

## ABSTRACT

FREGOSI, ANNA. Calibration of Thermal Soil Properties in the Shallow Subsurface. (Under the direction of Carl Kelley.)

We use nonlinear least squares methods and Bayesian inference to calibrate soil properties using models for heat and groundwater transport in the shallow subsurface. We first assume a constant saturation in our domain and use the analytic solution to the heat equation as a model for heat transport. We compare our results to those using the finite element code, Adaptive Hydrology (ADH). We then use ADH to simulate heat and groundwater transport in an unsaturated domain. We use the Model-Independent Parameter Estimation (PEST) software to solve the least squares problem with ADH as our model. In using Bayesian inference, we employ the Delayed Rejection Adaptive Metropolis (DRAM) Markov chain Monte Carlo algorithm to sample from the posterior densities of parameters in both models. We find our results are consistent with those found using soil samples with empirical methods.

© Copyright 2015 by Anna Fregosi

All Rights Reserved



Calibration of Thermal Soil Properties in the Shallow Subsurface

by  
Anna Fregosi

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2015

APPROVED BY:

---

Owen Eslinger

---

Ralph Smith

---

Brian Reich

---

Carl Kelley  
Chair of Advisory Committee

## DEDICATION

May this work bring glory to God.

“I can do all this through him who gives me strength.” -Philippians 4:13

## **BIOGRAPHY**

Anna Meade Fregosi was born in Columbia, South Carolina and graduated from Chapin High School in 2005. In 2009, she graduated from Mississippi State University with a degree in mathematics. She began graduate school in 2009 at North Carolina State University and received her Master's degree in mathematics in 2012. She married Daniel Jesse Fregosi in June 2013.

## ACKNOWLEDGEMENTS

This work has been partially supported by National Science Foundation Grants DMS-1406349, CDI-29 0941253, and by Army Research Office Grant W911NF-11-1-0367.

I greatly appreciate the support of many people who have helped me complete this work.

Thank you Gary Howell, Matt Haught, Adam Weast, Tim Penderghest, and Jeff Hensley for patiently helping me develop my computer programming skills. Thank you to others from USACE for supporting me in my research, career, and personal goals: Amanda Hines, Corey Winton, SaBra Neal, Stacy Howington, and especially Owen and Yolanda Eslinger. Thank you Brian Reich and Ralph Smith for your feedback and instruction on the statistical methods in this work. Thank you to my advisor Tim Kelley for your guidance and for the opportunity to work on this project.

Thank you to my classmates and friends at NC State: Anne Costolanski, Deena Giffen, Dave Mokrauer and family, Alyssa Armstrong, Lake Bookman, Melissa Ngamini, Ahlam Elashgh, Colleen McCarthy, Zack Kenz, Erin Solfiell, Li Wang, and Jeb Collins. I appreciate your hospitality, friendship, and support!

Thank you to my roommates Julianne Bower, Erin Bruns, Shirley Law and Addie Loeffler for your support. I also extend great thanks to many, many people at the Brooks Avenue Church of Christ for your generosity in every way. There are too many wonderful people to name. I could not have done this without you.

Thank you to my grandparents Troy Welch and the late Benita Welch, and James and Marion Meade. Your example and sacrifices enabled me to have this opportunity and gave me confidence. Thanks also to John and Claire Morrow, Elizabeth Young, Melissa Barbier, Rachel Strader, Tracy Brewster, Kaitlin Oswalt, Ellen McAngus, Susie Boles and Di Wu for their support!

Thank you to my brother Nathan Meade for his kindness and encouragement and also to

my in-laws Ron Fregosi and Lorrie Lehmann. Thank you all for cheering me on.

Without the sacrifice and constant support of my parents, Tabb and Tena Meade, this work would not have been possible. Thank you. Thank you to my husband, Danny, for all you have done to help me finish this project, particularly for your encouragement at the beginning and for your advice and support as I finish.

Above all, I praise God for allowing me to begin, endure, and finish this work.

# TABLE OF CONTENTS

<b>LIST OF TABLES</b>	ix
<b>LIST OF FIGURES</b>	xi
<b>Chapter 1 Introduction</b>	1
1.1 Background/motivation	1
1.2 Our contributions	3
1.3 Organization	4
<b>Chapter 2 Models</b>	5
2.1 An Analytic Model	5
2.2 A Numerical Model	8
2.2.1 Constant Saturation Model	8
2.2.2 Variable Saturation Model	9
2.2.3 Boundary Conditions	10
2.2.4 Numerical Solution	14
2.3 Other Models	15
2.3.1 Amplitude and Phase Methods	15
2.3.2 Johansen and Lu	15
2.3.3 Discrete Fourier Transform	20
<b>Chapter 3 Calibration Methods</b>	24
3.1 Nonlinear Least Squares	24
3.1.1 Iterative Methods	25
3.2 Bayesian Inference	28
3.2.1 Likelihood	29
3.2.2 Prior Density	30
3.2.3 Posterior Density	33
<b>Chapter 4 Sampling from the Posterior Distribution: Markov Chain Monte Carlo</b>	36
4.1 Monte Carlo Integration	37
4.2 Markov Chains	39
4.2.1 Definitions	39
4.2.2 Properties	40
4.2.3 Convergence Theorems	42
4.3 Markov Chain Monte Carlo	43
4.3.1 Gibbs Sampling	44
4.3.2 Metropolis-Hastings	46
4.3.3 Delayed Rejection Adaptive Metropolis	48
4.3.4 Convergence Diagnostics	52
4.3.5 Prediction Intervals	57

<b>Chapter 5 Results: Calibration of the Apparent Thermal Diffusivity</b>	59
5.1 Data	59
5.2 Least Squares (PEST) with Numerical Model (ADH)	63
5.2.1 Results: Dataset A	66
5.2.2 Results: Dataset B	72
5.2.3 Results: Dataset H	75
5.3 Bayesian Inference (DRAM) with Analytic Model	79
5.3.1 Diagnostics	80
5.3.2 Values for Apparent Thermal Diffusivity	84
5.3.3 Pairwise Scatterplots	85
5.3.4 Prediction Intervals	88
<b>Chapter 6 Results: Calibration of the Thermal Soil Parameters in a Variably Saturated Domain</b>	95
6.1 Data	96
6.1.1 Dataset A	96
6.1.2 Dataset H	98
6.2 Least Squares (PEST) with Numerical Model (ADH)	99
6.2.1 Dataset A	100
6.2.2 Dataset H	100
6.3 Bayesian Inference (DRAM) with Numerical Model (ADH)	104
6.3.1 Simulations with Constant Saturation	104
6.3.2 Variable Saturation	119
<b>Chapter 7 Conclusions</b>	127
<b>REFERENCES</b>	129
<b>APPENDICES</b>	139
Appendix A Data	140
A.1 Instruments	140
A.1.1 Subsurface Instruments	141
A.1.2 Radiometer	142
A.1.3 Air Temperature and Relative Humidity	143
A.1.4 Barometric Pressure	144
A.1.5 Radiation	144
A.1.6 Wind Speed and Direction	144
A.1.7 Precipitation	144
A.1.8 Storage of Data	145
Appendix B Adaptive Hydrology	146
B.1 Implementation	146
B.1.1 Boundary Condition File	146
B.1.2 Mesh File	147
B.1.3 Hotstart Files	147

