involve
a journal of mathematics

Parameter identification and sensitivity analysis to a thermal diffusivity inverse problem

Brian Leventhal, Xiaojing Fu, Kathleen Fowler and Owen Eslinger

msp

2015 vol. 8, no. 3

INVOLVE 8:3 (2015)
dx.doi.org/10.2140/involve.2015.8.385

Parameter identification and sensitivity analysis to a thermal diffusivity inverse problem

Brian Leventhal, Xiaojing Fu, Kathleen Fowler and Owen Eslinger

(Communicated by Suzanne Lenhart)

The solution to inverse problems is an application shared by mathematicians, scientists, and engineers. For this work, a set of shallow soil temperatures measured at eight depths between 0 and 30 cm and sampled every five minutes over 24 hours is used to determine the diffusivity of the soil. Thermal diffusivity is a modeling parameter that impacts how heat flows through soil. In particular, it is not known in advance if the subsurface region is homogeneous or heterogeneous, which means the thermal diffusivity may or may not depend on depth. To this end, it is not clear which assumptions may apply to represent the physical system embedded within the parameter estimation problem. Analytic methods and a simulation based least-squares approach to approximate the diffusivity are compared to fit the temperature profiles to different heat flow models. The simulation is based on a spatially dependent, nonsteady-state discretization to a partial differential equation. To complete the work, a statistical sensitivity study using analysis of variance is used to understand how errors that arise in the modeling phase impact the final model. We show that for the analytic methods, errors in the initial fitting of the temperature data to sinusoidal boundary conditions can have a strong impact on the thermal diffusivity values. Our proposed framework shows that this soil sample is heterogeneous and that modeling depends significantly on data used as top and bottom boundary conditions. This work offers a protocol to determine the soil type and model sensitivities using analytic, numerical, and statistical approaches and is applicable to modifications of the classic heat equation found across disciplines.

1. Introduction

Inverse problems arise routinely across science and engineering disciplines. Using a mathematical approach to such parameter estimation problems avoids the tedious task of trial-and-error to match a mathematical model to experimental data. For this work, we consider a heat transport model in the shallow subsurface and use both

MSC2010: primary 35K05, 49N45, 62J10; secondary 35Q93.
Keywords: inverse problems, subsurface flow, sensitivity analysis.

analytic and numerical approaches to fit data. Part of the challenge is that the nature of the subsurface is not known in advance; thus it is not clear which model applies or whether assumptions made to apply analytic models are reasonable. In applying numerical approaches, assumptions on the types of boundary conditions can significantly impact the results. In the presence of such uncertainty and the possible addition of experimental error, the identified parameters may give suboptimal fits or provide values far from truth. This work offers a protocol to determine the soil type and model sensitivities using analytic, numerical, and statistical approaches by comparing common approaches to heat flow in the shallow subsurface and studying how choices made during the modeling phase can impact the results of the inverse problem.

The propagation of heat in the subsurface can be modeled by the second-order partial differential equation

$$\frac{\partial T}{\partial t} = K\frac{\partial^2 T}{\partial z^2}, \tag{1}$$

where $T(z,t)$ is the time-dependent temperature distribution at depth $z > 0$ for $t > 0$. Thermal diffusivity, K cm/min, which describes how easily heat propagates through the medium, is proportionally related to thermal conductivity such that

$$K = \frac{\hat{k}}{\rho c}, \tag{2}$$

where $\hat{k}$ is the thermal conductivity, ρ is the density and c is the heat capacity. Although in (1), K is often assumed to be constant in practice, due to the complex nature of the subsurface, K is usually spatially dependent. We refer to these as homogeneous and heterogeneous soils respectively. Heat flow in the heterogeneous case would be described by

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left[K\frac{\partial T}{\partial z}\right] = K\frac{\partial^2 T}{\partial z^2} + \frac{\partial K}{\partial z}\frac{\partial T}{\partial z}, \tag{3}$$

where now $K = K(z)$. Analytic solutions to various forms of these models exist [Carslaw and Jaeger 1986; Powers 2006; Narasimhan 2009] and have been studied for decades. Alternatively, given the spatially distributed thermal conductivity along with initial and boundary temperature, the temperature distribution over time can be approximated numerically.

The inverse of this problem is the focus of this study. Mathematical approaches can be used to help guide practitioners on the nature of the subsurface since it is not known in advance how much the soil type actually varies. Specifically, subsurface temperature data monitored at seven depths between 0 and 30 cm and logged over time were used to determine the thermal diffusivity of the test site. Figure 1 below

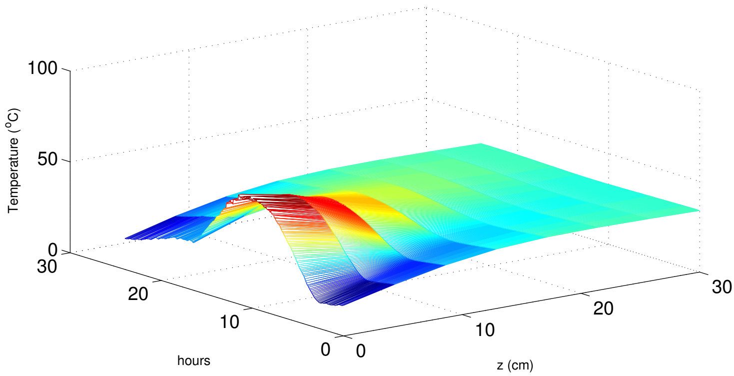

Figure 1. Temperature data.

shows temperature data as a function of time and depth. Analytic methods for determining K from temperature profiles have been proposed in the literature and implemented using data taken from the Loess Plateau in China [Gao et al. 2009]. Some of those methods are compared here but are based on the assumption that the soil is homogeneous. We compare these approaches to a simulation-based approach using a numerical approximation to the heterogeneous model in (3) and a minimization of the least-squares error between the model output and the temperature data. Since all of these methods include choices made during the modeling phase, we conduct a sensitivity study to understand how these choices impact the final model. The sensitivity study is based on a statistical analysis of variance.

