Winter 2017

Data Analysis and Modeling to Optimize Thermal Treatment Cost and Performance

Jack C. Parker  
*University of Tennessee, jpark@utk.edu*

Ungtae Kim  
*Cleveland State University, u.kim@csuohio.edu*

Alyson Fortune  
*TerraTherm*

Steffen Griepke  
*TerraTherm*

James P. Galligan  
*TerraTherm*

Publisher’s Statement


Original Citation


How does access to this work benefit you? Let us know!
Authors
Jack C. Parker, Ungtae Kim, Alyson Fortune, Steffen Griepke, James P. Galligan, and Amber Bonarrigo
Data Analysis and Modeling to Optimize Thermal Treatment Cost and Performance

by Jack C. Parker, Ungtae Kim, Alyson Fortune, Steffen Griebke, James P. Galligan, and Amber Bonarrigo

Abstract

The objective of in situ thermal treatment is typically to reduce the contaminant mass or average soil concentration below a specified value. Evaluation of whether the objective has been met is usually made by averaging soil concentrations from a limited number of soil samples. Results from several field sites indicate large performance uncertainty using this approach, even when the number of samples is large. We propose a method to estimate average soil concentration by fitting a log normal probability model to thermal mass recovery data. A statistical approach is presented for making termination decisions from mass recovery data, soil sample data, or both for an entire treatment volume or for subregions that explicitly considers estimation uncertainty which is coupled to a stochastic optimization algorithm to identify monitoring strategies to meet objectives with minimum expected cost. Early termination of heating in regions that reach cleanup targets sooner enables operating costs to be reduced while ensuring a high likelihood of meeting remediation objectives. Results for an example problem demonstrate that significant performance improvement and cost reductions can be achieved using this approach.

Introduction

In situ thermal cleanup technologies have largely evolved from methods developed for enhanced oil recovery applications (Schumacher 1980; U.S. EPA 2004; Kingston et al. 2010). These technologies have been used to treat a wide range of volatile organic chemicals (VOCs) including chlorinated solvents, nonchlorinated volatile organic compounds, petroleum hydrocarbons, and semivolatile organic compounds (Vinegar et al. 1999; Beyke and Fleming 2005; Truex et al. 2007). Thermal methods are often used to treat contaminant source regions containing low solubility nonaqueous phase liquids (NAPLS), which are recalcitrant to many other methods. Thermal technologies utilize heat to enhance the removal of contaminants from the subsurface primarily by increasing the contaminant vapor pressure for VOCs and by decreasing the viscosity of low volatility NAPL. Heating may also enhance contaminant removal by increasing aqueous solubility, aqueous and vapor phase diffusion coefficients, and/or biotic and abiotic decay rates. Depending on the operating temperature and contaminant properties, heating may also decrease soil-water sorption coefficients and/or NAPL interfacial tension and liquid viscosities.

In situ thermal technologies in common commercial use include thermal conductive heating (TCH) that employs heating elements in wells to heat soil primarily by thermal conduction (Fan and Udell 1995; Hansen et al. 1998), electrical resistive heating (ERH) involving application of electrical current to an electrode network to heat soil by resistive energy dissipation (Heron et al. 1998a, 1998b; Beyke and Fleming 2005; Powell et al. 2007), and steam enhanced extraction (SEE) which heats the aquifer by steam injected into a network of injection wells with vapor and liquid phase recovery from multiphase extraction wells (Wu 1977; Itamura and Udell 1993; Davis 1997, 1998). Mass recovery rates during thermal treatment can be measured by monitoring mass recovery rates in extracted vapor and liquid phases as applicable to the various technologies.

The total energy input required to achieve a specified cleanup objective is strongly dependent on the boiling point of the contaminant(s) of concern, which depends on the chemical composition of NAPL (if present) and boiling point(s) of the pure contaminant(s). Chlorinated solvents and other chemicals can exhibit heterogeneous azeotropic behavior in which the boiling point of a NAPL-water mixture is less than the boiling point of solvent or water alone (Ginellling and Onken 1997; U.S. EPA 2004; Ponton 2009). The heterogeneous mixture boiling point may be estimated from Dalton’s law of partial pressures as the temperature when solvent vapor pressure plus water vapor pressure equals the ambient (atmospheric plus hydrostatic) pressure. After the coking point is reached, mole fractions of water and solvent in liquid and vapor phases will remain constant (in the absence of mass transfer limitations) until the NAPL phase is depleted. Azeotropic boiling points and mole fractions for several chlorinated solvents are shown in Table 1. For azeotropic systems with a boiling point less than 100°C, the water-NAPL system boils first when the coking point is reached. After NAPL is boiled off, dissolved and adsorbed solvent will continue to volatilize as the temperature gradually increases to
the aqueous phase boiling point. Thus, most solvent may be volatilized well before water reaches a full boil (contingent on spatial variability in temperature and contaminant distribution), which may substantially reduce the energy requirements in the absence of coboiling behavior. However, if very low residual soil concentrations are targeted, heating above the coboiling point to the free water boiling point, and/or holding the system at the target temperature longer may be necessary.

After a specific technology suited to site conditions is selected, heat balance calculations are typically performed considering treatment zone (TZ) geometry, well configurations, subsurface heat transfer characteristics (heat capacity, thermal conductivity, and advection rates), and cost and remediation time trade-offs associated with well spacing, heating rates, and energy required to reach the target temperature. Operational monitoring is employed to determine when remediation objectives have been met and operation can be terminated. The most common criteria for thermal system termination are maintaining a specified aquifer temperature for a defined period or reducing the average contaminant soil concentration below a target level. We will focus on the latter because knowledge of contaminant mass remaining after thermal treatment is an important variable to predict effects of thermal treatment on downgradient dissolved plume attenuation (Rao et al. 2001). Thermal system monitoring also commonly utilizes measurements of contaminant mass recovery rates based on fluid flow rates and concentrations from recovery wells.

The design and operation of all in situ remediation systems is complicated by the high degree of spatial and temporal variability inherent to geologic systems and by the hard reality that characterization of this variability is difficult and costly. It is imperative to come to terms with uncertainty and to design and operate systems with due consideration of uncertainty so that desired outcomes can be achieved with acceptable probability. To manage noisy data, Levine (2010) proposed comparing the upper confidence limit (UCL) of a moving aquifer average of aquifer concentration at a specified probability level with the compliance concentration as a criterion for regulatory closure.

Remediation system design is to a great degree a problem of managing uncertainty. Conventional approaches based on best estimates of system properties have a potential likelihood of failure to meet remediation targets, to overshoot budgets or both. Considerable work has been performed on optimization of long-term monitoring to trade-off costs against the value of information (Loaiciga et al. 1992; Reed et al. 2000; U.S. EPA 2000; U.S. EPA 2007). Stochastic optimization methods employ Monte Carlo models to define probability distributions of remediation performance and cost for a given design and use optimization algorithms to determine design variables that minimize probability-weighted cost subject to performance constraints (Cardiff et al. 2010; Parker et al. 2011).

In this paper, we will evaluate the limitations of soil concentration measurements for assessing thermal system performance and making termination decisions. We propose an alternative method to estimate average soil concentration from mass recovery data that is more reliable and less costly than soil sampling and present a robust statistical method for pooling data at various scales to make termination decisions with a consistent level of certainty. Coupled with a stochastic cost optimization method to identify optimal performance monitoring strategies, significant cost savings and performance improvements are demonstrated.

### Thermal Mass Recovery Model

Our objective is to develop a practical empirical model to estimate average soil concentration of contaminant within a defined treatment volume from field measurements of cumulative mass recovery over time. The approach will be applicable to most thermal systems, with the exception of those involving high temperature thermal decomposition or chemical destruction by oxidation, hydrolysis, or other mechanisms. For SEE, estimation of contaminant mass recovery rates over time requires monitoring and analysis of total fluids recovered from multiphase extraction wells. For TCH and ERH, periodic measurements of gas concentrations and flow rates from vapor extraction wells are required, although mass recovery in extracted liquids may be necessary if groundwater extraction is performed to control advective heat losses or to maintain hydraulic control. Total volatiles in extracted gas can be monitored economically using photoionization detector and flame ion-
ization detector sensors with occasional sample analyzes, using more accurate gas chromatography units to provide speciation information to calibrate sensor data. Time integration of mass recovery rates provides an accurate and cost effective means of determining cumulative mass recovery. Although commonly used to monitor integrated recovery rates over entire thermal systems, subsets of recovery wells or even individual recovery wells may be monitored separately to measure mass recovery from different areas within a site at relatively low cost compared to soil sampling.

Thermal mass recovery vs. time is often approximated using a normal probability distribution model with a mean equal to the time recovery rate is at a maximum ($\Delta t_{\text{peak}}$), which corresponds closely to the mixture boiling point. However, since the normal distribution model is symmetrical about the mean while elapsed time cannot be negative, a normal distribution that fits data prior to $\Delta t_{\text{peak}}$ must predict essentially 100% recovery within $2 \times \Delta t_{\text{peak}}$. For many sites, this will significantly underestimate treatment duration. To describe recovery curves more accurately, we propose to use a lognormal cumulative distribution function (CDF) in the form

$$M(t) = M_o N\left(\ln(t); \ln(\Delta t_{\text{peak}}), S_{\text{therm}}\right) + M'_{\text{ext}} t$$

where $M(t)$ is the cumulative mass recovered from the treated region after operating duration $t$, $M_o$ is the initial mass in the monitored volume, $\Delta t_{\text{peak}}$ is the time to reach maximum recovery rate. $N[x;m,S]$ is the normal CDF of $x$ with mean $m$ and standard deviation $S$. $S_{\text{therm}}$ is the standard deviation of the lognormal distribution, and $M'_{\text{ext}}$ is a steady-state mass inflow rate to the monitored zone from adjacent soil, for example, due to groundwater flow (a summary of abbreviations and symbols used in this paper is given in Appendix S4, Supporting information). For small $S_{\text{therm}}$ values (<0.2), the lognormal model closely approximates a normal distribution, but exhibits increasingly positive skewness as $S_{\text{therm}}$ increases. Mass recovery rate, $M'(t)$, can be computed by differentiating Equation 1. Making use of the normal probability density function (PDF), $N[x;m,S] = dN[x;m,S]/dx$, yields

$$M'(t) = \frac{dM(t)}{dt} = M_o \frac{t}{\Delta t_{\text{peak}}} N\left(\ln(t); \ln(\Delta t_{\text{peak}}), S_{\text{therm}}\right) + M'_{\text{ext}}.$$ (2)

Normalized cumulative mass recovery and recovery rate vs. time curves for the above model with $M'_{\text{ext}} = 0$ are illustrated in Figure 1 for a range of $S_{\text{therm}}$ values. As $S_{\text{therm}}$ increases, the curves become increasingly skewed to larger times.