B.1.4	Met File . . . . .	148
B.1.5	Sup File . . . . .	148
Appendix C	Diagram of Simulations . . . . .	149
C.1	PEST + ADH . . . . .	149
C.2	DRAM + ADH . . . . .	150
C.2.1	run_DRAM_Desert . . . . .	150
C.2.2	model_prior.m . . . . .	150
C.2.3	model_ssq.m . . . . .	150
C.2.4	Prediction Intervals . . . . .	153
C.2.5	Parallel Simulations . . . . .	154
Appendix D	Additional Diagnostics . . . . .	156
D.1	Constant Saturation . . . . .	156
D.1.1	Least Squares (PEST) with Numerical Model (ADH) . . . . .	156
D.1.2	DRAM . . . . .	159
D.2	Variable Saturation . . . . .	164
D.2.1	PEST . . . . .	165
D.2.2	24-hour Constant Saturation Results . . . . .	165
D.2.3	DRAM . . . . .	168



## LIST OF TABLES

2.1	Parameters in thermal conductivity term. . . . .	19
2.2	Known constants in equation for $k$ . . . . .	20
5.1	Six thermal material parameters of interest. . . . .	63
5.2	Additional values used in ADH or to calculate $k$ . . . . .	64
5.3	Date of ADH simulation for each dataset. . . . .	65
5.4	Test cases for Dataset A initial conditions. . . . .	66
5.5	Weights for Dataset A. . . . .	68
5.6	Singular values of Jacobian computed with all initial parameters for Dataset A. . . .	69
5.7	Fixed values and initial values for PEST optimization, Dataset A constant saturation.	70
5.8	Results of PEST optimization for Dataset A. . . . .	70
5.9	Singular values of Jacobian computed with all initial parameters for Dataset B. . . .	72
5.10	Fixed values and initial values for PEST optimization, Dataset B constant saturation.	73
5.11	Results of PEST optimization for Dataset B. . . . .	73
5.12	Weights for Dataset H [35]. . . . .	75
5.13	Fixed values and initial values for PEST optimization, Dataset H constant saturation.	77
5.14	Results of PEST optimization for Dataset H. . . . .	78
5.15	Mean, standard deviation and Monte Carlo standard error (MCSE) for $\alpha$ from Dataset A. . . . .	83
5.16	Statistics of $k$ estimates for Dataset A. Units are in $m^2/hr$ [36]. . . . .	84
5.17	Statistics of $k$ estimates for Dataset B. Units are in $m^2/hr$ [36]. . . . .	85
5.18	Statistics of $k$ estimates for Dataset H. Units are in $m^2/hr$ [36]. . . . .	85
6.1	Singular values of Jacobian computed with all initial parameters for Dataset H, unsaturated case. . . . .	101
6.2	Fixed groundwater parameters for unsaturated case, Dataset H. These values are for both the gravel and deep soil layers. . . . .	102
6.3	Results of PEST optimization for Dataset H, unsaturated case. . . . .	102
6.4	Specific heat optimization results for Dataset H (Fraser/Highlands), unsaturated and dry cases (PEST + ADH). . . . .	103
6.5	Initial conditions for constant saturation 24-hour simulations. . . . .	105
6.6	Results from PEST optimization, Dataset A constant saturation 24-hour case. . . .	106
6.7	Results from PEST optimization, Dataset B constant saturation 24-hour case. . . .	106
6.8	Results from PEST optimization, Dataset H constant saturation 24-hour case. . . .	106
6.9	Mean, standard deviation and Monte Carlo standard error (MCSE) for each param- eter from Dataset A. . . . .	111
6.10	Statistics of $k$ estimates for Dataset A. Units are in $m^2/hr$ [36]. . . . .	117
6.11	Statistics of $k$ estimates for Dataset B. Units are in $m^2/hr$ [36]. . . . .	117
6.12	Statistics of $k$ estimates for Dataset H. Units are in $m^2/hr$ [36]. . . . .	118
6.13	Mean, standard deviation and Monte Carlo standard error (MCSE) for each param- eter from Dataset H. . . . .	122

6.14	Statistics of specific heat estimates for Dataset H. Units are in (W-hr/g-K) [35]. . .	126
D.1	Mean, standard deviation and Monte Carlo standard error (MCSE) for $\alpha$ from Dataset B. . . . .	162
D.2	Mean, standard deviation and Monte Carlo standard error (MCSE) for $\alpha$ from Dataset H. . . . .	164
D.3	Mean, standard deviation and Monte Carlo standard error (MCSE) for each parameter from Dataset B. . . . .	170
D.4	Mean, standard deviation and Monte Carlo standard error (MCSE) for each parameter from Dataset H. . . . .	172

## LIST OF FIGURES

4.1	Example of Markov chain following a trajectory through the parameter space. . . . .	53
4.2	Example of Markov chain rejecting too many values. . . . .	54
4.3	Example of Markov chain convergence to stationary distribution. . . . .	54
4.4	Example of autocorrelation plot with satisfactory rate of decrease. . . . .	56
4.5	Example of autocorrelation plot with high autocorrelation (unsatisfactory). . . . .	56
5.1	Shortwave radiation, longwave radiation and temperature at 1cm depth for Dataset A. The effects of spikes in shortwave radiation are visible in the plot of temperature at 1cm depth. . . . .	61
5.2	Shortwave radiation, longwave radiation and temperature at 1cm depth for Dataset H. . . . .	62
5.3	Temperature data and ADH output for Cases A and B, 6:00 on Julian day 125. . . .	67
5.4	Temperature data and ADH output for Cases A and B, 12:00 on Julian day 125. . .	68
5.5	Dataset A model and temperature data at 1cm and 30cm depth over calibration period. . . . .	71
5.6	Dataset B model and temperature data at 1cm and 30cm depth over calibration period. . . . .	74
5.7	Dataset H model and temperature data at 1cm and 30cm depth over calibration period. . . . .	79
5.8	Chain history from 1 to 3,000 samples. . . . .	80
5.9	Chain history from 1 to 10,000 samples. . . . .	81
5.10	Chain history from 10,000 to 100,000 samples. . . . .	82
5.11	Autocorrelation for burned-in portion of chain. . . . .	83
5.12	Pairwise marginal posterior densities for the model parameters, Dataset A. . . . .	86
5.13	Pairwise marginal posterior densities for the model parameters, Dataset B. . . . .	87
5.14	Pairwise marginal posterior densities for the model parameters, Dataset H. . . . .	88
5.15	Model, data and prediction intervals for Dataset A. . . . .	89
5.16	Model, data and prediction intervals for Dataset B. . . . .	90
5.17	Model, data and prediction intervals for Dataset H. . . . .	91
5.18	Model, data, prediction intervals, and PEST+ADH results for Dataset A. . . . .	92
5.19	Model, data, prediction intervals, and PEST+ADH results for Dataset B. . . . .	93
5.20	Model, data, prediction intervals, and PEST+ADH results for Dataset H. . . . .	94
6.1	Dataset A precipitation and saturation over the ramp-up and calibration periods. . .	97
6.2	Dataset A temperature at 1cm and 15cm depth over the ramp-up and calibration periods. . . . .	97
6.3	Dataset H precipitation and saturation over the ramp-up and calibration periods. . .	98
6.4	Dataset H temperature at 1cm and 15cm depth over the ramp-up and calibration periods. . . . .	99
6.5	Dataset H model and temperature data at 1cm and 30cm depth over calibration period. . . . .	103