We proceed by describing the methods used to determine the thermal diffusivity, both analytically and numerically, and then presenting those results in Sections 2 and 3. We follow with the sensitivity analysis in Section 4 and point the way towards future work in Section 5.

2. Analytic approaches

We consider four methods that approximate K values explicitly using temperature values at different depths. The methods are based on the homogeneous model in (1).

If we consider boundary conditions of the form

$$T(0,t) = T_a + A \sin(\omega t + \phi) \tag{4}$$

and

$$\lim_{z\to\infty} T(z,t) = T_a, \tag{5}$$

an analytic solution to (1) is given by

$$T(z,t) = T_a + A e^{-z/D} \sin\Big(\omega t - \frac{z}{D} + \phi\Big), \tag{6}$$

with $D = \sqrt{2K/\omega}$. Here, (4) states that the surface temperature varies as a sinusoidal function whose parameters include the time-average temperature T_a (°C), amplitude A (°C), radial frequency ω (rad s^{-1}) and phase constant ϕ (rad). The bottom boundary condition (5) indicates that as depth increases sufficiently, the soil temperature is not affected by the surface temperature and thus maintains a constant value.

Four analytic methods were used to approximate the thermal conductivities at seven locations. The seven locations are between different depths, i.e., between 0 and 1 cm, between 1 and 5 cm, between 5 and 10 cm, and continuing until 30 cm deep. The methods described in [Gao et al. 2009; Horton et al. 1983] call for a homogeneous soil thermal conductivity profile. With thermal conductivity assumed homogeneous, the analytic methods call for only two depths to estimate the conductivity. To perform the analytic methods, the raw data temperatures need to be approximated by a sinusoidal curve of the form

$$T_1(z_1, t) = \overline{T}_1 + A_1 \sin(\omega t + \phi_1), \tag{7}$$

$$T_2(z_2, t) = \overline{T}_2 + A_2 \sin(\omega t + \phi_2), \tag{8}$$

where A_1, A_2 are half of the difference between the daytime maximum and nighttime minimum amplitudes for the soil depths. Furthermore, $\overline{T}_1, \overline{T}_2$ are the arithmetic averages of the daytime maximum soil temperature and the nighttime minimum soil temperature at depths z_1, z_2. The initial phases of the soil temperature, ϕ_1 and ϕ_2, are obtained using a least-squares fit (as opposed to using a spline to fit the data) because numerical values for those parameters are needed in the analytic models to determine the conductivity. The resulting least-squares problem is nonlinear and a variety of optimization methods would apply. Since a genetic algorithm [Holland 1973] was being used in the project elsewhere, it was used here as well. Genetic algorithms require no gradient information for minimization and are thus attractive choices for an off-the-shelf optimization approach.

Since sinusoidal approximation is only needed at two depths for the analytic models, the two depths whose sinusoidal curves give the least error compared to the raw error are used to compute the thermal conductivity. With each producing a residual of 10^{-1}, the data located one centimeter and five centimeters deep were used. Tables 1 and 2 show the results of the fit curve for each of the seven days of data. Table 3 shows the top boundary condition sinusoidal parameters as well.

The four methods considered for this experiment are the amplitude method, the phase method, the arctangent method and the logarithmic method. Essentially, if we assume that K is independent of depth (i.e., the media is homogeneous) and that the boundary temperature is sinusoidal, then the analytic solution of the one dimensional heat equation can be used to approximate K. The amplitude and phase methods are directly based on the analytic solution above. The arctangent and

day	A_1	ω	ϕ_1	$\bar{T}_1$
1	$1.45 \cdot 10^1$	$6.51 \cdot 10^{-3}$	3.13	$3.98 \cdot 10^1$
2	$1.54 \cdot 10^1$	$5.58 \cdot 10^{-3}$	5.58	$3.99 \cdot 10^1$
3	$1.45 \cdot 10^1$	$5.46 \cdot 10^{-3}$	5.46	$3.88 \cdot 10^1$
4	$1.38 \cdot 10^1$	$5.16 \cdot 10^{-3}$	5.16	$3.73 \cdot 10^1$
5	$1.58 \cdot 10^1$	$3.94 \cdot 10^{-3}$	3.94	$3.35 \cdot 10^1$
6	$1.64 \cdot 10^1$	$5.49 \cdot 10^{-3}$	5.49	$3.74 \cdot 10^1$
7	$1.73 \cdot 10^1$	$3.94 \cdot 10^{-3}$	2.44	$3.28 \cdot 10^1$

Table 1. Parameters obtained at a depth of 1 cm.