The cumulative mass recovered and average soil concentration at a given time may be computed from Equation 1 assuming no net accumulation due to $M'_{\text{ext}}$ (i.e., $M'_{\text{ext}}$ is truly steady-state) as

$$M_{\text{rem}}(t) = M_o - M(t)$$

$$C_{\text{soil}}(t) = \frac{M_{\text{rem}}(t)}{\rho_{\text{soil}}V_{\text{soil}}}$$ (3)

where $M_{\text{rem}}(t)$ is mass remaining at time $t$, $\rho_{\text{soil}}$ is soil bulk density, and $V_{\text{soil}}$ is the treated soil volume. The rate of change of average soil concentration may be derived from Equation 3 as

$$\frac{dC_{\text{avg}}(t)}{dt} = -\frac{M'(t)}{\rho_{\text{soil}}V_{\text{soil}}}.\quad (4)$$

An alternative method of estimating soil concentration from recovery data, independent of the lognormal model, is to extrapolate the current recovery rate forward in time assuming a constant rate reduction factor (i.e., second derivative of ln recovery curve) as

$$C_{\text{avg}}(t) = -\int_0^t \frac{M'(t)(1+M'_{\text{in}})\tau}{\rho_{\text{soil}}V_{\text{soil}}} d\tau = -\frac{M'(t)}{\rho_{\text{soil}}V_{\text{soil}} \ln(1+M'_{\text{in}})}.$$ (5)

where $M'(t)$ is the current observed mass recovery rate, $M'_{\text{in}} = d(\ln M)/dt$ is the rate reduction factor, an $\tau$ is a dummy integration variable. We will address the magnitude and variability of $M'_{\text{in}}$ later.

**Field Verification of Model**

Data from eight thermal treatment sites (Table 2) were used to evaluate the lognormal recovery model. Perchloroethene (PCE) was the primary contaminant at all sites except Site 6 for which TCE was the main contaminant. Nonlinear regressions were performed to fit lognormal model parameters to cumulative mass recovery and recovery rate data for the field sites (calibration method details are given in Appendix S1). Owing to much greater “noise” in rate measurements, $w_{\text{max}}$ values computed from cumulative
## Table 2

Model Calibration Results for Initial Calibrations ($w_{mass} = 1$) and Final Calibrations ($w_{mass} = 0.5$) for Eight Field Sites

<table>
<thead>
<tr>
<th>Site ID</th>
<th>Thermal Method</th>
<th>Treatment Zone Volume (m$^3$)</th>
<th>Actual Mass Recovery (kg)</th>
<th>$t_{start}$ (d)</th>
<th>$t_{stop}$ (d)</th>
<th>$M_0$ (kg)</th>
<th>$\Delta t_{peak}$ (d)</th>
<th>$S_{term}$</th>
<th>$S_{ln_mass}$</th>
<th>$S_{ln_rate}$</th>
<th>Initial Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TCH</td>
<td>11,100</td>
<td>2353.4</td>
<td>0</td>
<td>210</td>
<td>2353</td>
<td>73</td>
<td>0.25</td>
<td>0.011</td>
<td>1.692</td>
<td>2353</td>
</tr>
<tr>
<td>2</td>
<td>TCH</td>
<td>6039</td>
<td>5248.5</td>
<td>0</td>
<td>147</td>
<td>5248</td>
<td>55</td>
<td>0.40</td>
<td>0.024</td>
<td>0.710</td>
<td>5390</td>
</tr>
<tr>
<td>3</td>
<td>TCH</td>
<td>78,000</td>
<td>75,331.7</td>
<td>0</td>
<td>244</td>
<td>75332</td>
<td>118</td>
<td>0.20</td>
<td>0.039</td>
<td>1.079</td>
<td>75332</td>
</tr>
<tr>
<td>4</td>
<td>TCH</td>
<td>2523</td>
<td>3400.6</td>
<td>0</td>
<td>100</td>
<td>3408</td>
<td>42</td>
<td>0.25</td>
<td>0.023</td>
<td>0.909</td>
<td>3401</td>
</tr>
<tr>
<td>4a</td>
<td>a</td>
<td>0</td>
<td>100</td>
<td>73</td>
<td>16</td>
<td>0.40</td>
<td>0.012</td>
<td>0.645</td>
<td>800</td>
<td>18</td>
<td>0.039</td>
</tr>
<tr>
<td>4b</td>
<td>a</td>
<td>10</td>
<td>100</td>
<td>1183</td>
<td>26</td>
<td>0.20</td>
<td>b</td>
<td>b</td>
<td>1209</td>
<td>24</td>
<td>b</td>
</tr>
<tr>
<td>4c</td>
<td>a</td>
<td>24</td>
<td>100</td>
<td>1478</td>
<td>27</td>
<td>0.22</td>
<td>b</td>
<td>b</td>
<td>1400</td>
<td>25</td>
<td>b</td>
</tr>
<tr>
<td>5</td>
<td>TCH</td>
<td>1180</td>
<td>349.2</td>
<td>0</td>
<td>76</td>
<td>356</td>
<td>21</td>
<td>0.63</td>
<td>0.019</td>
<td>0.412</td>
<td>356</td>
</tr>
<tr>
<td>5a</td>
<td>a</td>
<td>0</td>
<td>76</td>
<td>95</td>
<td>8</td>
<td>0.90</td>
<td>0.023</td>
<td>0.356</td>
<td>95</td>
<td>8</td>
<td>0.023</td>
</tr>
<tr>
<td>5b</td>
<td>a</td>
<td>17</td>
<td>76</td>
<td>270</td>
<td>8</td>
<td>0.70</td>
<td>b</td>
<td>b</td>
<td>205</td>
<td>7</td>
<td>0.70</td>
</tr>
<tr>
<td>5c</td>
<td>a</td>
<td>33</td>
<td>76</td>
<td>205</td>
<td>25</td>
<td>0.99</td>
<td>b</td>
<td>b</td>
<td>71</td>
<td>25</td>
<td>0.99</td>
</tr>
<tr>
<td>6</td>
<td>ERH</td>
<td>13,340</td>
<td>528.1</td>
<td>0</td>
<td>200</td>
<td>570</td>
<td>78</td>
<td>0.65</td>
<td>0.035</td>
<td>0.504</td>
<td>568</td>
</tr>
<tr>
<td>6a</td>
<td>a</td>
<td>0</td>
<td>200</td>
<td>100</td>
<td>23</td>
<td>0.80</td>
<td>0.021</td>
<td>0.505</td>
<td>95</td>
<td>23</td>
<td>0.60</td>
</tr>
<tr>
<td>6b</td>
<td>a</td>
<td>23</td>
<td>200</td>
<td>280</td>
<td>40</td>
<td>0.63</td>
<td>b</td>
<td>b</td>
<td>280</td>
<td>40</td>
<td>0.63</td>
</tr>
<tr>
<td>6c</td>
<td>a</td>
<td>83</td>
<td>200</td>
<td>155</td>
<td>40</td>
<td>0.70</td>
<td>b</td>
<td>b</td>
<td>160</td>
<td>40</td>
<td>0.70</td>
</tr>
<tr>
<td>7</td>
<td>TCH</td>
<td>10,703</td>
<td>1250.6</td>
<td>0</td>
<td>108</td>
<td>1270</td>
<td>23</td>
<td>0.80</td>
<td>0.083</td>
<td>0.813</td>
<td>1341</td>
</tr>
<tr>
<td>7a</td>
<td>a</td>
<td>0</td>
<td>108</td>
<td>250</td>
<td>4</td>
<td>0.61</td>
<td>0.014</td>
<td>0.566</td>
<td>250</td>
<td>4</td>
<td>0.50</td>
</tr>
<tr>
<td>7b</td>
<td>a</td>
<td>7</td>
<td>60</td>
<td>835</td>
<td>17</td>
<td>0.37</td>
<td>b</td>
<td>b</td>
<td>835</td>
<td>17</td>
<td>0.44</td>
</tr>
<tr>
<td>7c</td>
<td>a</td>
<td>54</td>
<td>108</td>
<td>201</td>
<td>22</td>
<td>1.02</td>
<td>b</td>
<td>b</td>
<td>195</td>
<td>22</td>
<td>0.87</td>
</tr>
<tr>
<td>8</td>
<td>SEE</td>
<td>5248</td>
<td>10,959.0</td>
<td>0</td>
<td>450</td>
<td>12,300</td>
<td>155</td>
<td>0.60</td>
<td>0.068</td>
<td>0.529</td>
<td>12,500</td>
</tr>
<tr>
<td>8a</td>
<td>a</td>
<td>0</td>
<td>450</td>
<td>7100</td>
<td>105</td>
<td>0.45</td>
<td>0.021</td>
<td>0.386</td>
<td>7100</td>
<td>100</td>
<td>0.40</td>
</tr>
<tr>
<td>8b</td>
<td>a</td>
<td>110</td>
<td>305</td>
<td>2800</td>
<td>100</td>
<td>0.30</td>
<td>b</td>
<td>b</td>
<td>2700</td>
<td>105</td>
<td>0.25</td>
</tr>
<tr>
<td>8c</td>
<td>a</td>
<td>190</td>
<td>305</td>
<td>1600</td>
<td>81</td>
<td>0.20</td>
<td>b</td>
<td>b</td>
<td>1600</td>
<td>100</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Note: Site identification numbers (IDs) with no letter are results for a single inflection model fit to the site data. Site IDs with letters denote parameters for multi-inflection model fit to data.

$a = \text{Total treatment zone volume for multiple "event" model is the same as that shown for single "event" model.}$

$b = S_{ln_{mass}}$ and $S_{ln_{rate}}$ values for multiple "event" model are aggregate values for the combined event model.
recovery and rate data variances were generally greater than 0.99. Given the low weight applicable to rate data and the observation that rate data, even with low weights, tended to make convergence of the nonlinear regression more difficult, we used \( w_{\text{mass}} = 1 \) for initial calibrations.