6.6	Dataset H saturation 15cm depth over calibration period. . . . .	104
6.7	Dataset A model and temperature data at 1cm and 30cm depth over calibration period. . . . .	107
6.8	Dataset B model and temperature data at 1cm and 30cm depth over calibration period. . . . .	107
6.9	Dataset H model and temperature data at 1cm and 30cm depth over calibration period. . . . .	108
6.10	Dataset A chain histories for the first 10 samples of each individual chain. . . . .	109
6.11	Dataset A chain histories for all eight chains. . . . .	110
6.12	Dataset A chain history for one individual chain. . . . .	110
6.13	Autocorrelation function for Dataset A. . . . .	111
6.14	95% Prediction intervals for Dataset A. . . . .	112
6.15	95% Prediction intervals for Dataset B. . . . .	113
6.16	95% Prediction intervals for Dataset H. . . . .	114
6.17	Pairwise scatterplots from burned-in portion of chain for Dataset A. . . . .	115
6.18	Pairwise scatterplots from burned-in portion of chain for Dataset B. . . . .	115
6.19	Dataset H pairwise plots, constant 24-hour case. . . . .	116
6.20	Estimates for $k$ from all methods for all datasets. . . . .	118
6.21	Estimates for $k$ from all methods for all datasets with log scale on y-axis. . . . .	119
6.22	Dataset H chain histories for the first 5 samples of each individual chain. . . . .	120
6.23	Dataset H chain histories for all eight chains. . . . .	120
6.24	Dataset H chain history for one individual chain. . . . .	121
6.25	Autocorrelation function for Dataset H. . . . .	122
6.26	95% Prediction intervals for Dataset H. Depths of 1 and 30cm. . . . .	123
6.27	95% Prediction intervals for Dataset H. All depths shown. . . . .	124
6.28	Pairwise scatterplots from burned-in portion of chain for Dataset H. . . . .	125
A.1	Met Station in Vicksburg . . . . .	140
A.2	Layout of Instruments . . . . .	141
A.3	Temperature Probe . . . . .	142
A.4	Flux Plate . . . . .	142
A.5	Radiometer . . . . .	143
C.1	Diagram of <code>model_ssq.m</code> . . . . .	151
D.1	Dataset A model and data over all depths for the constant saturation case (PEST+ADH). 157	
D.2	Dataset B model and data over all depths for the constant saturation case (PEST+ADH). 158	
D.3	Dataset H model and data over all depths for the constant saturation case (PEST+ADH). 159	
D.4	Chain history from 1 to 3,000 samples, Dataset B. . . . .	160
D.5	Chain history from 1 to 10,000 samples, Dataset B. . . . .	160
D.6	Chain history from 10,000 to 100,000 samples, Dataset B. . . . .	161
D.7	Autocorrelation for burned-in portion of chain, Dataset B. . . . .	161
D.8	Chain history from 1 to 3,000 samples, Dataset H. . . . .	162
D.9	Chain history from 1 to 10,000 samples, Dataset H. . . . .	163

D.10 Chain history from 10,000 to 100,000 samples, Dataset H. . . . .	163
D.11 Autocorrelation for burned-in portion of chain, Dataset H. . . . .	164
D.12 Dataset H model and temperature data all depths over calibration period, unsaturated case. . . . .	165
D.13 Dataset A model and temperature data all depths over calibration period, constant 24-hour case. . . . .	166
D.14 Dataset B model and temperature data all depths over calibration period, constant 24-hour case. . . . .	167
D.15 Dataset H model and temperature data all depths over calibration period, constant 24-hour case. . . . .	168
D.16 Dataset B entire simulation for all chains, constant saturation 24-hour case. . . . .	169
D.17 Dataset B first few iterations across all chains, constant saturation 24-hour case. . . . .	169
D.18 Dataset B autocorrelation function, constant saturation 24-hour case. . . . .	170
D.19 Dataset H entire simulation for all chains, constant saturation 24-hour case. . . . .	171
D.20 Dataset H first few iterations across all chains, constant saturation 24-hour case. . . . .	171
D.21 Dataset H autocorrelation function, constant saturation 24-hour case. . . . .	172
D.22 95% Prediction intervals for Dataset A, constant saturation 24-hour case with ADH as model. . . . .	173
D.23 95% Prediction intervals for Dataset B, constant saturation 24-hour case with ADH as model. . . . .	174
D.24 95% Prediction intervals for Dataset H, constant saturation 24-hour case with ADH as model. . . . .	175

# Chapter 1

## Introduction

In this work, we use both optimization and Bayesian statistics to perform parameter calibration for heat transport models in the shallow subsurface. We first consider the constant saturation case and then proceed to model heat transport in the variably saturated case. The Bayesian approach to parameter calibration allows us to quantify uncertainty in our problem.

### 1.1 Background/motivation

Models of soil temperature are valuable in disciplines including weather and climate modeling, soil science, civil engineering and remote sensing [36, 66, 78]. We find [11, 18, 51, 96] and [75] particularly useful for gaining an understanding of the thermal properties of soil.

Common models for heat transport in the shallow subsurface are based on the heat equation in one dimension [36, 43, 53, 97]. We detail the heat equation for this context in Chapter 2. Thermal properties of soil are the thermal conductivity, volumetric heat capacity and thermal diffusivity,  $k$ . Since  $k$  is quotient of the conductivity and the volumetric heat capacity, analysis is often focused on one of either the conductivity or diffusivity [32, 43]. Water content plays a significant role in the temperature of the soil, so to have an accurate model for heat transport we must couple our temperature model with a moisture model [35, 36, 66]. Our first results are

based on the assumption of constant water content and are computed with precipitation-free data. In this case, we consider the thermal diffusivity of soils. This property describes the rate at which heat travels through the soil and has units of area over time [75]. By simplifying the subsurface transport processes to conduction only and incorporating thermal effects of the vapor phase, we consider the apparent thermal diffusivity [32, 53, 57, 100].

Many estimates for thermal diffusivity are based on harmonics of the heat equation. Pearce and Gold [76] consider the annual component of the solution and determine diffusivity estimates based on the slopes of the amplitude and phase of variations plotted against depth. Carson [29] estimates diffusivity for the annual and daily temperature cycles based on the amplitude and phase of variations as well. Carson concluded the amplitude and phase methods were unreliable for the diurnal cycle. However, in [100], Wierenga et al. found the methods appropriate under the assumption of a constant water content and clear weather. Results using a finite difference scheme to solve the heat equation together with a simple optimization scheme for estimating  $k$  are presented in [100].

Horton et al. in [53] compares six methods for estimating the apparent thermal diffusivity. The amplitude and phase methods are considered along with two analogous methods using multiple harmonics of the solution to the heat equation: the arctangent and logarithmic methods. The last two methods are the harmonic method, which uses a harmonic expansion of sine terms, and the finite difference methods for solving the heat equation [53]. Horton et al. concludes the amplitude and phase methods are insufficient for estimating  $k$  under cloudy conditions, the numerical method is appropriate when a large amount of data is available, and the harmonic method produces the most reliable estimates given the least amount of data [53]. Methods for using the Laplace transform to estimate  $k$  are discussed in [95] and [32]. Different forms of boundary conditions for the heat equation are evaluated in [86], with the conclusion that the choice of boundary condition is important. Finally, methods for empirical estimates for the thermal conductivity are developed in [57] and [66].