day	A_2	ω	ϕ_2	$\bar{T}_2$
1	$8.73 \cdot 10^1$	$5.75 \cdot 10^{-3}$	3.25	$3.75 \cdot 10^1$
2	$8.65 \cdot 10^1$	$5.69 \cdot 10^{-3}$	1.55	$3.80 \cdot 10^1$
3	$8.12 \cdot 10^1$	$5.11 \cdot 10^{-3}$	1.67	$3.69 \cdot 10^1$
4	$7.60 \cdot 10^1$	$5.02 \cdot 10^{-3}$	1.19	$3.62 \cdot 10^1$
5	$8.37 \cdot 10^1$	$5.60 \cdot 10^{-3}$	2.66	$3.66 \cdot 10^1$
6	$9.93 \cdot 10^1$	$3.87 \cdot 10^{-3}$	1.97	$3.31 \cdot 10^1$
7	$9.39 \cdot 10^1$	$4.51 \cdot 10^{-3}$	2.95	$3.47 \cdot 10^1$

Table 2. Parameters obtained at a depth of 5 cm.

day	amplitude	ω	phase	$\bar{T}$
1	$1.74 \cdot 10^1$	$6.81 \cdot 10^{-3}$	3.00	$3.76 \cdot 10^1$
2	$2.14 \cdot 10^1$	$5.68 \cdot 10^{-3}$	2.33	$3.82 \cdot 10^1$
3	$1.96 \cdot 10^1$	$5.21 \cdot 10^{-3}$	2.18	$3.61 \cdot 10^1$
4	$1.84 \cdot 10^1$	$4.85 \cdot 10^{-3}$	2.81	$3.35 \cdot 10^1$
5	$2.05 \cdot 10^1$	$4.79 \cdot 10^{-3}$	2.40	$3.34 \cdot 10^1$
6	$2.28 \cdot 10^1$	$5.48 \cdot 10^{-3}$	2.59	$3.51 \cdot 10^1$
7	$2.26 \cdot 10^1$	$5.27 \cdot 10^{-3}$	2.92	$3.48 \cdot 10^1$

Table 3. Parameters obtained for the boundary condition.

logarithmic methods are based on the notion that a Fourier series can reduce errors introduced by the assumption that a single sinusoidal wave is sufficient to estimate the surface temperature. We state these approaches here and point the reader to [Gao et al. 2009; Horton et al. 1983] for more details.

The amplitude method:

$$K = \frac{\omega(z_1 - z_2)^2}{2\ln(A_1/A_2)^2}. \tag{9}$$

The phase method:

$$K = \frac{\omega (z_1 - z_2)^2}{2(\phi_1 - \phi_2)^2}. \tag{10}$$

The arctangent method: This method is based on the notion that soil temperature can be described by a Fourier series,

$$T = \overline{T} + \sum_{i=1}^{n} (a_i \sin(i \omega t) + b_i \cos(i \omega t)).$$

With $n = 2$, K can be estimated with

$$K = \left(\frac{\omega \Delta z^2}{2 \arctan \frac{(T_1 - T_3)(T_2' - T_4') - (T_2 - T_4)(T_1' - T_3')}{(T_1 - T_3)(T_1' - T_3') + (T_2 - T_4)(T_2' - T_4')}} \right)^2, \tag{11}$$

where temperatures T_j and T_j' are recorded at 6 hour time intervals and two different depths z_1, z_2.

The logarithmic method: Using the same assumptions as the arctangent method, K can be expressed as

$$K = \left(\frac{0.012 \Delta z}{\ln \frac{(T_1 - T_3)^2 + (T_2 - T_4)^2}{(T_1' - T_3')^2 + (T_2' - T_4')^2}} \right)^2. \tag{12}$$

Table 4 shows the results for each method for the seven days studied. As can be observed, the amplitude method and logarithmic method estimate the thermal conductivity on the same order of magnitude over the seven days. The other two methods, phase and arctangent, estimate the thermal conductivities with significant variability. They do not hold the order of magnitude constant over the seven days, thus producing significantly different results from other methods.

day	amplitude	phase	arctangent	logarithm
1	$1.11 \cdot 10^{-2}$	2.18	$9.43 \cdot 10^{-2}$	$1.45 \cdot 10^{-2}$
2	$9.39 \cdot 10^{-3}$	$6.29 \cdot 10^{-2}$	$1.63 \cdot 10^{-1}$	$3.92 \cdot 10^{-3}$
3	$9.91 \cdot 10^{-3}$	$7.73 \cdot 10^{-2}$	$2.86 \cdot 10^{-1}$	$4.56 \cdot 10^{-3}$
4	$1.02 \cdot 10^{-3}$	$9.62 \cdot 10^{-1}$	5.19	$3.04 \cdot 10^{-3}$
5	$8.43 \cdot 10^{-3}$	$1.98 \cdot 10^{-2}$	$2.75 \cdot 10^{-2}$	$1.21 \cdot 10^{-3}$
6	$9.69 \cdot 10^{-3}$	$5.12 \cdot 10^{-1}$	$2.95 \cdot 10^{-2}$	$2.91 \cdot 10^{-3}$
7	$7.93 \cdot 10^{-3}$	$1.30 \cdot 10^{-1}$	$5.09 \cdot 10^{-2}$	$3.42 \cdot 10^{-3}$

Table 4. Estimated conductivities (cm/min).

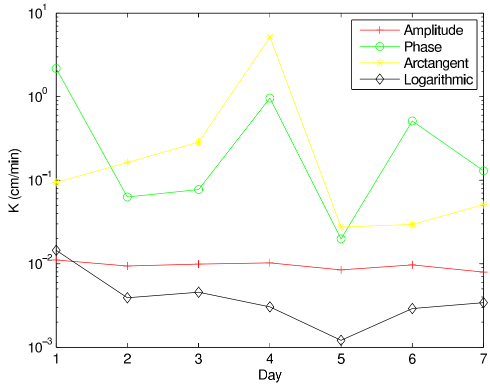

Figure 2. Estimated conductivities.