Initial calibrations with \( w_{\text{mass}} = 1 \) were performed by fitting \( M_r, \Delta t_{\text{peak}}, \) and \( S_{\text{therm}} \) to a single mass recovery function for the entire treatment duration for each field site. Fitted parameter values are summarized in Table 2 along with \( S_{\text{therm}} \) and \( S_{\text{therm}} \) vs. values computed using the calibrated parameters. The results (Figure 2, Table 2) show reasonably good agreement. Estimated \( S_{\text{therm}} \) values range from 0.2 to 0.8 for the various sites and appear to be uniformly distributed with an equal number of sites having \( S_{\text{therm}} \) values above and below 0.5.

The lognormal model predicts thermal treatment duration (\( \Delta t_{\text{rem}} \)) to be a function of \( S_{\text{therm}}, \Delta t_{\text{peak}}, M_r, \) and target mass remaining \( M_r \). Figure 3 illustrates this relationship normalized as \( \Delta t_{\text{rem}}/\Delta t_{\text{peak}} \). To obtain a mass reduction of 99-99.99% (i.e., \( M_r/M_0 \) from 0.01 to 0.0001), \( \Delta t_{\text{rem}}/\Delta t_{\text{peak}} \) ranges from approximately 1.6 to 2.1 for \( S_{\text{therm}} = 0.2 \), from 2.5 to 4.4 for \( S_{\text{therm}} = 0.4 \), from 4.0 to 9.3 for \( S_{\text{therm}} = 0.6 \), and from 6.4 to 19.6 for \( S_{\text{therm}} = 0.8 \). Note that \( M_r/M_0 \) is equivalent to the ratio of target average soil concentration to initial average soil concentration. The results emphasize the strong dependence of thermal treatment duration on \( S_{\text{therm}} \) as well as initial and target soil concentrations.

Of the sites studied, half had \( S_{\text{therm}} \) values less than 0.5 with actual treatment termination between 2 to 3 times \( \Delta t_{\text{peak}} \) at model-estimated mass recovery ratios averaging 99.8%. The other half of the sites, with \( S_{\text{therm}} \) values greater than 0.5, exhibited marked positive skew in the recovery curves and terminated at model-estimated mass recovery ratios averaging only 95% after operating for 2 to 4.5 times \( \Delta t_{\text{peak}} \). The results demonstrate the relative difficulty of achieving high recovery ratios at sites with high \( S_{\text{therm}} \) values.

Closer inspection of observed and simulated mass recovery curves indicates multiple inflection points occur for many sites, reflecting multiple recovery rate peaks. This may be attributed to operations initiated on different TZs at various times, for example, due to incremental start-up of component systems including pilot tests, to variations in well spacings or geologic properties that affect heat or mass transfer rates, or to spatial variability in cocontaminants that affect boiling point. The behavior of multi-infection sites may be modeled by superposition of multiple lognormal recovery events—that is, by modeling each “event” by a different parameterization of Equation 1 and then summing results for all events at each date (i.e., after converting from operating time to calendar time) to obtain site-wide mass recovery curves, recovery rate curves, and average soil concentrations. For each added event, two additional model parameters are introduced to align calendar and operating times, namely the stop time for the prior event \( t_{\text{stop}} \) and the start time for the following event \( t_{\text{start}} \). We take \( t_{\text{stop}} \) and \( t_{\text{start}} \) values relative to the first start time (designated as zero), while \( \Delta t_{\text{peak}} \) values are given relative to \( t_{\text{start}} \) for each separate event. Values for these parameters may be operationally known. If not, they may be estimated during the calibration process.

Initial calibrations with \( w_{\text{mass}} = 1 \) were repeated using multilognormal models for five sites that exhibited multiple inflections. Results for these calibrations are given in Table 2 with letters following the site number to identify each lognormal event (e.g., 4a, 4b, and 4c). A graphical illustration of cumulative mass recovery and recovery rate curves vs. time for single- and multi-infection calibrations is given for Site 7 (Figure 4).

Initial calibration results reveal a tendency for calibrated parameters to underpredict recovery rates at late times. Because small deviations in recovery curve tails may result in significant errors in inferred final average soil concentrations, a final round of calibrations was performed to refine asymptotic tail behavior. Only data from the last 20 to 30 d of recovery operations were used in these regressions and more weight was given to recovery rate data using \( w_{\text{mass}} = 0.5 \). This procedure resulted in small increases in \( S_{\text{therm}} \) values (Table 2). For single-event calibrations, \( S_{\text{therm}} \) values ranged from 0.2 to 0.8 for initial calibrations and 0.24 to 0.93 for final calibrations across all sites. For multi-event calibrations, the range in average \( S_{\text{therm}} \) values between sites was from 0.27 to 0.86 for initial calibrations and from 0.30 to 0.86 for final calibrations.
the difference between maximum and minimum $S_{\text{therm}}$ values for the same site varied from a relatively narrow 0.30 to a high of 0.86, indicating that while $S_{\text{therm}}$ values for different areas within a site are likely to be less variable than differences between sites, differences within sites may sometimes be as great as differences between sites.

Uncertainty in estimated mass remaining using the recovery model depends on uncertainty in the computed mass recovered as well as the asymptotic recovery, $M_{\alpha}$. Standard deviations in $\text{ln}$ model-predicted mass, $S_{\text{ln mass}}$, and $\text{ln}$ recovery rate, $S_{\text{ln rate}}$, were computed from deviations between $\text{ln}$ predicted and $\text{ln}$ measured quantities and are tabulated for each calibration in Table 2. As an overall measure of model uncertainty, the root mean square error (RMSE) in $\text{ln}$ cumulative mass recovery and $\text{ln}$ recovery rate was computed as

$$S_{\text{ln rmse}} = \left( \frac{S_{\text{ln mass}}^2 + S_{\text{ln rate}}^2}{2} \right)^{1/2}.$$  

(6)

Average values of $S_{\text{ln mass}}$, $S_{\text{ln rate}}$, and $S_{\text{ln rmse}}$ are summarized in Table 3 for initial and final calibrations of sites with single inflection points, for multi-inflection sites modeled with a single inflection model and for multi-inflection sites modeled with a multi-inflection model. Not surprisingly, the multi-inflection sites exhibited lower uncertainty when a multi-inflection model was used. Final calibrations that used rate as well as mass recovery data to refine the initial calibrations reduced uncertainty for all cases. We recommend using the two-step calibration method as it is more robust than attempting to calibrate in a single step with rate and mass data. The second calibration using mass and rate data produced a small increase in $S_{\text{ln mass}}$ and a larger decrease in $S_{\text{ln rate}}$ yielding a net decrease in $S_{\text{ln rmse}}$. Curiously, single inflection sites had significantly higher error than multi-inflection sites calibrated with a single inflection model. We attribute this to the small sample size and unique features of the individual sites. Across all formulations, average $S_{\text{ln rmse}}$ is about 0.50 with mass data only and 0.41 after calibration refinement using rate and mass data.

### Soil Monitoring Data Accuracy and Uncertainty

Three of the study sites (1, 4, and 5) had soil concentration measurements before and after thermal treatment that can be used to estimate initial and final contaminant mass and its uncertainty for comparison with estimates from mass recovery data. Prior to undertaking this comparison, we wish to consider the estimation of mean concentration and its confidence limits from soil concentration data, as there are many factors that must be considered to avoid, or at least limit, errors.

The following methods were used to compute two-sided 95% confidence limits for initial (preremediation) and final (postremediation) average soil concentration.

Method 1—Normal distribution model. This approach is well-known and is often used by remediation contractors. Confidence limits are computed as

$$m_{\pm CL} = m \pm t(\alpha, n_{\text{sample}}) \frac{S}{\sqrt{n_{\text{sample}}}}$$

(7)

where $m=\text{average}(x_1, \ldots, x_{n_{\text{sample}}})$ and $S=\text{stddev}(x_1, \ldots, x_{n_{\text{sample}}})$ are the arithmetic average and standard deviation of soil concentration measurements on $n_{\text{sample}}$ soil samples, $m_{\pm CL} = m_{\pm CL, \alpha}$ is the lower confidence limit (LCL) of the mean value

### Table 3

<table>
<thead>
<tr>
<th>Calibration case</th>
<th>Initial Calibration</th>
<th>Final Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_{\text{ln mass}}$</td>
<td>$S_{\text{ln rate}}$</td>
</tr>
<tr>
<td>a. Single inflection sites</td>
<td>0.025</td>
<td>1.160</td>
</tr>
<tr>
<td>b. Multi-inflection site treated as single</td>
<td>0.046</td>
<td>0.633</td>
</tr>
<tr>
<td>c. Multi-inflection site treated as multi</td>
<td>0.018</td>
<td>0.492</td>
</tr>
<tr>
<td>d. Best site model (cases a and c)</td>
<td>0.021</td>
<td>0.742</td>
</tr>
</tbody>
</table>
computed with a negative sign on the right hand side. \( m_{\text{UCL}, \text{f}} = m_{\text{LCL}, \text{f}} \) is the UCL computed with a positive sign, and \( t(\alpha, n) \) is the two-sided \( t \)-value for probability level \( \alpha \).

Method 2—Lognormal distribution model. This approach accommodates the asymmetric nature of positively skewed high variance populations. Confidence limits of the arithmetic mean concentration are computed from the lognormal model as

\[
m_{\pm\text{CL}, \alpha} = \exp \left( \ln m + t(\alpha, n) \frac{S_{\ln}}{\sqrt{n}} \right)
\]

where \( m \) is the arithmetic average concentration computed as \( m = \exp(m_{\ln} + 0.5S_{\ln}) \) where \( m_{\ln} = \text{average}(\ln x_1, \ldots, \ln x_{\text{amp}}) \) and \( S_{\ln} = \text{std}(\ln x_1, \ldots, \ln x_{\text{amp}}) \).