In this work we also consider the specific heat of the soil. The volumetric heat capacity is

the product of the density and specific heat of the soil. The specific heat may also be referred to as the heat capacity. Early methods for estimating the specific heat of a soil involved heating combinations of water and soil and using a calorimeter to determine changes in temperature [19, 41, 62, 85, 90]. In [41], Fritton analyzes error in these methods. Kersten [62] performs a thorough analysis of soil properties using a calorimeter and determines that the methods for estimating specific heat using mixtures of soil and water is appropriate. Kersten [62] also determines a proportional relationship between specific heat and temperature. This relationship is also determined in [60], describing a way to determine specific heat using temperature measurements and known values for specific heat.

Methods for estimating the specific heat using heat-pulse probes are described in [20, 22, 23, 27, 63, 64, 73, 74, 83, 89]. Least squares and maximum likelihood estimate methods for estimating the specific heat using heat-pulse data are described in [17, 21, 99]

The effect of surface fires on thermal properties of soil are explored in [26, 69, 70]. Finally, Yadav, in [103], found that specific heat is independent of compaction of soils.

## 1.2 Our contributions

We collaborate with the U.S. Army Engineer Research and Development Center (ERDC) in this work. We use data from three different locations to calibrate our model and parameters. The data was collected over a 40-day period and includes measurements of temperature above and below the surface, precipitation, radiation, wind speed and direction, humidity, pressure, and subsurface heat flux. In addition, we utilize soil samples collected from each site to determine estimates for  $k$ . My specific contributions are listed below.

- The Adaptive Hydrology (AdH) model, created at ERDC, is a 3-D finite element simulation for heat transport. I determine scalar estimates for  $k$  using AdH coupled with PEST parameter estimation software. We find these estimates to be consistent with estimates using other methods. In the variably saturated case, I use the same methods to determine



an estimate for the specific heat of the soil.

- I use the framework of Bayesian inference to analyze the probability densities for the apparent thermal diffusivity and specific heat in varied domains. This contribution allows uncertainty quantification of the heat transport model. Using the Bayesian approach, I compare scalar estimates for  $k$  with those found using other methods as well as compute prediction intervals for future observations.

There are several differences between the two calibration methods listed above. The first calibration method, using PEST, stems from the frequentist view of statistics. The parameters are considered fixed and unknown. We optimize to produce scalar values for individual parameters. The second calibration method reflects the Bayesian view of statistics. The parameters are considered unknown and random. We describe the parameters, given data, by their probability distributions.

### 1.3 Organization

In Chapter 2 we describe the models we use for heat transport and groundwater in the shallow subsurface. In Chapter 3 we describe the least squares and Bayesian inference calibration methods. We detail Markov Chain Monte Carlo methods in Chapter 4. In Chapters 5 and 6 we present and discuss our results for the domains with constant and variable saturation, respectively.

## Chapter 2

# Models

In this chapter we describe two models for heat transport in the shallow subsurface. The first model is the analytic solution to the heat equation, assuming a constant water content, with Dirichlet boundary conditions. The Dirichlet boundary condition at the surface accommodates the diurnal temperature cycle. We assume the temperature approaches a constant average temperature as depth increases. The second model is the numerical solution to a more general form of the heat equation with a Neumann boundary condition at the surface. In this model the Neumann boundary condition at the surface accounts for shortwave and longwave radiation, sensible heat exchange, latent heat exchange and precipitation [35]. This more general boundary condition provides a more realistic model. In Section 2.3, we mention similar models to which we will later compare our results.

### 2.1 An Analytic Model

We use Fourier’s law of heat conduction to develop our model for heat transport. The amount of heat conducted across a unit area for a unit amount of time is the heat flux density,  $q$  ( $\text{J m}^{-2} \text{s}^{-1}$ ) [11, 97]. Fourier’s law says that the heat flux  $q$  in a homogeneous medium is proportional

to the temperature gradient  $\nabla u$  ( $\text{K m}^{-1}$ ) and opposite in direction [18, 50, 97]:

$$q = -\kappa \nabla u. \quad (2.1)$$

The constant of proportionality  $\kappa$  ( $\text{J m}^{-1} \text{s}^{-1} \text{K}^{-1}$ ) in (2.1) is the thermal conductivity of the medium. It is the quantity of heat conducted per unit time, per unit area [11]. The thermal conductivity of water is about  $0.6 \text{ J m}^{-1} \text{s}^{-1} \text{K}^{-1}$  [66, 78] and the thermal conductivity of soil is  $\mathcal{O}(1) \text{ J m}^{-1} \text{s}^{-1} \text{K}^{-1}$  [66].

The specific heat  $c_p$  ( $\text{J kg}^{-1} \text{K}^{-1}$ ) per unit mass is the rate of change of amount of heat in a unit mass of soil with respect to temperature change of one unit [50, 97]. The product of  $c_p$  and the mass per unit volume or density,  $\rho$  ( $\text{kg m}^{-3}$ ), is the volumetric heat capacity. It is the amount of heat per unit volume needed to change temperature by one unit. The rate of change of the quantity of heat in a unit volume with respect to time is thus

$$(c_p \rho) \frac{\partial u}{\partial t} \quad (2.2)$$

The principle of the conservation of energy gives us the form of the heat equation for a homogenous soil [50, 97]. For soil with volumetric heat capacity  $(c_p \rho)$  and thermal conductivity  $\kappa$ , we write the heat equation in a homogeneous domain  $\Omega$ :

$$(c_p \rho) u_t = \nabla \cdot (\kappa \nabla u), \quad (2.3)$$

where  $u$  is temperature as a function of time  $t$  and space [35].

The homogeneity assumption implies  $(c_p \rho)$  and  $\kappa$  do not depend on space or time. Thus we can combine them to form the equation

$$u_t = k \nabla^2 u, \quad (2.4)$$

where  $k = \frac{\kappa}{c_p \rho}$  ( $\text{m}^2 \text{s}^{-1}$ ) is the thermal diffusivity. Values for soil thermal diffusivity are  $\mathcal{O}(10^{-7})$   $\text{m}^2 \text{s}^{-1}$  [53].

We now focus on the 1-D heat conduction equation in a seminfinite domain:

$$u_t = k u_{zz}, \quad (t, z) \in (0, T] \times [0, \infty), \quad (2.5)$$

where  $z$  is depth from the surface  $z = 0$ .

We assume the temperature oscillates with frequency  $\phi$  about an average value  $\hat{C}$  at all depths [50, 97]. Thus, at the surface, the temperature is [97]

$$u(t, 0) = C \sin(2\pi\phi t + \omega) + \hat{C}, \quad (2.6)$$

where  $C$  is the amplitude of the temperature oscillations at the surface and  $\omega$  is a phase shift. We assume the temperature approaches the constant  $\hat{C}$  as  $z \rightarrow \infty$  [97]. The assumption of the harmonic behavior of the temperature with time implies that we do not need to specify an initial condition for  $t = 0$  [97].

The solution to (2.5) for the given boundary conditions and assumptions is [36, 97]:

$$u(t, z) = C e^{-\alpha z} \sin(2\pi\phi t - \alpha z + \omega) + \hat{C}, \quad (2.7)$$

where  $\alpha > 0$  is a damping parameter.