The thermal conductivity changes with days shown by Figure 2. As there are noteworthy differences between each method, the arctangent and phase methods seem to have the most significant change.

The thermal conductivities found can be used to determine temperature profiles for each method. As seen in Figure 3, not one method provides an accurate estimation of the data profile. Large errors in the original sinusoidal fitting or the inaccuracy of assuming thermal conductivity homogeneity could account for this difference in estimation. To this end, although attractive for their simplicity, the analytic methods do not provide an accurate approximation to the data.

The assumption that thermal conductivity is homogeneous throughout the soil may be inaccurate. Instead of assuming homogeneity from 0 to 30 cm, homogeneity can be assumed on small subintervals. This assumption is reasonable if the porous media is layered so that it is homogeneous in $x-y$ directions and constant on layered intervals in the z direction. These results are shown in Table 5. By displaying a change in thermal conductivity with depth in Figure 5, results either confirm that homogeneity was an inaccurate assumption or that errors from the initial fitting are having an impact on the final values.

depth (cm)	amplitude	phase	arctangent	logarithmic
1–5	$8.50 \cdot 10^{-3}$	$7.20 \cdot 10^{-2}$	$2.52 \cdot 10^{-1}$	$2.90 \cdot 10^{-3}$
5–10	$1.70 \cdot 10^{-2}$	$3.70 \cdot 10^{-2}$	$2.40 \cdot 10^{-2}$	$1.40 \cdot 10^{-3}$
10–15	$7.70 \cdot 10^{-3}$	$3.17 \cdot 10^{-2}$	$8.73 \cdot 10^{-2}$	$3.20 \cdot 10^{-3}$
15–20	$3.58 \cdot 10^{-2}$	$8.60 \cdot 10^{-1}$	$1.64 \cdot 10^{-2}$	$5.79 \cdot 10^{-4}$
20–25	$6.19 \cdot 10^{-2}$	$6.03 \cdot 10^{-2}$	$8.09 \cdot 10^{-1}$	$3.88 \cdot 10^{-5}$
25–30	$2.05 \cdot 10^{-2}$	$2.88 \cdot 10^{-2}$	$4.90 \cdot 10^{-2}$	$4.26 \cdot 10^{-3}$

Table 5. Differences in conductivities at each depth on day 7.

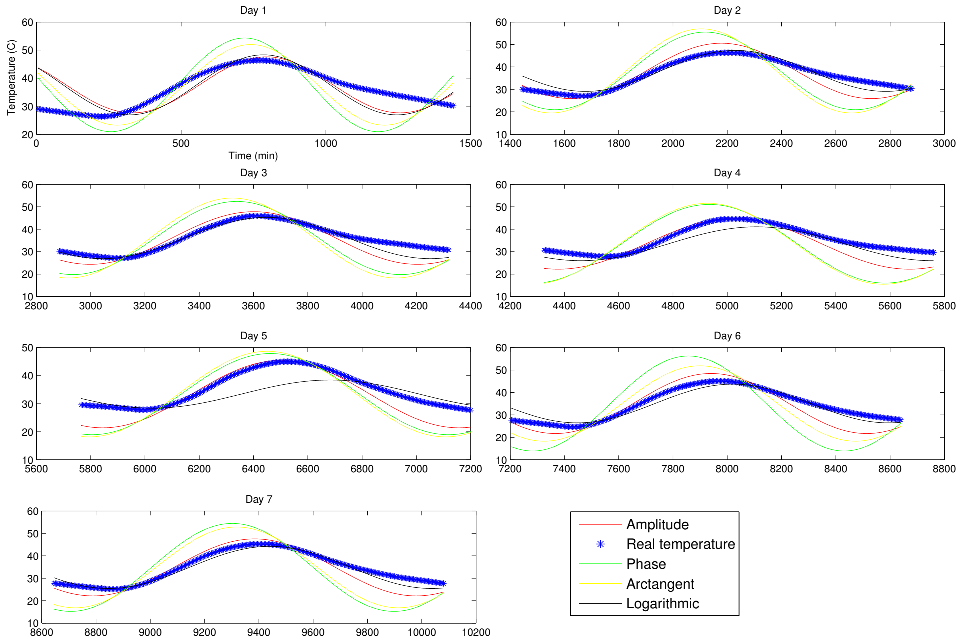

Figure 3. Temperature distributions.

We will show later that small variations in fitting the sinusoidal curve impact the analytic solutions greatly. Moreover, with regard to the inaccurate assumption of homogeneity, a new approach is taken below. We proceed by analyzing the simulation-based optimization approach to conduct the experiment with the assumption that thermal conductivities are heterogeneous.

3. Simulation-based approach

In the first approach, thermal conductivities are calculated using four analytic methods. However, the results indicate that the assumption of homogeneity may not be valid. To this end, an optimization framework where the least-squares error (LSE) between data and a simulated temperature profile facilitates the incorporation of spatially varying thermal conductivities. For the simulation, finite differences were used to discretize (3) in space with backward Euler in time. To validate the simulation tool, results were compared to a problem with a known solution using a forcing term $f(z,t)$ on the right hand side and a known function $K(z)$ to ensure accurate truncation error. To account for the fact that data would be used in the subsequent study for K, we use a spline to describe the variation of K in space and then differentiate it to obtain $\partial K/\partial z$.