Method 3—Alternate lognormal model. Although the lognormal model is a more realistic approximation of high variance populations, estimates of the arithmetic mean from lognormal model parameters can be sensitive to deviations from the lognormal model especially in the tail (Reimann and Filzmoser 2000) or to truncation of nondetects (Helsel 2010). Method 3 uses Equation 8, but instead of computing the arithmetic mean \( m \) from lognormal parameters, it is computed as in Method 1.

Omitting nondetect values from statistical calculations can result in overestimation of the sample mean and underestimation of variance (Helsel 2010). Including nondetects for statistical calculations with values set at the detection limits will somewhat attenuate errors in sample means, but may do little to attenuate underestimation of variance. Setting nondetects below the detection limit will reduce lognormal statistics errors for analyses involving single contaminants as in the present case. Normal distribution means (arithmetic averages) and standard deviations are less sensitive to the treatment of nondetects than lognormal statistics.

Analyses of synthetic lognormal datasets with an \( S_{\ln} \) of about 2.5 (typical for field sites) were performed to evaluate handling of samples below detection limits by assigning nondetects a numerical value equal to the detection limit times a factor \( F \). For datasets with about 20% nondetects, arithmetic averages were insensitive to \( F \), while \( S_{\ln} \) values were most accurately estimated using \( F \) values between 0.1 and 0.5 (0.2 was optimal). With 40% nondetects, average values remained insensitive, while the sensitivity of \( S_{\ln} \) values to \( F \) increased. The most accurate results were obtained using \( F=0.1 \) (\( S_{\ln} \) was underestimated by 25% using \( F=0.5 \)). For greater than 50% nondetects, it was not possible to obtain accurate averages and \( S_{\ln} \) values using a single \( F \) value. With 80% nondetects, the best compromise was obtained using \( F=0.001 \), which underestimated the average value and overestimated \( S_{\ln} \), but yielded similar 95% confidence limits. For the present study, \( F=0.1 \) was used to assign numerical values to all reported nondetects.

Average soil concentrations from soil sampling rounds prior to thermal treatment for the three sites were computed by each of the above methods and multiplied by estimates of dry soil mass within the treatment volumes to determine total pre-remediation contaminant mass \( (M) \) and their confidence limits. Confidence limits of pre-remediation mass estimates were also computed from the \( M(t) \) model for comparison with soil sample-based values by

\[
M_{\pm\text{CL}, \alpha} = \exp(M_{\ln} + t(\alpha, \infty)S_{\ln, \text{rms}})
\]

where \( M_{\pm\text{CL}, \alpha} = M_{\pm\text{CL}, \alpha} \) is the LCL of the mean value computed with a negative sign on the right-hand side, \( M_{\pm\text{UCL}, \alpha} = M_{\pm\text{UCL}, \alpha} \) is the UCL computed with a positive sign, \( M_{\ln} \) is the calibrated pre-remediation mass summed across all TZs, and \( S_{\ln, \text{rms}} \) is the RMSE for the final single inflection model calibration for Site 1 and the final multi-inflection model calibrations for Sites 4 and 5.

Estimates of \( M_i \) and its confidence limits for the above methods are tabulated in Table 4 along with measured total mass recovered during thermal treatment. Not surprisingly, the \( M(t) \) model yields estimates of pre-remediation mass with narrow confidence limits consistent with actual

### Table 4

<table>
<thead>
<tr>
<th>Site</th>
<th>No. Samples</th>
<th>Parameter</th>
<th>Estimated Mass in Treatment Zone Prior to Remediation (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>78</td>
<td>Mean value or best estimate</td>
<td>1645</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% LCL of mean</td>
<td>1004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% UCL of mean</td>
<td>2286</td>
</tr>
<tr>
<td>4</td>
<td>124</td>
<td>Mean value or best estimate</td>
<td>1616</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% LCL of mean</td>
<td>463</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% UCL of mean</td>
<td>2769</td>
</tr>
<tr>
<td>5</td>
<td>46</td>
<td>Mean value or best estimate</td>
<td>282</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% LCL of mean</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% UCL of mean</td>
<td>561</td>
</tr>
</tbody>
</table>
recovery data from which model results are derived. Best estimates of contaminant mass using soil concentration data with Method 1 consistently underestimate actual recovery. Upper and lower confidence limits appear to have a downward bias compared to actual recovery, as expected for data with positive skew. Method 2 best estimates consistently overpredict actual recovery by 6-45%. Lower and upper confidence limits appear to be biased high. The LCL for Site 1 using Method 2 exceeds actual recovery by a factor of nearly 3. Method 3 shows the least erratic behavior with confidence limits that bracket actual recovery.

The results suggest that the soil data exhibit greater positive skew than the lognormal model accommodates, resulting in an inconsistency between the actual arithmetic mean soil concentration and the mean inferred from lognormal model parameters. Method 3 largely avoids this discrepancy by using the actual arithmetic mean. Some of the observed differences in mass estimates may be due to averaging data with equal weights for all data points. Geostatistical methods might reduce such errors, although it was not possible for the datasets considered, because sample coordinate information was not available. It is also possible that the selection of sampling locations itself was biased. For example, if sampling focused on identification of “hot spots” upward bias would likely occur. Unintended bias may be avoided using pseudo-random sampling algorithms (ITRC 2012).

Postremediation average soil concentrations and confidence limits are tabulated in Table 5 for the same sites based on the three soil data analysis methods, the M(t) model (Equation 3), and the rate extrapolation method (Equation 5). Calculations for the rate extrapolation method were performed using moving averages of measured recovery rates to attenuate noise. Values of \( M_{\text{est}} \) in Equation 5 at termination dates ranged from about -0.5 to -0.01, which were used to estimate a range of \( C_{\text{est}} \) roughly interpreted as 95% confidence limits at the time thermal treatment ceased.

Although the \( M(t) \) model and recovery rate extrapolation method are based on the same underlying data, the assumptions and computational approaches are very different. Reasonable agreement between the \( M(t) \) model and the recovery rate extrapolation method support the validity and accuracy of both methods. The difference between the initial and final calibration results indicates that model refinement to weight late time mass and recovery rates is important to obtain accurate estimates of mass remaining.

Method 1 confidence limits are clearly unreliable as all three cases show physically impossible negative LCL. Method 2 results are erratic. The Method 2 LCL for Site 1 exceeds the value based on mass recovery, while the Site 5 UCL is less than the concentration estimated from mass recovery methods. Site 4 confidence limits using Method 2 are far above the value inferred from mass recovery data. Method 3 results do not appear to be much better.

All three methods of analyzing soil concentration data yield final average soil concentrations that are at least an order of magnitude lower than those obtained from mass recovery data for Site 5. It is tempting to conjecture that the mass recovery methods are overestimating the final average concentration for Site 5 rather than the converse. However, a quick look at the mass recovery curve for Site 5 (Figure 2) reveals that the curve was still climbing rather steeply at the time treatment was terminated and would likely have taken another several weeks to flatten. The measured recovery rate on the last day of operation for Site 5 was just under 0.5 kg PCE per day, suggesting on the order of 5 kg of PCE was remaining in the system at termination. Dividing this by the estimated soil mass in the TZ (9.5 × 10^7 kg) indicates the average soil concentration was about 5 mg/kg when terminated, which is far greater than the average concentration of 0.067 mg/kg estimated from soil data by Method 3.

Surprisingly, the reliability of average soil concentration results shows no evident relationship with the number of soil samples. While confidence interval widths increase substantially from Site 5, which had 85 samples, to Site 4 with only 14 samples, the most egregiously erroneous best estimate occurred for Site 5, which had the most samples. Results based on soil concentration data indicate that while numerical precision improves with more samples, the accuracy does not necessarily converge to full-scale reality. For

<table>
<thead>
<tr>
<th>Site</th>
<th>No. Samples</th>
<th>Parameter</th>
<th>Soil Data</th>
<th>M(t) Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Initial Calibration</td>
<td>Final Calibration</td>
</tr>
<tr>
<td>1</td>
<td>58</td>
<td>Mean value or best estimate</td>
<td>2.407</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% LCL of mean</td>
<td>-1.398</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% UCL of mean</td>
<td>6.212</td>
<td>0.045</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>Mean value or best estimate</td>
<td>2.600</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% LCL of mean</td>
<td>-0.657</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% UCL of mean</td>
<td>5.857</td>
<td>0.476</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>Mean value or best estimate</td>
<td>0.286</td>
<td>9.707</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% LCL of mean</td>
<td>-0.013</td>
<td>5.921</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% UCL of mean</td>
<td>0.584</td>
<td>15.915</td>
</tr>
</tbody>
</table>
confidence limits of Site 5 to bracket the average soil concentration estimated from mass recovery data. A residual
in error of about 2 would need to be added to \( S_{\log}^{1/2} \) in Equation 8. In most cases, a term of this magnitude would
be greater than \( S_{\log}^{1/2} \) even with a small number of samples. Perhaps Site 5 is an anomaly, but the large number of
centric results for the three sites (Sites 1, 4, and 5) do not engender confidence in the reliability of contaminant mass
estimates based on soil sample data.

The foregoing indicates that estimates of average soil concentration from soil sample data exhibit large uncer-
tainty. Uncertainty associated with estimates based on mass recovery data is generally lower, but still significant.
Furthermore, \( S_{\log} \) values exhibit significant a priori uncertainty which has a large effect on treatment duration to reach a
given cleanup objective. While \( S_{\log} \) values can be progressively refined by regression analyses as treatment pro-
gresses, uncertainty in treatment duration, and hence cost, associated with both measurement and model uncertainties
should be factored into treatment design, as discussed in the following sections.

Performance Monitoring Strategies and Termination Criteria

As stated previously, the objective we adopt for thermal treatment is the commonly used criteria that contaminant
mass in the source zone should be reduced below a value corresponding to a specified average soil concentration,
which may be stated as

\[
C_{\text{soil}}^{\text{avg}} \leq C_{\text{stop}}
\]  

(10)

where \( C_{\text{soil}}^{\text{avg}} \) is the arithmetic average soil concentration in the source zone and \( C_{\text{stop}} \) is a stipulated cleanup target. In
practice, difficulties arise in the application of Equation 10 because the true value of \( C_{\text{soil}}^{\text{avg}} \) is never known exactly. If we compute the average
value of soil concentration from a number of soil samples, \( C_{\text{soil}}^{\text{avg}} \) and substitute this value for the true \( C_{\text{soil}}^{\text{avg}} \) in Equation 10, there will be a substantial likelihood
of erroneously terminating treatment before the target criterion is met owing to deviations between \( C_{\text{soil}}^{\text{avg}} \) and \( C_{\text{soil}}^{\text{avg}} \).