We only consider the diurnal oscillations and fix  $\phi = 1/24$  hours. Thus our analytic model for heat transport in the subsurface is

$$u(t, z) = C e^{-\alpha z} \sin\left(\frac{\pi}{12}t - \alpha z + \omega\right) + \hat{C}. \quad (2.8)$$

We relate  $\alpha$  to the soil thermal diffusivity through

$$\alpha = \sqrt{\frac{\pi}{24k}}. \quad (2.9)$$

We will often denote  $\theta = (C, \alpha, \omega, \hat{C})$  as our set of model parameters.

## 2.2 A Numerical Model

The Adaptive Hydrology (AdH) model produced by the US Army Corps of Engineers (USACE) simulates both heat transport and subsurface flow by simulating the solution to a more general form of the heat equation using finite elements.

### 2.2.1 Constant Saturation Model

For a 3-D soil domain with constant saturation the equation for heat transport, as seen in (2.3), is

$$(c_p \rho) u_t - \nabla \cdot (\kappa \nabla u) = 0. \quad (2.10)$$

We explain the numerical model in terms of a heterogenous domain. The soil is composed of a solid phase, liquid phase (water) and gaseous phase (air) [33]. In practice, the domain may be divided into multiple materials with different properties, where heat transport in each material is governed by the equations in this section.

We define the volumetric heat capacity ( $c_p \rho$ ) by the volumetric heat capacities of the domain components. Letting the subscripts  $s$ ,  $w$  and  $g$  denote soil, water and air, respectively and letting  $\eta_i$  be the volume fraction for each component, we have [33]

$$(c_p \rho) = \eta_s c_{p,s} \rho_s + \eta_w c_{p,w} \rho_w + \eta_g c_{p,g} \rho_g. \quad (2.11)$$

The pore space is the space between particles or aggregates of the solid phase. The fraction of pore space volume in the total medium volume is the porosity,  $\phi$  [18, 80, 98]. If water completely fills the pore space, we say the soil is saturated. Otherwise, the pore space is filled with a mixture of air and water and is unsaturated. The fraction of water volume present with respect to the volume of the pore space is the degree of saturation or normalized water content

$S_w$  [51]. In terms of water content and porosity we have [35]:

$$(c_p \rho) = (1 - \phi)c_{p,s}\rho_s + \phi(S_w c_{p,w}\rho_w + (1 - S_w)c_{p,g}\rho_g). \quad (2.12)$$

We also define the thermal conductivity with a mixture model. Johansen in [57] defines the dimensionless Kersten number  $K_e$  as a normalized conductivity that depends only on saturation. If  $\kappa_d$  is the thermal conductivity of dry soil and  $\kappa_s$  is the thermal conductivity of saturated soil, then [35, 66]

$$\kappa = (\kappa_s - \kappa_d)K_e + \kappa_d. \quad (2.13)$$

Details regarding Equation (2.13) are in Subsection 2.3.

The thermal diffusivity is thus

$$k = \frac{(\kappa_s - \kappa_d)K_e + \kappa_d}{(1 - \phi)c_{p,s}\rho_s + \phi(S_w c_{p,w}\rho_w + (1 - S_w)c_{p,g}\rho_g)}. \quad (2.14)$$

### 2.2.2 Variable Saturation Model

For unsaturated domains, the equation for heat transport must be coupled with the equation for moisture transport through the saturation  $S_w$ . We again use the heat equation (2.10):

$$(c_p \rho)u_t - \nabla \cdot (\kappa \nabla u) = 0.$$

Now, however, the volumetric heat capacity  $c_p \rho$  depends on the changing saturation  $S_w$  in the following way (as seen in Equation (2.12)):

$$(c_p \rho) = (1 - \phi)c_{p,s}\rho_s + \phi(S_w c_{p,w}\rho_w + (1 - S_w)c_{p,g}\rho_g). \quad (2.15)$$

Thus, we first solve for the saturation  $S_w$  and then use  $S_w$  in (2.15) to determine the volumetric heat capacity.

The saturation is a function of the pressure head  $\psi$  [35, 54, 92], which is the height of a column of water relative to some reference point [52, 98]. We use the van Genuchten model [24, 35, 59, 92, 94] to relate saturation and pressure head

$$S_w(\psi) = S_r + \frac{(1 - S_r)}{[1 + (\alpha|\psi|)^n]^m}, \quad \psi \leq 0, \quad (2.16)$$

where  $S_r$  is residual saturation and  $\alpha$ ,  $m$ ,  $n$  depend on the domain. We note that  $m = 1 - \frac{1}{n}$ . The parameters  $m$  and  $\alpha$  depend on the slope of the saturation-pressure head curve [94]. The parameter  $n$  relates to the pore-size distribution.

Soil moisture transport also depends upon the hydraulic conductivity,  $K$  of the soil. The hydraulic conductivity describes the flow rate of a liquid through a porous medium [35, 52, 102]. If the soil is saturated, we describe  $K_s$ , the saturated hydraulic conductivity. The relative permeability  $k_r(\psi)$  is related to the soil structure and also describes flow through the soil [35, 37]. We use the Mualem and van Genuchten models to describe the relationship [35, 59, 94] between relative permeability and pressure head:

$$k_r(\psi) = \frac{[1 - (\alpha|\psi|)^{n-1}[1 + (\alpha|\psi|)^n]^{-m}]^2}{[1 + (\alpha|\psi|)^n]^{m/2}}, \quad \psi \leq 0. \quad (2.17)$$

We use the mixed form of Richards' equation for moisture transport [35, 54, 92]

$$S_s S_w(\psi) \frac{\partial \psi}{\partial t} + \phi \frac{\partial S_w(\psi)}{\partial t} = \nabla [K_s k_r(\psi) \nabla (\psi + z)], \quad (2.18)$$

where the specific storage  $S_s$  describes the fluid compressibility [92],  $z$  is depth and  $t$  is time.

### 2.2.3 Boundary Conditions

We specify boundary conditions for (2.10) which depend on shortwave and longwave radiation, sensible heat exchange, latent heat exchange and precipitation [35]. For our numerical simulations, we incorporate data into these boundary conditions.

- **Radiation** We first explain a few concepts of solar radiation. A black body is an object which absorbs all radiation it receives [13, 80]. The amount of energy  $F$  a black body at temperature  $T$  emits is given by the Stefan-Boltzmann Law [13, 18, 35]:

$$F = \sigma T^4, \quad (2.19)$$

where  $\sigma = 5.67 \times 10^{-8}$  (W m<sup>-2</sup> K<sup>-4</sup>) is the Stefan-Boltzmann constant [13].

Energy from the sun, which acts as a black body [13], reaches the earth in the form of shortwave solar radiation [13]. There are two types of solar irradiance: direct and diffuse. Direct irradiance is not scattered or absorbed before it reaches the surface of the earth, while diffuse irradiance has been scattered by the atmosphere [13, 80]. The total global solar radiation is the sum of both the direct and diffuse radiation.

Now we describe the interaction between solar radiation and the earth. The earth absorbs some of the shortwave solar radiation and re-emits some of it in the form of longwave radiation [13, 80], because the temperature of the earth is much lower than that of the sun. The shortwave radiation not absorbed by the ground surface is reflected [13]. The albedo  $\alpha_g$  of the ground surface is the fraction of shortwave radiation reflected [13, 35, 80]. Thus the amount of shortwave radiation absorbed by the surface is [35]

$$SW^\downarrow = (1 - \alpha_g)SW_{total}. \quad (2.20)$$

We use  $SW_{total} = SW_{measured}$ , the measured shortwave radiation (data). There is no shortwave solar radiation at night.