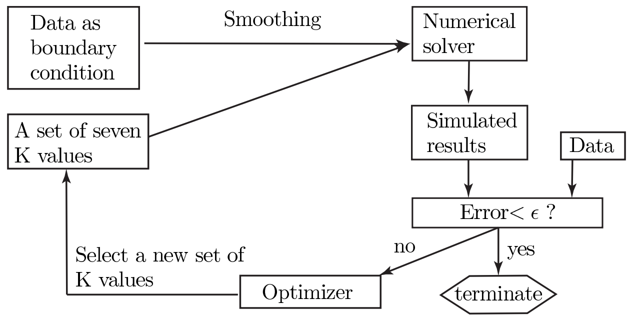

Figure 4. Structure of the optimization scheme.

In this new context, the logged data T^{obs} is an $N_t \times N_z$ matrix, where N_t is the number of time points and N_z is the number of spatial nodes. The least-squares problem is then

$$\min_{K \in \Omega} J(K) = \frac{\frac{1}{2}\sum_{i=1}^{N}(\widehat{T}_i(K) - T_i^{\text{obs}})^2}{\frac{1}{2}\sum_{i=1}^{N}(T_i^{\text{obs}})^2}, \tag{13}$$

where Ω represents reasonable bound constraints on K. The simulated temperature profile $\widehat{T}(K)$ is obtained by numerically solving the heat equation. Since the temperature at the surface has been recorded from the meteorology station, a Dirichlet boundary condition can be easily incorporated.

Because evaluation of the objective function in (13) requires output from a simulation tool, we use sampling methods for the optimization since gradient information is not available. To proceed, we use the same genetic algorithm that was used to fit the sinusoidal boundary condition from the above study. The optimization framework is displayed in Figure 4. At each iteration, the optimizer will pick a set of six K values based on the bounds Ω and previous function evaluations. This vector of K values is then used as input for the numerical solver for the heat equation that outputs the temperature profile. The simulated profile is compared to actual data to obtain the error at the current iteration. The optimization terminates when the error becomes sufficiently small, resulting with the current set of K values as a potential optimal solution.

With the simulation tool in place, we fit the temperature profile in each layer at 24 hours by optimizing conductivities at 1, 5, 10, 15, 20, 25, and 30 cm. As a first attempt, we assumed that the conductivity varies linearly between these locations and was constant from 30 cm to the location of the bottom boundary condition and between the top of the domain and 1 cm. To this end, since the mean subsurface temperature is not known, we also include the temperature at the bottom depth as a decision variable. We used a depth of 70 cm to enforce the bottom boundary condition and used $\Delta z = 0.1$ cm and $\Delta t = 0.1$ minutes. The temperature data used

depth	K (cm/min)
1 cm	$3.2809 \cdot 10^{-2}$
5 cm	$3.1440 \cdot 10^{-1}$
10 cm	$2.6486 \cdot 10^{-2}$
15 cm	$3.9509 \cdot 10^{-1}$
20 cm	$1.9201 \cdot 10^{-1}$
25 cm	$9.8476 \cdot 10^{-3}$
30 cm	$2.6704 \cdot 10^{-1}$

Table 6. Preliminary optimization results, for a temperature of 35.3°C, LSE = $5.9633 \cdot 10^{-5}$ and E = 0.895°C.

over space and time is given above in Figure 1. The temperatures range from about 13 to 67°C in the first 24 hours. Table 6 shows the optimal values obtained for each depth at the 24th hour. The last two rows show the least-squares error (LSE) and the maximum temperature difference (E) over each depth.

The results are promising and the temperature fit can be seen in Figure 5. The maximum error across all depths over time is only 6.2°C, which is a significantly lower than the corresponding first day results in Figure 3. These results confirm that the data likely corresponds to heterogeneous soil. However, in general, this is not known in advance. Thus, it is important to understand the strengths and weaknesses of all methods applied here. To this end, the sensitivity study presented in the next section quantifies how errors in these modeling components impact the overall quality of the inverse problem solution.

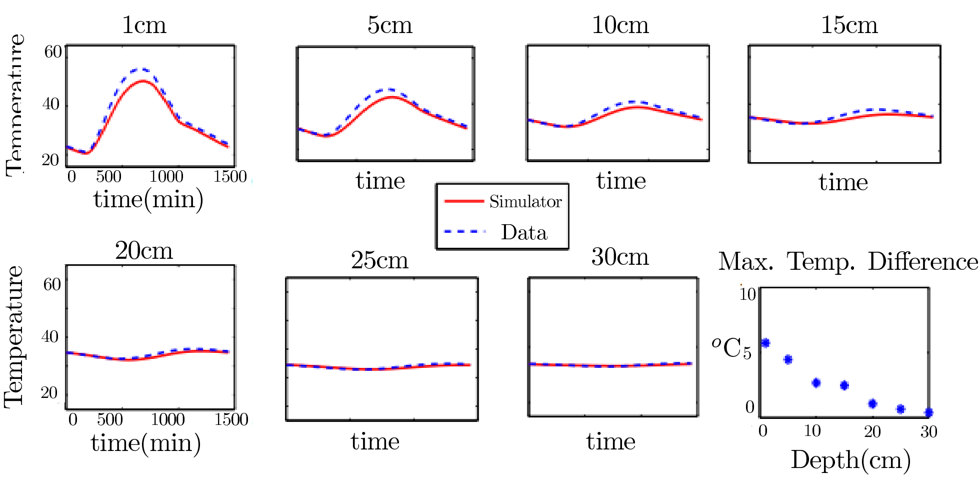

Figure 5. Comparison of simulated temperatures and data at each sensor location over time and maximum temperature difference.