A practical way to contend with this uncertainty is to employ the statistical termination criterion

\[
C_{\text{UCL}}^{\text{UCL}} \leq C_{\text{soil}}^{\text{avg}}
\]  

(11)

where \( C_{\text{UCL}}^{\text{UCL}} \) is the UCL of estimated average soil concentration at exceedance probability \( \alpha \) (i.e., significance
level). As \( C_{\text{UCL}}^{\text{UCL}} > C_{\text{soil}}^{\text{avg}} \) for any \( \alpha > 0.5 \), Equation 11 is a
more stringent stop criteria than Equation 10 when \( C_{\text{soil}}^{\text{avg}} \) is implicitly substituted for \( C_{\text{soil}}^{\text{avg}} \). The difference between
\( C_{\text{UCL}}^{\text{UCL}} \) and \( C_{\text{soil}}^{\text{avg}} \) is a safety factor to reduce the likelihood of an erroneous decision to terminate early.

High variance properties of quantities that are physically constrained to be non-negative, such as contaminant con-
centration, necessarily exhibit positively skewed distributions. Normal probability distributions cannot describe such
behavior. Lognormal distributions capture the major features of such data and are commonly used as a mathematically
expedient approximation. If the average concentration is estimated from \( n_{\text{soil}} \) soil samples, then

\[
C_{\text{UCL}}^{\text{UCL}} = \exp \left( \ln \left( C_{\text{soil}}^{\text{avg}} \right) + t_{(\alpha, N)} \right) = C_{\text{soil}}^{\text{avg}} \leq C_{\text{soil}}^{\text{avg}}
\]  

(12)

which yields the termination criteria in terms of \( C_{\text{soil}}^{\text{avg}} \) for \( n_{\text{soil}} \) soil samples

\[
C_{\text{soil}}^{\text{UCL}} - t_{(\alpha, N)} \sqrt{S_{\text{soil}}^{\text{avg}}/n_{\text{soil}}} \leq C_{\text{soil}}^{\text{avg}}
\]  

(13)

where \( C_{\text{soil}}^{\text{avg}} \) is the arithmetic average of \( n_{\text{soil}} \) samples, \( S_{\text{soil}}^{\text{avg}} \) is the standard deviation of \( \ln \) concentration values, and \( t_{(\alpha, N)} \) is the \( t \)-value for one-sided significance level \( \alpha \) with \( N \) degrees of freedom. If \( S_{\text{soil}}^{\text{avg}} \) is computed from the
first \( n_{\text{soil}} \) samples then \( N = n_{\text{soil}} - 1 \), while if \( S_{\text{soil}}^{\text{avg}} \) is based on prior
site characterization data or experience with other sites, then \( N = \infty \). Note that for \( \alpha = 0.05 \), which denotes the 95% upper
confidence limit (UCL), \( t \) is 1.64 with \( N = \infty \), while for \( \alpha = 0.5 \), \( t \) is 0, in which case \( C_{\text{UCL}}^{\text{UCL}} = C_{\text{soil}}^{\text{avg}} \) indicating an equal probability of over- or underestimating the average value.

Equations 12 and 13 are approximate because the actual probability distribution will not exactly follow a lognormal
curve (Reimann and Filzmoser 2000) and because they assume uncertainty in the arithmetic average concentration has
the same variance as the geometric mean.

As an example, consider a site with a cleanup target \( C_{\text{soil}}^{\text{UCL}} \) of 1 mg/kg. It is planned to take \( n_{\text{soil}} = 20 \) soil samples to assess whether the objective has been met. Assume a prior estimate of \( S_{\text{soil}}^{\text{avg}} = 2.9 \). If we want a 95% probability (\( \alpha = 0.05 \)) that the actual average soil concentration will be less than 1 mg/kg when we terminate treatment, then \( t = 1.64 \) and
Equation 12 indicates that the average concentration \( C_{\text{soil}}^{\text{avg}} \) computed from 20 samples needs to be less than
0.345 mg/kg to achieve the desired reliability. If we increase the number of samples to 50, we could terminate with the
same confidence when \( C_{\text{soil}}^{\text{avg}} \leq 0.510 \) mg/kg.

An alternative to termination based on soil sample measurements is to estimate average soil concentration from
cumulative mass recovery data as described by Equation 3. As mass recovery data is intrinsically integrated over a
defined bulk soil volume \( V_{\text{soil}} \), no sample averaging operation is required and Equation 13 can be modified as

\[
C_{\text{UCL}}^{\text{UCL}} = \exp \left( \ln \left( \frac{M_{\text{rem}}}{\rho_{\text{soil}} \cdot \rho_{\text{soil}}} \right) + t_{(\alpha, \infty)} S_{\text{ln}}^{\text{soil}} \right) \leq C_{\text{soil}}^{\text{avg}}
\]  

(14)

where \( M_{\text{rem}} = M(t) - M_{\text{soil}} \) is the best estimate of mass remaining based on recovery data, and \( S_{\text{ln}}^{\text{soil}} \) is the standard error of the
natural log of mass remaining.

An important aspect of the statistical stop criteria is that \( C_{\text{UCL}}^{\text{UCL}} \) decreases with decreasing measurement uncertainty
\( S_{\text{ln}}^{\text{soil}} \) for soil data and \( S_{\text{ln}}^{\text{rec}} \) for mass recovery data) and with an
increasing number of soil samples, which allows earlier ter-
minalation of heating at a given confidence level. As we have shown, \( S_{\text{ln}}^{\text{soil}} >> S_{\text{ln}}^{\text{rec}} \), which lends a significant advantage to
mass recovery data for termination decisions. For decisions based on soil data, increasing the number of samples reduces
uncertainty in principle (assuming unbiased sampling), which allows earlier termination and reduces heating costs. However,
this benefit must be balanced against higher sampling costs.
The foregoing statistical stop criteria may be applied to an entire thermal treatment volume to terminate operation of all heating units simultaneously, or alternatively since time to reach \( C_{\text{UCL,global}}^{\text{pool}} \) will vary spatially, it may be possible to reduce operating costs by applying stop criteria independently to smaller regions to terminate heating earlier in areas that reach cleanup objectives sooner. For example, anticipating that regions with higher initial contaminant concentrations are likely to take longer to cleanup, a system designer may consider dividing a site into multiple TZs based on ranges of preremediation soil concentrations determined from soil boring data. It may also be cost advantageous to further divide TZs into multiple monitoring zones (MZs). The total number of MZs may thus range from 1 to a value equal to the number of TZs per TZ summed over all TZs.

Trade-offs will arise when trying to optimize the number of TZs and MZs. Specifically, while more TZs and/or MZs offer the possibility of cost savings by terminating heating earlier for areas that cleanup more quickly, more total soil samples and/or more mass recovery measurements (and associated plumbing) will be needed to make reliable decisions at smaller scales. The potential savings from early termination may thus be offset by additional monitoring costs.

In addition to termination decisions at the MZ level (smallest decision level), we may concurrently apply termination rules at the TZ level to terminate all MZs within a TZ based on their aggregated data, or at the site level to terminate all TZs (and their MZs) based on aggregated data for the entire site. Note that since \( n_{\text{stop}} \) for an entire TZ is equal to the sum of \( n_{\text{stop}} \) values for its MZs, \( C_{\text{UCL,a}}^{\text{pool}} \) for a TZ based on site data will always be lower than that for the individual MZs. Therefore, it may be possible for an entire TZ to meet its aggregated termination criteria before the individual MZs, or likewise for site-wide criteria to be met before individual TZs.

The following decision logic is proposed for making termination decisions at various scales after each measurement/sampling event:

1. Tabulate and analyze the most recent sampling/monitoring data.
2. If \( C_{\text{UCL,global}}^{\text{pool}} \leq C_{\text{stop,global}}^{\text{pool}} \) then terminate treatment in all TZs and MZs.
3. If \( C_{\text{UCL,TZ}} \leq C_{\text{stop,boul}}^{\text{pool}} \) then terminate treatment for TZ.
4. If TZi has multiple MZs and \( C_{\text{UCL,MZ}_j} \leq C_{\text{stop,boul}}^{\text{pool}} \) then terminate treatment for MZj.
5. Repeat step 4 for all MZs in TZj.
6. Repeat steps 3-5 for all TZs.
7. Repeat steps 1-6 for next sampling/monitoring event until all treatment is terminated.

where \( C_{\text{UCL,global}}^{\text{pool}} \) is a site-wide UCL based on statistically pooled soil sampling and mass recovery data across all TZs and MZs, \( C_{\text{UCL,TZ}} \) is the value computed using pooled data from all MZs within TZi, \( C_{\text{UCL,MZ}_j} \) is the value from pooled data from a single MZj within TZi, \( C_{\text{stop,global}}^{\text{pool}} \) is a site-wide cleanup criteria specified by the site owner in consultation with regulators, and \( C_{\text{stop,boul}}^{\text{pool}} \) is a termination criteria for individual TZs or MZs that may be equal to or less than \( C_{\text{stop,global}}^{\text{pool}} \).

Details regarding the calculation of \( C_{\text{UCL,global}} \), \( C_{\text{UCL,TZ}} \), and \( C_{\text{UCL,MZ}_j} \) to obtain equal reliability termination decisions at all scales are given in Appendix S2.

The MZ stop criteria that apply when soil and/or mass recovery data are used to make individual MZ termination decisions which may be obtained by statistically pooling information from both soil and mass recovery data. If mass recovery data is employed, then mass recovery must be monitored independently for each MZ. The standard deviation of individual ln soil concentration measurements within TZi, is characterized by \( \sigma_{\text{lnTZ}_i} \), and we assume that all MZs within a given TZi have the same uncertainty (assuming random sampling locations). Different measurement types are weighted inversely proportional to their variance (Kool et al. 1987) and the pooled standard deviation is computed as a weighted root mean square.

The pooled data termination criteria for an entire TZ can be computed from the volume-weighted average AZ soil concentrations within the TZ and its pooled standard deviation and site-wide termination criteria may be obtained by upscaling TZ statistical parameters in the same manner.