Nonblack bodies emit energy at a rate [18]

$$F = \epsilon \sigma T^4, \quad (2.21)$$

where  $0 < \epsilon < 1$  is a property of the material called the emissivity and  $T$  is the temperature



of the nonblack body. For a soil with emissivity  $\epsilon_g$  and ground temperature  $T_g$ , the total longwave radiation emitted is [35]

$$LW^\uparrow = \epsilon_g \sigma T_g^4. \quad (2.22)$$

We use measured temperature values for  $T$ . Values of  $\epsilon_g$  are usually between 0.9 and 0.95 [13].

Longwave radiation emitted from the atmosphere and clouds is also absorbed by the surface [35]:

$$LW^\downarrow = \epsilon_g LW, \quad (2.23)$$

We use  $LW = LW_{measured}$ , the measured longwave radiation (data). The earth continues to absorb and emit longwave energy at night.

- **Sensible Heat Exchange** Sensible heat exchange is the exchange of heat between the surface and the air which does not include chemical processes [35]. The sensible heat flux is proportional to the product of the temperature difference, air density  $\rho_a$  and air specific heat  $c_{p,a}$  [75]:

$$H \propto \rho_a c_{p,a} (T_g - T_a), \quad (2.24)$$

where  $T_a$  is air temperature at a height  $Z$  above the ground and  $T_g$  is ground temperature [35]. The thermal resistance to the sensible heat flux depends on the wind speed,  $W_s$  and the surface aerodynamic roughness [35, 75]. We use the logarithmic wind profile equation in our formulation [75] and define the sensible heat exchange at the surface by

$$H = \left( W_0 + \left( \frac{\rho_a c_{p,a}}{S_c} \right) \left( \frac{K}{\log(\frac{Z}{Z_0})} \right)^2 W_s \right) (T_g - T_a), \quad (2.25)$$

with constants: windless coefficient  $W_0$ , Schmidt number  $S_c$ , von Karman constant  $K$  and roughness coefficient  $Z_0$  [35].

- **Latent Heat Exchange** Latent heat is absorbed or released during phase changes such as evaporation, condensation, freezing and melting [13, 35, 56]. The latent heat of water vaporization,  $\lambda$  ( $\text{J kg}^{-1}$ ), is the energy released or absorbed during condensation or evaporation [35, 75]. The latent heat flux at the surface is the product of the mass flux,  $E$  ( $\text{kg m}^{-2} \text{s}^{-1}$ ), due to evaporation and the latent heat of water vaporization [75]:

$$LE = \lambda E. \quad (2.26)$$

The mass flux  $E$  is related to the change in specific humidity,  $\Delta q_h$  [75]:

$$E \propto \rho_a \Delta q_h. \quad (2.27)$$

The specific humidity depends on the vapor pressure  $e$  (kPa) and atmospheric pressure  $P$  (kPa) [75]. We define  $q_h$  ( $\text{kg kg}^{-1}$ ) by [75]

$$q_h \approx \frac{0.622e}{P}, \quad (2.28)$$

where we use the ratio of the molecular mass of water to the molecular mass of dry air, 0.622, [10, 12]. Thus for vapor pressures at the surface  $e_0$  and at some height above the surface,  $e_a$ , [75]

$$E \propto -\rho_a \frac{0.622}{P} (e_0 - e_a). \quad (2.29)$$

For soil surface resistance  $r_0$  [35]:

$$LE = \frac{(0.622)\rho_a\lambda}{r_0P} (e_0 - e_a). \quad (2.30)$$

**Precipitation** Precipitation, *Precip*, also plays a role in the exchange of heat at the surface [35]. We express the thermal effects of precipitation as the product of the measured precipitation, specific heat of air, and air temperature [15, 35, 58].

$$PH = (Precip)c_{p,a}T_a \quad (2.31)$$

The preceding terms all contribute to the heat flux at the surface of our domain. Our flux boundary condition is [35]

$$Flux = SW^\downarrow + LW^\downarrow - LW^\uparrow + H + LE + PH. \quad (2.32)$$

For the numerical simulations, we specify a Dirichlet boundary condition at the bottom of our domain. This boundary condition may be the output of a previous simulation or a value chosen by the analyst.

For the groundwater portion of the numerical model, we specify a Neumann flow boundary condition of precipitation rate at the surface and a Dirichlet boundary condition of total head at the bottom of our domain.

#### 2.2.4 Numerical Solution

The ADaptive Hydrology (ADH) numerical model for heat and groundwater transport in the shallow subsurface is produced by the US Army Corps of Engineers [14,15,36,79]. ADH is a finite element code for saturated and unsaturated flow in 2-D and 3-D [15]. For 3-D simulations, ADH uses linear tetrahedral elements on an unstructured conforming mesh with continuous linear Lagrange basis functions [15, 35]. ADH also features a mesh refinement scheme [15, 35, 79]. Time integration is performed using implicit Euler with an adaptive step size [84]. The user may select a maximum step size. ADH may be implemented in serial or parallel [79].

For details regarding the implementation of ADH, see Appendix B. We refer to [45] for a details of the finite element method.

## 2.3 Other Models

As mentioned in Chapter 1, there are other ways to determine estimates for  $k$ . We later compare our results using the methods in Sections 2.1 and 2.2 to some of these other methods. Here we briefly explain the amplitude and phase methods, the two empirical methods of Lu and Johansen, and the use of the discrete Fourier transform in modeling heat transport.

### 2.3.1 Amplitude and Phase Methods

Many estimates for thermal diffusivity are based on harmonics of the heat equation. Consider the amplitude,  $C_i$ , of the temperature fluctuations at depth  $z_i$  over a period  $P$ . For data at depths  $z_1$  and  $z_2$  with amplitudes  $C_1$  and  $C_2$ , respectively, the thermal diffusivity can be estimated by the amplitude method [29]:

$$k = \frac{\pi(z_1 - z_2)^2}{P \left( \ln \left( \frac{C_1}{C_2} \right) \right)^2}. \quad (2.33)$$

Consider also the maxima of temperature fluctuations over a period  $P$ . The phase method requires the times  $t_1$  and  $t_2$  of maxima at depths  $z_1$  and  $z_2$  [29, 36]:

$$k = \frac{P(z_2 - z_1)^2}{4\pi(t_2 - t_1)^2}. \quad (2.34)$$

### 2.3.2 Johansen and Lu

Our estimates of  $k$  from experiments depend on data from the meteorological station, soil samples and expert opinion. Here we detail the two approaches we use in determining these experimental estimates. We first describe the formulation of both the Johansen [57] and Lu [66] models for soil thermal conductivity. Then we describe our use of these models to determine estimates for  $k$ .

## Formulation

As explained before, the thermal diffusivity

$$k = \frac{\kappa}{c_p \rho} \quad (2.35)$$

is the quotient of the thermal conductivity  $\kappa$  and volumetric heat capacity  $c_p \rho$ .

We use the methods of Lu [66] and Johansen [57] to estimate the thermal conductivity  $\kappa$ . Both models have the form

$$\kappa = (\kappa_{sat} - \kappa_{dry})K_e + \kappa_{dry}, \quad (2.36)$$

where  $\kappa_{sat}$  and  $\kappa_{dry}$  are the saturated and dry thermal conductivities, respectively, and the Kersten number  $K_e$  depends on saturation and soil texture.