4. Sensitivity analysis

Analysis of variance (ANOVA) is a way to determine whether model parameters have an effect on the model output by comparing the ratio of the variation between sample means to the variation within each sample. For this study, we consider how the parameters in both analytic and simulation-based approaches impact the estimation of K and the model fit. The starting point for the procedure is to sort each parameter into groups. Analysis is done by considering changes in a response as the group changes. Specifically, ANOVA is a hypothesis test with null hypothesis $H_0 = \mu_1 = \mu_2 = \cdots = \mu_k$, where k is the number of experimental groups. Each μ represents the mean of the single parameter, often called a factor, that is being found by the values in each experimental group. When rejecting the null hypothesis, the alternative hypothesis states that at least one mean is different from another; however, it does not specify which one. The experimental groups are different equally spaced intervals for a single variable. The ANOVA examines the source of variation by finding the sum of squares of deviation from the mean for each of these groups. Using a statistical F-test, the procedure is able to determine whether or not at least one mean is deviating from the others. The F-test will produce a p-value; if this value is below a significance of 0.05 then the null hypothesis is rejected. If the significance is above 0.05, the null hypothesis is failed to be rejected. For this work, we seek to understand the sensitivity of parameters for both the analytic and the numerical approaches to matching the temperature data.

4.1. ***Sensitivity analysis of analytic methods.*** Even if a soil sample is homogeneous, there could be errors within the initial sinusoidal fitting of the data due to experimental noise. A sensitivity study can be used to understand how errors in this fitting will impact the resulting temperature profile, in particular, if we consider a hypothetical problem with known model parameters. In other words we sampled variations of the parameters in (7) and (8) and determined how they impacted the ability to identify the conductivity. Specifically, we varied $A_1, A_2, \overline{T}_1, \overline{T}_2, \phi_1$ and ϕ_2 and compared the calculated K to the known value. Using a Latin hypercube sampling (LHS) approach to assure a uniform distribution of selections with intervals surrounding the true values, we considered 1,600 values of each parameter. The bounds used for the LHS sampling are displayed in Table 7 as well as the true parameter value. Following the sampling, the parameters were grouped and an analysis of variance (ANOVA) was performed to show how errors in the initial least-squares fit impact the thermal conductivities from the four analytic methods.

We consider one response for each of the four analytic methods to determine K. These are found by taking the difference between the true conductivity and the conductivity found using the perturbed parameter values. Values for each of the independent parameters were grouped into eight subsets determined by equal sized

parameter	lower bound	upper bound
$A_1 = 5.5974°C$	3°C	7°C
$A_2 = 2.2885°C$	1°C	5°C
$\bar{T}_1 = 20°C$	18°C	22°C
$\bar{T}_2 = 20°C$	18°C	22°C
$\phi_1 = 0.776$	−1	1
$\phi_2 = -0.1880$	−1	1

Table 7. True parameter values and LHS bounds.

ranges within the lower and upper bound of the parameter. ANOVA compares the variance of the objective function within each group to that same variance between the groups. If this ratio is sufficiently small, then the objective function is sensitive to changes in that parameter. This test provides a p-value that establishes a confidence level for sensitivity.

ANOVA results are easily visualized through main effect plots, one developed for each parameter analyzed. Large changes in dependent variable values within each plot show the method is sensitive to changes of that independent parameter. In other words, a flat line means little sensitivity to variation of the parameter value. The vertical axis shows the mean value of the response for values of the parameter of that specific group. A p-value is found to numerically measure the sensitivity, with a p-value close to zero indicating that the parameter is sensitive. The main effects results of the analysis of variance for the amplitude, phase, arctangent and logarithmic methods are shown in Figures 6–9.

As seen, all methods are most sensitive to variations of the amplitude parameter. Thus, errors in estimating the amplitude result in large changes in thermal conductivity results from the four analytic methods. It appears that variations of the other parameters have an impact but are not nearly as significant as variations within the amplitude.

4.2. ***Sensitivity analysis of a heterogeneous system.*** The simulation-based approach uses an optimization algorithm to determine a temperature profile. This technique calls for variation with the bottom boundary condition and seven thermal conductivities. A similar study using ANOVA is conducted to understand the impact of each of these parameters on the model fit by considering the LSE as the output. As with the analytic results, a Latin hypercube sampling is used to sample all parameters. The bounds for the LHS were $[20, 45]$ degrees Celsius for the bottom temperature and $[10^{-4}, 10^{-1}]$ for each thermal conductivity.

Parameters are considered to be sensitive if their corresponding p-value is less than 0.05; p-values are given in Table 8. ANOVA results are displayed visually

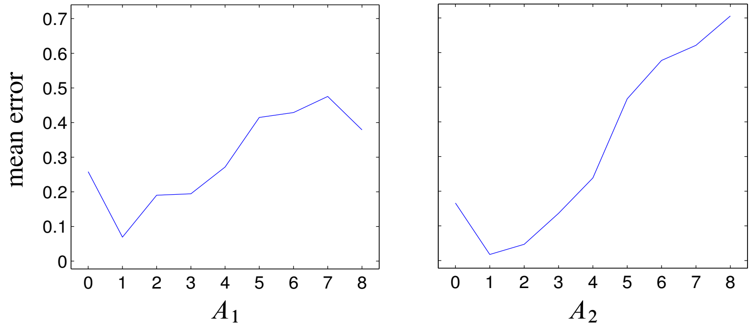

Figure 6. Amplitude method: both A_1, A_2 result in $p \approx 0$.