UCL values at each decision scale will increase as the number of soil samples per sampling event increases, making it easier to meet stop criteria at a prescribed probability level. Increasing the number of soil samples thus enables earlier system termination, but there will be a trade-off between the cost for additional samples vs. operating cost reductions for earlier termination. Similarly, increasing the frequency of sampling will enable heating to be terminated earlier on average, which is likely to produce some savings, but at the expense of greater sampling costs. However, sampling and analytical costs are typically small relative to other operating costs, making a net cost reduction likely. Increasing the number of MZs also has the potential to reduce operating costs by terminating some areas sooner at the cost of more measurements. Additional performance monitoring variables that will affect decision uncertainty and cost include the number of MZs per TZ, the number of locations sampled per sampling event within each MZ, the number of depths sampled per boring during each sampling event, the initial date for soil sampling, the time interval between sampling events, and the choice of measurement methods (i.e., soil samples, mass recovery data, or both). Optimization of performance monitoring parameters may be used to minimize total cost for specific site conditions, as discussed in the following section.

**Design Optimization Approach**

We have identified various factors that will affect the performance reliability and cost of thermal remediation—some of which are inherent properties of the site and others that can be manipulated and hence treated as design variables. Owing to the large number of factors, uncertainty in true values of many properties, and complexity of interactions, ad hoc design approaches are likely to be suboptimal in terms of performance and/or cost. We wish to evaluate potential performance improvement and cost reductions for thermal treatment associated with various monitoring strategies by the application of optimization methods using the
Stochastic Cost Optimization Toolkit (SCOToolkit) program (Parker et al. 2011; Kim et al. 2012; Lee et al. 2012), which performs optimization analyses to determine design parameters that minimize expected (i.e., probability-weighted) total cost to meet specified remediation criteria. However, our focus here will be on optimization of monitoring parameters to meet specified source cleanup objectives without direct consideration of downstream plume behavior.

Each Monte Carlo simulation considers equally probable realizations of thermal model parameters (e.g., $M_e$ and $S_{therm}$) with termination decisions based on “noisy” data. At the termination of each realization, the performance and cost is evaluated and an optimization algorithm is used to iteratively adjust specified design variables to minimize the “expected” (i.e., probability-weighted average) cost. Cost function details are described in Appendix S3.

To encourage the optimization algorithm to identify design parameter values that have a high probability of meeting remediation objectives, the optimization objective function adds a user-defined “penalty cost” for each Monte Carlo realization if the “true” site-wide average soil concentration (i.e., with “noise-free” measurements) exceeds the global stop criteria. The penalty cost may be a real cost, for example, the anticipated cost to implement “Plan B” if the initially proposed approach fails (e.g., a plume containment system), or it may be a fictitious value selected to yield a desired probability of success. Design parameters are determined to minimize the expected cost including any penalty costs. However, the penalty cost is not included in reported expected costs for optimized designs.

Description of Optimization Problem

We consider a hypothetical problem involving thermal treatment of a dense nonaqueous phase liquid (DNAPL) PCE source in an unconfined aquifer using TCH. The site consists of 1.5 m of gravelly fill over interbedded silt and clay with some sand lenses to a depth of 9.1 m, and with clay from 9.1 to 12.2 m over bedrock. A water table occurs at 4.5 m with an average Darcy velocity of 0.2 m/year. Three TZs were identified based on site characterization data (Table 6, Figure 5). Thermal treatment is planned using TCH from the surface to the maximum observed PCE depth of 4.6 m.

The site cleanup objective ($C_{soil_{global}}$) is to achieve average soil concentration over the entire treatment volume <1 mg/kg. Optimization analyses that couple thermal treatment and dissolved plume migration could be used to determine the most cost-effective approach to meet groundwater criteria (Parker et al. 2010, 2012). However, we focus here solely on optimizing thermal treatment operational variables based on specified soil cleanup criteria.

Contaminant mass in each TZ ($M_{op}$) for each Monte Carlo realization prior to commencing thermal treatment is generated assuming a lognormal distribution as

$$M_{op} = \rho_{soil} V_{TZ} \exp(m_{ln_{soilTZ}} + N_{rnd}(0,1))$$

where $V_{TZ}$ is the bulk TZ volume, $m_{ln_{soilTZ}}$ is the mean ln soil concentration in the TZ computed as (ln $C_{max}$ + ln $C_{min}$)/2 based on values in Table 6 assuming $C_{max}$ and $C_{min}$ represent ± one standard deviation confidence limits of a lognormal population. $S_{therm}$ is the ln standard deviation of $M_e$ which is assumed to be 0.7, and $N_{rnd}(0,1)$ is a normally distributed random variable with zero mean and unit standard deviation. If TZs are divided into multiple MZs, the initial mass in the jth MZ is generated such that the total equals $M_{op_{TZ}}$. Based on results from the eight field sites discussed previously, $S_{therm}$ uncertainty is characterized by a uniform distribution with a range from 0.2 to 0.8.

The number of heating and vapor recovery wells and related infrastructure, for example, problem were determined based on heat balance calculations taking into consideration capital and operating cost trade-offs with

<table>
<thead>
<tr>
<th>Variable</th>
<th>TZs for Multi-TZ Cases</th>
<th>Single TZ Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>TZ area (m²)</td>
<td>609</td>
<td>288</td>
</tr>
<tr>
<td>$C_{\min}$ (mg/kg)</td>
<td>2</td>
<td>200</td>
</tr>
<tr>
<td>$C_{\max}$ (mg/kg)</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td>Number of heating wells</td>
<td>54</td>
<td>25</td>
</tr>
<tr>
<td>Number of recovery wells</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>$S_{TZ_{op/day}}$ (Sk)</td>
<td>2.015</td>
<td>0.933</td>
</tr>
<tr>
<td>Costs independent of TZ</td>
<td>1953.0</td>
<td></td>
</tr>
<tr>
<td>$S_{cap}$ (Sk)</td>
<td>2.137</td>
<td></td>
</tr>
<tr>
<td>$S_{cost_{op/day}}$ (Sk)</td>
<td>0.0055</td>
<td></td>
</tr>
<tr>
<td>$S_{maz}$ (Sk/kg)</td>
<td>0.200</td>
<td></td>
</tr>
<tr>
<td>$S_{M_{opz}}$ (Sk/d)</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td>$S_{M_{opz}}$ (Sk/d)</td>
<td>2.175</td>
<td></td>
</tr>
<tr>
<td>$S_{initial}$ (Sk)</td>
<td>0.280</td>
<td></td>
</tr>
</tbody>
</table>

* $C_{\max}$ and $C_{\min}$ are regarded as ± one standard deviation confidence limits.

Figure 5. Plan view of treatment zones for optimization problem.
continuous heating. The number of heating wells ($N_{\text{HZ}}$) and vapor recovery wells ($N_{\text{VW}}$) for each TZ and unit cost values computed from cost sensitivity analyses (see Appendix S3) are summarized in Table 6. The estimated time to reach the PCE azeotropic boiling point, $\Delta t_{\text{faze}}$, for the design was 65 d, which is treated as deterministic in Monte Carlo simulations. Vapor recovery is assumed to continue for 2 weeks following termination of heating at all wells.

In addition to considering cases with the site divided into three TZs with approximately known contaminant levels, we also consider the entire site treated as a single TZ. For consistency with the multi-TZ analyses, $M_I$ values for the single “lumped” TZ realizations are computed as the sum of values for the multi-TZ analyses.

Three performance monitoring strategies are considered:

1. **Soil data only.** Soil sampling is assumed to commence at a time $\Delta t_{\text{mon1}}$ after beginning thermal treatment and is repeated at time intervals $\Delta t_{\text{mon2}}$. At each sampling event, $N_{\text{boringMZ}}$ borings per MZ are advanced with $N_{\text{mpboring}}$ samples taken per boring at different depths. The total number of samples per sampling event per MZ is thus $N_{\text{boringMZ}} \times N_{\text{mpboring}}$.

2. **Mass recovery data only.** Cumulative mass recovery data for each MZ is used to estimate mass remaining and average soil concentration is computed from Equations 1 to 3 using calibrated model parameters. Mass recovery data is assumed to be available weekly to conservatively account for time to process data and implement decisions.

3. **Mass recovery and soil data.** Method 2 is used to make preliminary termination decisions, which are not implemented until soil sampling data confirm the decision. Soil sampling commences 1 week after mass recovery termination signals for the signaled regions only and is repeated at time intervals $\Delta t_{\text{mon3}}$ until pooled soil concentration data and mass recovery measurements satisfy termination criteria.

To generate “noisy” measurement data and to compute confidence limits for termination decisions, estimates of $S_{\text{HZTZ}}^{\text{imp}}$ for soil measurements and $S_{\text{n}}^{\text{me}}$ for mass recovery measurements are needed in Equation 13. Based on RMSE estimates for the eight field sites discussed above, we assume that mass recovery data have an uncertainty of $S_{\text{n}}^{\text{me}} = 0.4$. For field sites with soil sample datasets, site-wide $S_{\text{HZTZ}}^{\text{imp}}$ values computed from the raw data ranged from 2.11 to 3.86. Based on anomalous behavior observed for average soil concentration confidence limits for the field sites discussed earlier, a considerably larger site-wide $S_{\text{HZTZ}}^{\text{imp}}$ value may be appropriate. We use a site-wide value of 3.5 for example cases that involve a single lumped TZ. For cases with the site divided into three TZs with approximately equal variances, $S_{\text{HZTZ}}^{\text{imp}}$ should be smaller than the site-wide value by a factor of about $(1/3)^{1/2}$, which yields an estimate of $S_{\text{HZTZ}}^{\text{imp}} \approx 2.0$ for each TZ in the three-TZ cases.

Stochastic cost optimization analyses for the hypothetical site were performed for six cases with the site treated as a single TZ or divided into multiple TZs, with fixed or optimized values for the exceedance probability $\alpha$ and local stop criteria ($C_{\text{Slop local}}$), and with design variables for the three monitoring strategies described above optimized. The maximum number of MZs for each TZ is taken equal to the number of recovery wells in the TZ. Our objective is to evaluate effects of various operational monitoring strategies and associated optimized variables on thermal treatment performance and cost. Six optimization cases are considered, which are summarized in Table 7.