The saturated thermal conductivity

$$\kappa_{sat} = \kappa_s^{1-\phi} \kappa_w^\phi \quad (2.37)$$

is determined by the thermal conductivity of the solid,  $\kappa_s$ , the thermal conductivity of water,  $\kappa_w$ , and porosity  $\phi$  [57, 66, 78]. The thermal conductivity of the solids

$$\kappa_s = \kappa_q^q \kappa_o^{1-q}, \quad (2.38)$$

where

$$\kappa_o = \begin{cases} 2.0 \text{ Wm}^{-1}\text{K}^{-1}, & q > 0.2 \\ 3.0 \text{ Wm}^{-1}\text{K}^{-1}, & q \leq 0.2. \end{cases} \quad (2.39)$$

is determined by the thermal conductivity of quartz  $\kappa_q = 7.7 \text{ Wm}^{-1}\text{K}^{-1}$ , quartz fraction,  $q$ , and thermal conductivity of other materials,  $\kappa_o$  [57, 66].

Lu and Johansen have different models for  $K_e$  and  $\kappa_{dry}$ .

## Johansen

For the Johansen model,  $K_e$  depends on the saturation  $S_w$ :

$$K_e = \begin{cases} 0.7 \log S_w + 1.0, & (S_w > 0.05), \text{ coarse soils} \\ \log S_w + 1.0, & (S_w > 0.1), \text{ fine soils.} \end{cases} \quad (2.40)$$

Our soils are classified as coarse soils.

The dry thermal conductivity

$$\kappa_{dry} = \frac{0.135\rho_b + 64.7}{2700 - 0.947\rho_b} \quad (2.41)$$

depends on the bulk soil density  $\rho_b$  [57].

## Lu

For the Lu model, the Kersten number

$$K_e = \exp\{\alpha_s[1 - S_w^{(\alpha_s - 1.33)}]\} \quad (2.42)$$

depends on saturation and a soil texture parameter  $\alpha_s$  [66]. If the the percentage of sand and gravel is greater than 0.4 then the soil is coarse textured [66]. Following [66], we define  $\alpha_s$  as

$$\alpha_s = \begin{cases} 0.96, & \text{coarse-textured soils} \\ 0.27, & \text{fine-textured soils.} \end{cases} \quad (2.43)$$

The dry thermal conductivity is given by the model [66]

$$\kappa_{dry} = -a\phi + b, \quad (2.44)$$

where  $a = 0.56$  and  $b = 0.51$  are found empirically for values of porosity between 0.2 and 0.6.

## Experiments

The values we use in calculating  $k$  are compiled in Tables 2.1 and 2.2. Table 2.1 contains values derived from the met station data, soil samples and expert opinion. Known physical constants are compiled in Table 2.2.

Core soil samples were analyzed in an ERDC lab as well as by Daniel B. Stephens & Associates, Inc. (DBS&A). DBS&A took measurements at a range of depths. The cores were also flooded, air dried and oven dried in order to obtain measurements at different saturation values.

Table 2.1: Parameters in thermal conductivity term.

Symbol	Description	Source
$\phi$	Porosity	DBS&A (Highlands); average from DBS&A over all samples and depths (Desert); DBS&A or ERDC lab (Tropical)
$q$	Quartz Fraction	Expert (Tropical, Desert); Median value of range given by expert (Highlands [35])
$S_w$	Saturation	Probe data water content rounded to nearest whole number percentage, multiplied by lab porosity (Tropical); DBS&A (Desert and Highlands)
$\rho_s$	Specific Gravity (Density) of Solid	Calibrated using DBS&A data (Tropical, Highlands); average from ERDC or DBS&A (Desert)
$c_{p,s}$	Specific Heat of Solid	Calibrated using DBS&A data (Tropical, Desert, Highlands)
	Percentage of Sand and Gravel	Median value of range given by expert (Highlands [35]), DBS&A (Desert), ERDC lab (Tropical)
$K_e$	Kersten number	Depends on $S_w$ , $q$ , $\alpha_s$ [66]
$\kappa_{dry}$	Dry Thermal Conductivity	Depends on density $\rho_b$ [57] or porosity $\phi$ [66]
$\kappa_{sat}$	Saturated Thermal Conductivity	Depends on $\phi$ , $q$
$\rho_b$	Dry/bulk density	Depends on $\phi$ and $\rho_s$
$\alpha_s$	Soil texture	Coarse (from percentage of sand and gravel) value of 0.96 [66]



Table 2.2: Known constants in equation for  $k$ .

Symbol	Description	Units	Value/Source
$\kappa_w$	Thermal Conductivity of Water	(W/m-K)	(0.58) [25]
$c_{p,w}$	Specific Heat of Water	(W-hr/g-K)	(0.001161472) [48]
$\rho_w$	Specific Gravity (Density) of Water	(g/cm <sup>3</sup> )	(10 <sup>6</sup> ) [33]
$c_{p,g}$	Specific Heat of Gas	(W-hr/g-K)	(0.000281111) [9]
$\rho_g$	Specific Gravity (Density) of Gas	(g/cm <sup>3</sup> )	(1000) Value from [33] rounded to one significant figure

We estimate  $k$  using Equations 2.35-2.44 and the values in Tables 2.1 and 2.2. The results are in Section 5.3.2, Tables 5.16-5.18. The results for the Lu and Johansen models are labelled Experiments A and B, respectively.

### 2.3.3 Discrete Fourier Transform

We also compare our results to those found in [36] using a discrete Fourier transform to determine estimates for  $k$ . In this approach, we again consider the 1-D heat conduction equation in a semi-infinite domain, Equation (2.5). We assume the boundary condition at the surface is a sum of sinusoids and

$$\lim_{z \rightarrow \infty} u(t, z) = 0. \quad (2.45)$$

With these boundary conditions, the solution becomes [36]

$$u(t, z) = \sum_{j=0}^N u_j, \quad (2.46)$$

for

$$u_j(t, z) = C_j e^{(-\alpha_j z)} \sin(2\pi\phi_j t - \alpha_j z + \omega_j). \quad (2.47)$$

We only consider frequencies  $\phi_j$  that can be written  $\phi_j = j\phi$ , where  $\phi$  is  $\frac{1}{24}$  hours. Using trigonometric identities, we can write

$$\begin{aligned} u_j(t, z) &= C_j e^{(-\alpha_j z)} \sin(-\alpha_j z + \omega_j) \cos(2\pi\phi_j t) \\ &\quad + C_j e^{(-\alpha_j z)} \cos(-\alpha_j z + \omega_j) \sin(2\pi\phi_j t). \end{aligned} \quad (2.48)$$

We set

$$A_j = C_j e^{(-\alpha_j z)} \sin(-\alpha_j z + \omega_j) \quad (2.49)$$

$$B_j = C_j e^{(-\alpha_j z)} \cos(-\alpha_j z + \omega_j) \quad (2.50)$$

so that (2.48) becomes

$$u_j(t, z) = A_j \cos(2\pi\phi_j t) + B_j \sin(2\pi\phi_j t). \quad (2.51)$$

At each depth  $z_i$ , we use the data and a discrete Fourier transform to determine the Fourier coefficients  $A_{i,j}$  and  $B_{i,j}$ . With the new coefficients  $A_{i,j}$  and  $B_{i,j}$ , (2.46) for a fixed depth  $z_i$  becomes

$$u(t, z_i) = \sum_{j=0}^N \left( A_{i,j} \cos(2\pi\phi_j t) + B_{i,j} \sin(2\pi\phi_j t) \right). \quad (2.52)$$