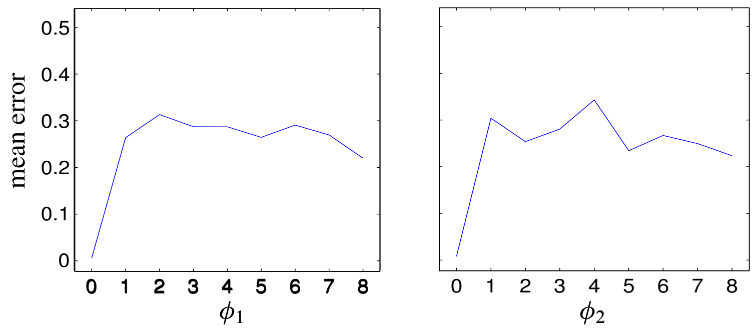

Figure 7. Phase method: both p-values > 0.05.

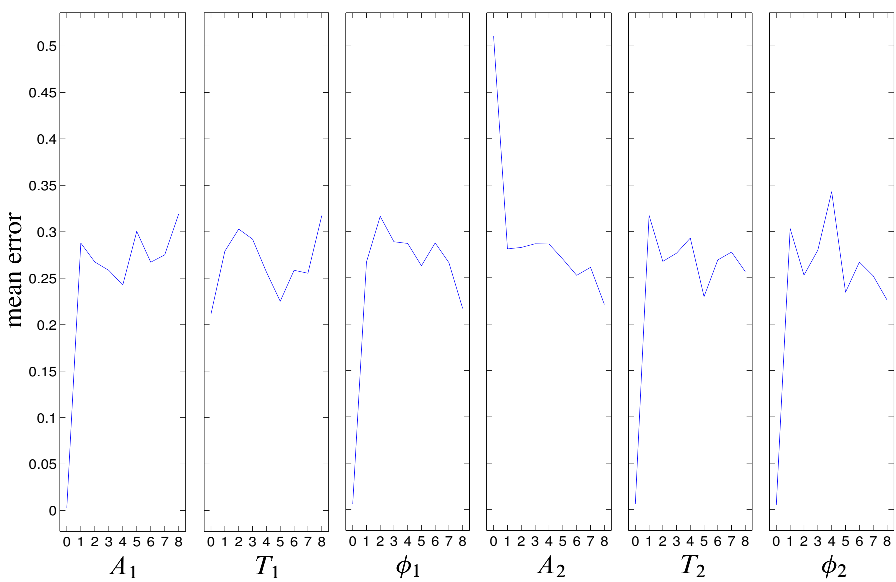

Figure 8. Arctangent method: all p-values > 0.05.

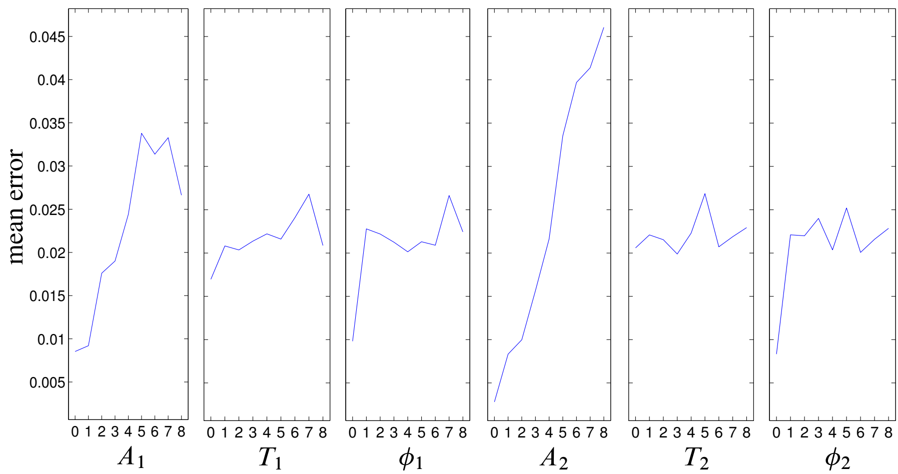

Figure 9. Logarithmic method: both A_1, A_2 result in $p \approx 0$.

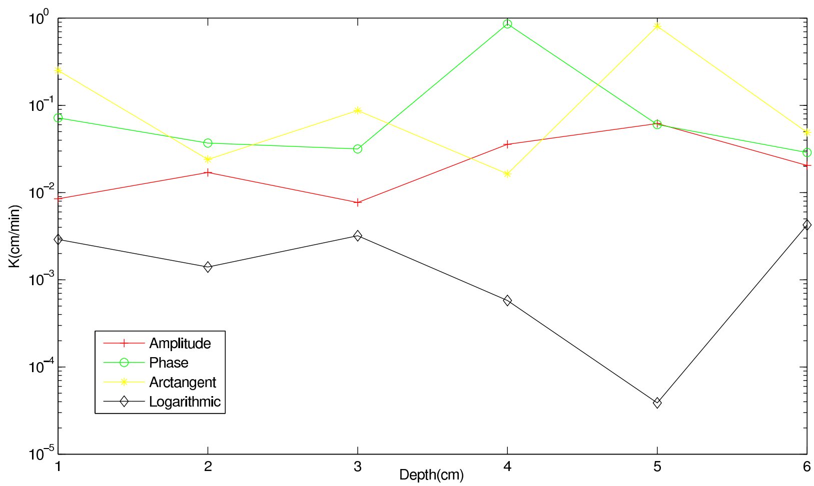

Figure 10. Day 7: conductivities at different depths.

using main effects plots shown in Figure 11. The horizontal axis shows the number of the group, where the intervals above were split into eight equal subintervals. The vertical axis is the average least-squares value corresponding to each group.