### Table 7

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring method</td>
<td>Soil</td>
<td>Soil</td>
<td>Soil</td>
<td>Soil</td>
<td>Recovery</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Probability of failure (%)</td>
<td>8</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td></td>
</tr>
<tr>
<td>Expected duration (d)</td>
<td>310</td>
<td>411</td>
<td>320</td>
<td>332</td>
<td>290</td>
<td>289</td>
<td></td>
</tr>
<tr>
<td>Costs (Sk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected total cost</td>
<td>3764</td>
<td>4099</td>
<td>3834</td>
<td>3612</td>
<td>3247</td>
<td>3580</td>
<td></td>
</tr>
<tr>
<td>95% UCL of total cost</td>
<td>4808</td>
<td>4525</td>
<td>4987</td>
<td>4100</td>
<td>3752</td>
<td>4289</td>
<td></td>
</tr>
<tr>
<td>Expected monitoring cost</td>
<td>67</td>
<td>189</td>
<td>79</td>
<td>129</td>
<td>29</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Expected other op. cost</td>
<td>1745</td>
<td>1957</td>
<td>1802</td>
<td>1530</td>
<td>1265</td>
<td>1580</td>
<td></td>
</tr>
<tr>
<td>Design variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{\text{HZ}}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Significance level ($\alpha$)</td>
<td>0.500</td>
<td>0.500</td>
<td>0.120</td>
<td>0.017</td>
<td>0.025</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>$C_{\text{Slop local}}$ (mg/kg)</td>
<td>1</td>
<td>0.094</td>
<td>0.390</td>
<td>0.948</td>
<td>0.701</td>
<td>0.783</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{HZTZ}}$</td>
<td>1</td>
<td>19</td>
<td>1</td>
<td>2,2.3</td>
<td>6,1,1</td>
<td>3,1,1</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{mpboring}}$</td>
<td>10</td>
<td>1</td>
<td>7</td>
<td>2,2,2</td>
<td>—</td>
<td>1,3,4</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{mpboring}}$</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>—</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>$\Delta t_{\text{mon1}}$ (d)</td>
<td>245</td>
<td>168</td>
<td>168</td>
<td>182</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>$\Delta t_{\text{mon3}}$ (d)</td>
<td>70</td>
<td>70</td>
<td>63</td>
<td>35</td>
<td>—</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

Bold values are fixed during optimization. Italic values are optimization results.

1Expected costs and durations are probability-weighted averages. See text for discussion.
Results of Optimization Analyses

Probability-weighted average ("expected") total costs for the various cases ranged from $3247 to $4099 k, with UCL\textsubscript{opt} from $3247 to $4987 k considering all quantifiable sources of uncertainty (Table 7). Expected durations ranged from 289 to 411 d. The UCL\textsubscript{opt} of total cost, considering all quantifiable sources of uncertainty, ranged from $3247 to $4987 k. Total costs are divided into monitoring costs, other operating costs (including energy), and fixed costs for design and construction. The latter were constant at $1953 k for all cases. Expected monitoring costs ranged from $29 to $189 k, and other operating costs from $1265 to $1957 k.

It may be noted that expected values for corresponding costs from the stochastic optimization analyses are higher than those reported in the literature (Baker et al. 2016; Heron et al. 2016). Costs are similar if normalized for treatment duration. Longer probability-weighted average treatment durations in the present study may be partly attributed to a publication bias in favor of sites with low S\textsubscript{therm} values that are mostly completed within about 2-3 times \( \Delta_{\text{peak}} \). Based on data from the sites reported here, the frequency of sites with S\textsubscript{therm} values >0.5 is about equal to that of lower S\textsubscript{therm} sites. However, average remediation achievement to 99% mass reduction for low S\textsubscript{therm} sites (0.5-0.8) is about 2.4 times \( \Delta_{\text{peak}} \) compared to 4.8 times \( \Delta_{\text{peak}} \) for high S\textsubscript{therm} sites (0.5-0.8). To reach 99.9% mass reduction, these ratios climb to 3.3 times \( \Delta_{\text{peak}} \) for low S\textsubscript{therm} sites and 8.3 times for high S\textsubscript{therm} sites. A second factor may be that high S\textsubscript{therm} sites tend to be terminated at lower mass recovery ratios owing to durations exceeding time and budget expectations. For the sites reported here, the average mass reduction ratio computed for low S\textsubscript{therm} sites was 99.8% vs. only 95% for high S\textsubscript{therm} sites. A third factor contributing to longer treatment durations and costs is that design optimizations were formulated to achieve a high probability (in most cases >99%) that remediation criteria will be met. It is likely that most systems designs are not this stringent. For these reasons, we believe the expected durations and costs from the stochastic optimization results are realistic considering the full range in S\textsubscript{therm} values that may occur and assuming comparable cleanup levels are met with a high probability of success regardless of site recalcitrance.

Results for the various cases are discussed below.

Opt1. The first four cases utilize soil data only to make termination decisions for thermal treatment and treat the entire treatment volume as a single TZ. For Opt1 \( \alpha = 0.5 \) (hence \( C_{\text{UCL}_{\alpha}} = C_{\text{UCL}_{0.5}} \) and \( C_{\text{UCL}_{\alpha}} = C_{\text{UCL}_{0.5}} \)) indicating that heating in individual MZs and individual TZs is terminated when the measured average soil concentration within the respective area is below the site-wide stop criteria. These operating procedures are typical of industry practice. The number of MZs in the TZ (\( N_{\text{MZ/TZ}} \)) soil borings per MZ for each sampling event (\( N_{\text{sample/mg}} \)) soil sample depths per boring (\( N_{\text{sample/mg}} \)), and time which soil monitoring commences (\( \Delta_{\text{start}} \)) and the time between successive sampling events (\( \Delta_{\text{interval}} \)) were optimized.

The fixed \( \alpha \) value of 0.5 in conjunction with the condition that \( C_{\text{opt/stop local}} = C_{\text{opt/stop global}} \) made it difficult to find a set of design variables that could reliably achieve the remediation target. The best that could be managed by optimization still suffered an 8% probability that the true average concentration will exceed the target value of 1 mg/kg. No exceedances greater than 10 mg/kg were predicted. With optimized values of only one MZ in the single TZ, 10 borings in the MZ with 4 sampling depths per boring, this case employs a total of 50 soil samples per sampling event, yielding reasonable monitoring costs that are consistent with industry practice ($67k). The expected total cost is $3764k with $1745k for operating costs other than monitoring with the UCL\textsubscript{opt} of total cost equal to $4808. The expected treatment duration of 310 d is 4.8 times \( \Delta_{\text{peak}} \), which is consistent with the range in S\textsubscript{therm} values and remediation times inferred from the field sites discussed earlier in this paper. Aside from the consideration of risks from higher S\textsubscript{therm} values, we regard Opt1 as a reasonable approximation of typical industry practice.

Opt2. This case is the same as Opt1 except that \( C_{\text{UCL}_{\alpha}} \) is optimized subject to the constraint that it be no greater than \( C_{\text{UCL}_{\alpha}} \). The optimized value of 0.094 mg/kg requires individual TZs to reach a significantly lower concentration than the site-wide target to terminate early, which permits site-wide termination to occur when remaining areas are at a higher average concentration. This flexibility allowed probability of failure to decrease to <1%, which enabled improved reliability as reflected by a lower total cost UCL\textsubscript{opt} of $4525 compared to $4808 for Opt1. However, improved reliability was achieved at the expense of a significantly longer expected remediation duration (411 d), and higher expected total cost ($4099k), monitoring cost ($189k), and other operating cost ($1957k).

Opt3. This case is the same as Opt2 except that \( \alpha \) is also optimized to a value of 0.120. The resulting design also achieves an exceedance probability of <1% but with a shorter expected duration (320 d) and lower expected monitoring, other operating and total costs ($79, $1802, and $3834k, respectively) using only one MZ with seven borings per MZ sampled at four depths. However, the total cost UCL\textsubscript{opt} for Opt3 ($4987k) is greater than that for Opt1 or Opt2.

Opt4. This case is identical to Opt3, except that the site is divided into three TZs that have less uncertainty in average soil concentration than the site as a whole. The optimized value of \( \alpha \) for this case is 0.017, while the value of 0.948 mg/kg for \( C_{\text{UCL}_{0.5}} \), which is essentially the same as the site-wide criteria and less aggressive than the values for Opt 2 (0.094) and Opt 3 (0.390). The two largest and least contaminated TZs (TZ1 and TZ2) are each divided into two MZs and TZ3 is divided into three MZs for a total of seven MZs. Two borings per sampling event are taken from each MZ and TZ if they have not already terminated. Four depths are sampled per all borings.

Although the expected treatment duration is slightly longer than that for Opt3 at 332 d, energy savings from early termination of MZs or TZs resulted in significantly lower expected nonmonitoring operating costs ($1530k), expected total cost ($5612k), and total cost UCL\textsubscript{opt} ($4100k). Opt4 has a lower expected total cost and UCL\textsubscript{opt} of total cost, as well as a significantly higher probability of success than Opt1.
Opt5. This case is the same as Opt4 with three TZs, except that mass recovery measurements for each MZ are used to make termination decisions (Method 2). No soil sampling is performed during thermal treatment or for confirmation after treatment. The results achieve an exceedance probability of less than 1% with an optimized $\alpha$ value of 0.025 and $c_{\text{stop local}}$ of 0.701 mg/kg. The number of MZs per TZ is six for TZ1 (the largest, least contaminated zone) and only one for TZ2 and TZ3, for a total of eight MZs. Expected monitoring costs for Opt5 ($S_{\text{opt}}$) are much lower than for any of the soil monitoring cases (Opt1-Opt4). Because mass recovery data have lower measurement uncertainty and are available with much higher frequency (weekly is assumed), termination decisions can be made much sooner on average than with soil sample data. This is evidenced by a lower expected treatment duration (289 d) than Opt1-Opt4, hence sharply lower expected nonmonitoring operating costs ($\$1265\text{k}$) and total cost ($\$3247\text{k}$). The expected total cost is 10% lower than the best case using soil data only (Opt4) with the same probability of success. The expected total cost is also 14% lower than that for Opt1, the surrogate for “typical practice” that has the additional liability of an 8% probability of failure. Furthermore, the UCL$_{95}$ for Opt5 total cost ($\$3752\text{k}$) is significantly less than corresponding values for all soil monitoring cases (Opt1-Opt4) and less than the expected total cost for all but Opt4.