From these results, we can see that several of the conductivities are sensitive. Related work analyzes the impact of errors in the boundary and initial data on solution to the inverse problem [Fu and Leventhal 2011]. Here we find that our solutions are not sensitive to this boundary condition. Often in practice, ANOVA is done in advance to understand which model parameters should be included in the optimization and thereby reduce the size of the design space. In this context, the analysis can be used to weight those sensors more heavily in a subsequent optimization study.

parameter	p-value
depth	0
bot. temp (°C)	$6.22207 \cdot 10^{-1}$
K_1	$3.4870 \cdot 10^{-5}$
K_5	$1.4413 \cdot 10^{-1}$
K_{10}	$1.2058 \cdot 10^{-1}$
K_{15}	$2.6220 \cdot 10^{-2}$
K_{20}	$1.3701 \cdot 10^{-7}$
K_{25}	$5.2570 \cdot 10^{-1}$
K_{30}	0

Table 8. ANOVA results.

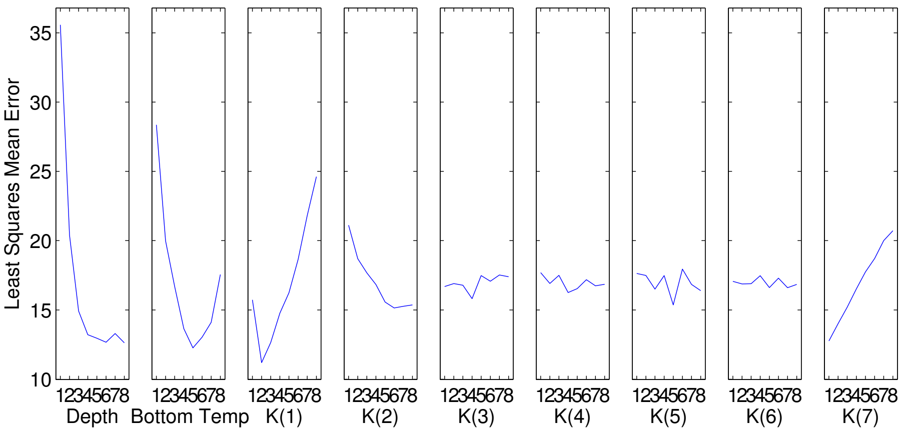

Figure 11. Main effects plots for simulation-based approach.

5. Conclusion

In this work, we have considered an inverse problem to determine the thermal conductivities for a heat transport model using temperature data in the shallow subsurface. Since it is not known in advance if the soil is homogeneous or heterogeneous, analytic and numerical approaches were used. Furthermore, sensitivity analyses can be paired with optimization and modeling problems to help understand how choices made during the solution procedure impact the quality of the results. These ideas provide a protocol for approaching these types of problems.

In this study, not one of the analytics methods for estimating thermal conductivity fit the temperature profile within the given degree of desired accuracy. Each parameter is significant in each method with amplitude being the most significant parameter. Thus small deviations in the amplitude cause large deviations within the resulting thermal conductivity. The amplitude and logarithmic methods display the general trend of temperature values however still not within the desired error. The numerical approach gave satisfactory results and a significantly smaller error than the analytic methods, indicating that this data corresponds to heterogeneous soil.

References

[Carslaw and Jaeger 1986] H. S. Carslaw and J. C. Jaeger, *Conduction of heat in solids*, 2nd ed., Clarendon, Oxford, 1986. Zbl 0584.73001

[Fu and Leventhal 2011] X. Fu and B. Leventhal, "Understanding the impact of boundary and initial condition errors on the solution to a thermal diffusivity inverse problem", *SIAM Undergrad. Res. Online* **4** (2011), 156–174.

[Gao et al. 2009] Z. Gao, L. Wang, and R. Horton, "A comparison of six algorithms to determine the soil thermal diffusivity at a site in the Loess Plateau of China", *Hydrol. Earth Syst. Sci.* **6** (2009), 2247–2274.

[Holland 1973] J. H. Holland, "Genetic algorithms and the optimal allocation of trials", *SIAM J. Comput.* **2** (1973), 88–105. MR 52 #12445a Zbl 0259.90031

[Horton et al. 1983] R. Horton, P. J. Wierenga, and D. R. Nielsen, "Evaluation of methods for determining the apparent thermal diffusivity of soil near the surface", *Soil Sci. Soc. Am.* **47**:1 (1983), 25–32.

[Narasimhan 2009] T. N. Narasimhan, "The dichotomous history of diffusion", *Phys. Today* **62**:7 (2009), 48–53.

[Powers 2006] D. L. Powers, *Boundary value problems and partial differential equations*, 5th ed., Elsevier, Amsterdam, 2006. Chapter 2. MR 2008a:35001 Zbl 1107.35001

Received: 2011-04-28 Revised: 2013-11-11 Accepted: 2013-12-20

leventbc@gmail.com — *Department of Psychology in Education, University of Pittsburgh, 5930 Wesley W. Posvar Hall, Pittsburgh, PA 15260, United States*

rubyfu@mit.edu — *Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States*

kfowler@clarkson.edu — *Department of Mathematics, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699, United States*

owen.j.eslinger@usace.army.mil — *US Army ERDC, 3909 Halls Ferry Road, Vicksburg, MS 39180, United States*

involve

msp.org/involve

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2015 is US $140/year for the electronic version, and $190/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve

2015 vol. 8 no. 3