Opt6. This case is similar to Opt4 and Opt5, except that monitoring is performed using Method 3, which employs mass recovery data by itself until a termination signal is obtained for site-wide, TZ, or MZ termination, after which soil data is collected periodically until pooled soil and recovery data confirm the decision. The results achieve an exceedance probability of less than 1% with a $c_{\text{stop local}}$ of 0.783 mg/kg and a stringent $\alpha$ value of 0.003. The number of MZs per TZ is three for TZ1 (largest, least contaminated) and one MZ for each of TZ2 and TZ3, for a total of five MZs. Only one boring per MZ is specified for TZ1 with three for TZ2 and four for TZ3 sampling four depths per boring for each location. The frequency of soil sampling after a termination signal based on mass recovery data is 28 d. Monitoring costs ($\$48\text{k}$) are not much higher than for Opt5 and the operating time of 289 d is essentially the same as for Opt5. However, other operating costs for Opt6 ($\$1580\text{)}$ are 25% higher than for Opt5, which is attributable to a 39% higher average energy utilization for Opt6 due to fewer early terminations of individual MZs and/or TZs. Relatively large uncertainty in soil data result in wider pooled confidence limits for Opt6 termination and a significantly higher UCL$_{95}$ of total cost.

**Summary and Conclusions**

Thermal treatment methods are effective technologies for remediation of DNAPL source zones due to their relatively low sensitivity to aquifer heterogeneity and DNAPL distributions. Nevertheless, significant uncertainty exists in the duration of heating required to meet remedial goals for a given system design. Normal distribution models for mass recovery as a function of time are unable to capture the positive skew of actual recovery data, which can lead to significant underestimation of the treatment duration necessary to reach cleanup objectives. We introduced a lognormal distribution model with recovery time duration characterized by the standard deviation in ln recovery time, $S_{\text{therm}}$, with values ranging from about 0.2 to 0.8 on a site-wide basis for field sites studied. For $S_{\text{therm}}=0.2$, remediation duration can range from 1.6×$\Delta t_{\text{peak}}$ (time to reach effective boiling point) to achieve a mass reduction of 99% to 2.4×$\Delta t_{\text{peak}}$ for 99.99% reduction. While for $S_{\text{therm}}=0.8$, treatment durations from about 6 to 20×$\Delta t_{\text{peak}}$ are predicted for the same mass reduction percentages.

We were unable to identify any significant correlations between $S_{\text{therm}}$ values and geologic complexity, DNAPL source complexity, or initial contaminant concentration of the sites. Furthermore, since differences in $S_{\text{therm}}$ values within a given site were as variable as differences between sites, estimates of $S_{\text{therm}}$ from pilot tests may not be predictive of the whole site. A priori uncertainty in $S_{\text{therm}}$ can result in significant uncertainty in site-wide treatment times. Iterative calibration of lognormal model parameters from mass recovery data provides incrementally refined estimates of $S_{\text{therm}}$ and other model parameters which enable extrapolation of contaminant mass remaining for use in making reliable real-time termination decisions. Uncertainty in $S_{\text{therm}}$ and other factors affecting treatment duration are taken into consideration in the design process using stochastic optimization methods.

We have proposed a strategy to turn the liability of uncertainty in time to reach cleanup objectives into a potential advantage by dividing the contaminated soil volume into TZs that exhibit different average contamination levels based on site characterization data and (optionally) further dividing TZs into MZs for purposes of making termination decisions. We also allow target soil concentrations for regions smaller than the full site (local stop criteria) to be specified at a value less than the site-wide stop criterion. Cleaning up less recalcitrant regions (lower initial soil concentration and/or $S_{\text{therm}} \gtrsim$) to a lower average concentration enables more recalcitrant regions (higher initial soil concentration and/or $S_{\text{therm}} \gtrsim$) to be terminated at a higher average concentration to achieve the same site-wide average, which offers the possibility of reducing overall treatment duration and total operating cost. The multilevel monitoring and termination strategy allows for site-wide termination as well as early termination of individual TZs or MZs within TZs.

Thermal system termination decisions are commonly made by comparing the average concentration computed from a round of soil samples directly with a cleanup target. However, averages from soil data are subject to large uncertainty even when the number of soil samples is large. An alternative method, which estimates average soil concentration from mass recovery measurements during thermal treatment using the lognormal mass recovery model, has been demonstrated that exhibits less uncertainty and lower cost than soil sampling. To explicitly account for uncertainty in average soil concentrations estimated from soil and/or mass recovery data, the multilevel termination strategy stops treatment when an UCL of estimated mean concentration at a specified probability is below the target concentration. We employ a statistical methodology for computing confidence limits at site-wide, TZ and MZ levels that allows termination decisions to be made at all scales with equal reliability.
To identify cost-optimal performance monitoring strategies to guide termination decisions, we incorporated the multiscale decision protocol into the stochastic cost optimization program SCOToolkit, which identifies design variables that minimize probability-weighted total cost considering uncertainty in site properties, model predictions, and monitoring data while maintaining a high likelihood of meeting remediation objectives.

Results for an example problem indicate that the practice of using computed average soil concentration (as opposed to an UCL) cannot achieve a high probability of meeting the target average soil concentration. Optimizing the confidence limit probability, local-scale cleanup level, number of MZs per TZ, soil borings per MZ for each sampling event, sample depths per boring, date for first sampling event, and time interval between sampling events for a site treated as a single TZ using only soil sampling data achieved cleanup objectives with a higher probability of success than a more conventional approach. Dividing the site into three TZs with different soil concentration ranges and optimizing the same variables reduced total cost by 6%. Optimizing confidence limit probability, local-scale cleanup level, and number of MZs per TZ with three TZs while using mass recovery data instead of soil data, achieved an additional 10% cost reduction. If confirmation of mass recovery-based results with soil sample data is desired or required, delaying each local termination decision until confirmed by soil sampling will increase the cost. Therefore, if confirmatory soil sampling is required, we recommend waiting until all heating units have been stopped based on mass recovery data before performing site-wide soil sampling.

In addition to computing the probability-weighted average cost for optimized designs, the method gives cost probability distributions that reflect uncertainty in measurements and calculations. An optimized example problem using only mass recovery data to make termination decisions (Opt 5) had a 16% lower expected total cost than a case that approximates typical industry practice (Opt 1), while the UCL of total cost for the former was 28% lower. Thus, the proposed methodology not only yields "expected" cost savings, but also sharply reduces the magnitude of potential cost overruns.

The MATLAB-based SCOToolkit program, as well as Excel spreadsheet templates for calibrating the mass recovery model and for making real-time termination decisions from mass recovery and/or soil sample data based on the methods described in this paper, can be obtained upon request to the second author at u.kim@csuohio.edu.

Acknowledgment

This research was conducted with funding from the U.S. Department of Defense Strategic Environmental Research and Development Program (SERDP) Environmental Restoration Program managed by Dr. Andrea Leeson under project ER-2310.

Supporting Information

The following supporting information is available for this article:

Appendix S1. Calibration of mass recovery model
Appendix S2. Risk adjusted statistical criteria for termination decisions

Appendix S3. Thermal treatment cost function
Appendix S4. List of abbreviations and symbols

References


**Biographical Sketches**

**Jack C. Parker,** Ph.D., corresponding author, is a research professor at the University of Tennessee. Previously, he served as a research scientist at Oak Ridge National Laboratory, as a principal investigator at the University of Tennessee, and as a professor at Virginia Tech. He has authored more than 200 publications dealing with multiphase transport, inverse modeling, error analysis and remediation optimization. He has given lectures, workshops, and courses in more than 20 countries. He can be reached at the Civil and Environmental Engineering Department, University of Tennessee, Knoxville, TN 37996; Tel.: (865) 974-7718; j.parker@utk.edu.

**Ungtae Kim,** Ph.D., is an assistant professor in the Department of Civil and Environmental Engineering, Cleveland State University, Cleveland, Ohio. He received a Ph.D. degree from Utah State University focusing on hydrology and water resources engineering. He implemented the Stochastic Cost Optimization Toolkit (SCOToolkit) code described in this paper. He has authored more than 60 publications in the field of watershed hydrology and groundwater transport/remediation. He can be reached at the Civil and Environmental Engineering Department, Cleveland State University, Cleveland, OH 44115; (216) 687-2599; u.kim@csuohio.edu.

**Alyson Fortune,** M.S., is a senior scientist with TerraTherm. She holds a B.S. in environmental science and a M.S. in atmospheric science. Alyson has more than 15 years of experience in the environmental industry, with experience including stationary source testing, vapor intrusion, and vapor-phase sampling and analysis. She manages thermal remediation treatability studies, conducts data quality reviews, maintains field monitoring systems, and performs other data management functions for TerraTherm. She can be reached at TerraTherm Inc., Gardner, MA 01440; (978) 730-1200.

**Steffen Griepke,** M.S., is a principal engineer and technical director at TerraTherm with a master’s degree in environmental engineering. He is responsible for TerraTherm’s subsurface designs and models, including selecting the most appropriate thermal method based on site geology, hydrogeology, contaminants and remedial goals. He also leads TerraTherm’s Technology Team and data management efforts. He can be reached at TerraTherm Inc., Gardner, MA 01440; (978) 730-1200.

**James Galligan,** M.S., P.E., is TerraTherm’s Vice President of Projects. He has more than 20 years of experience with in situ and ex situ remediation system design, installation, operation and troubleshooting, including over 10 years with ISIFR projects. Mr. Galligan has conducted numerous remedial technology pilot tests, feasibility studies and life-cycle cost evaluations. He has extensive experience in cost estimation and health and safety management. He can be reached at TerraTherm Inc., Gardner, MA 01440; (978) 730-1200.

**Amber Bonarrigo,** B.S., is a geologist and data manager at TerraTherm with a B.S. in geology. She assists with TerraTherm’s subsurface design and modeling as well as processing field data to estimate mass removal rates, average site temperature, utility usage, and other performance statistics at active thermal sites. She can be reached at TerraTherm Inc., Gardner, MA 01440; (978) 730-1200.