To compute estimates for  $k$ , we use

$$\frac{A_{i,j}}{B_{i,j}} = \frac{\sin(-\alpha_j z_i + \omega_j)}{\cos(-\alpha_j z_i + \omega_j)} = \tan(-\alpha_j z_i + \omega_j), \quad (2.53)$$

to define  $\gamma_{i,j} = \arctan\left(\frac{A_{i,j}}{B_{i,j}}\right) = -\alpha_j z_i + \omega_j$ . Then, we compute

$$E_{i,j} = \frac{A_{i,j}}{\sin(\gamma_{i,j})} = C_j e^{(-\alpha_j z_i)}. \quad (2.54)$$

Using the natural logarithm, Equation (2.54) becomes

$$\ln(E_{i,j}) = \ln(C_j e^{(-\alpha_j z_i)}) = \ln(C_j) + \ln(e^{(-\alpha_j z_i)}) = \ln(C_j) + (-\alpha_j z_i). \quad (2.55)$$

We use linear regression to approximate  $\alpha_j$  and the relation

$$\alpha_j = \sqrt{\frac{\pi \phi_j}{k}} \quad (2.56)$$

to determine estimates for  $k$  [36].

We compare our results to the results in [36], where  $\phi_j$  is fixed at  $\frac{1}{24}$ .

## Resampling

We later compare our results to those in [36] determined by resampling each dataset.

To compute an estimate using amplitude and phase methods, Equations (2.33) and (2.34) respectively, we use data at two depths. We have temperature probe data at seven depths, allowing  ${}^7C_2$  possible pairs of depths. We only consider 24-hour time periods. If the dataset contains  $N$  days, we generate  $N$  estimates for  $k$  for a fixed pair of depths. Overall, we compute  $N({}^7C_2)$  estimates for  $k$ .

To compute an estimate using the DFT-based method, we use data from 2-7 depths. Thus, in this case, the number of possible combinations of depths is

$${}^7C_2 + {}^7C_3 + {}^7C_4 + {}^7C_5 + {}^7C_6 + {}^7C_7. \quad (2.57)$$

We also use between 1 and  $N$  days of data to compute an estimate. Thus, the number of

consecutive 24-periods is  $\left(\sum_{j=1}^N j\right)$ . The total number of estimates of  $k$  using the DFT-based method for each dataset is [36]

$$\left(\sum_{j=1}^N j\right)\left({}^7C_2 + {}^7C_3 + {}^7C_4 + {}^7C_5 + {}^7C_6 + {}^7C_7\right). \quad (2.58)$$

Resampling with the DFT-based method produces more estimates for  $k$  than does resampling with the amplitude and phase methods.

## Chapter 3

# Calibration Methods

In this chapter we discuss two methods for calibrating our model parameters: nonlinear least squares and Bayesian inference. The first approach uses traditional least squares methods to find parameter values that minimize the residual sum of squares. The Bayesian approach involves constructing a distribution for the parameters themselves.

### 3.1 Nonlinear Least Squares

With the nonlinear least squares approach to parameter calibration, we use deterministic methods to obtain fixed scalar or vector solutions to the nonlinear least squares problem. The nonlinear least squares problem is to minimize the objective function [61]:

$$f(\theta) = \frac{1}{2} \sum_{i=1}^n (r_i(\theta))^2 = \frac{1}{2} R(\theta)^T R(\theta), \quad (3.1)$$

where  $R(\theta)$  is the residual vector  $r_i(\theta) = y_i - u_i(\theta)$  and  $n$  is the number of observations. We say  $p$  is the number of parameters and write the gradient of  $f$  as [61]

$$\nabla f(\theta) = (\partial f / \partial \theta_1, \dots, \partial f / \partial \theta_p) = R'(\theta)^T R(\theta), \quad (3.2)$$

where  $R'(\theta)$  is the  $n \times p$  Jacobian of  $R$  [61]:

$$(R'(\theta))_{ij} = \frac{\partial r_i}{\partial \theta_j}, \quad 1 \leq i \leq n, 1 \leq j \leq p. \quad (3.3)$$

The Hessian of  $f$  is the  $p \times p$  matrix [61]

$$\nabla^2 f(\theta) = \left( \frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \right) = R'(\theta)^T R'(\theta) + R''(\theta)^T R(\theta). \quad (3.4)$$

We note that a matrix  $A$  is positive semidefinite if

$$\theta^T A \theta \geq 0 \text{ for all } \theta \in R^p, \quad (3.5)$$

positive definite if

$$\theta^T A \theta > 0 \text{ for all } \theta \in R^p, \theta \neq 0 \quad (3.6)$$

and symmetric if  $A = A^T$ .

The necessary conditions for optimality say that at a local minimizer  $\theta^*$  of  $f$  [61]

$$\nabla f(\theta^*) = R'(\theta^*)^T R(\theta^*) = 0 \quad (3.7)$$

and  $\nabla^2 f(\theta^*)$  is positive semidefinite.

### 3.1.1 Iterative Methods

We might use an iterative method like Newton's method to find a local solution of (3.7), which is a local minimizer of  $f$ . For Newton's method to converge, our initial iterate  $\theta_0$  must be sufficiently close to  $\theta^*$  [61]. We update the current iterate,  $\theta_c$  with a new iterate  $\theta_+$  by

$$\theta_+ = \theta_c - (\nabla^2 f(\theta_c))^{-1} \nabla f(\theta_c) \quad (3.8)$$

or, in terms of the residual,

$$\theta_+ = \theta_c - [R'(\theta_c)^T R'(\theta_c) + R''(\theta_c)^T R(\theta_c)]^{-1} (R'(\theta_c)^T R(\theta_c)). \quad (3.9)$$

Since the term  $R''(\theta_c)^T R(\theta_c)$  is computationally expensive, we consider the Gauss-Newton update [61]:

$$\theta_+ = \theta_c - (R'(\theta_c)^T R'(\theta_c))^{-1} R'(\theta_c)^T R(\theta_c). \quad (3.10)$$

The limitations of the Gauss-Newton method are that we must have a good initial iterate, a small residual at the true solution,  $R(\theta^*)$ , and  $R'(\theta_c)^T R'(\theta_c)$  must be nonsingular, meaning  $n \geq p$  [61].

### Line Search Methods

We can accommodate for the initial iterate limitation of the Gauss-Newton method using line search algorithms. These methods update the iterations in descent directions  $d$  using the Armijo rule to determine the step size. The general form of the iteration is

$$\theta_+ = \theta_c + \lambda d, \quad (3.11)$$

where  $\lambda$  is called the steplength [61].

**Definition 3.1** *A vector  $d \in R^p$  is a descent direction for  $f$  at  $\theta$  if [61]*

$$\left. \frac{df(\theta + td)}{dt} \right|_{t=0} = \nabla f(\theta)^T d < 0. \quad (3.12)$$

For example, the steepest descent direction is  $d = -\nabla f(\theta)$  and the method of steepest descent is defined by the update:

$$\theta_+ = \theta_c - \lambda \nabla f(\theta_c). \quad (3.13)$$

We tune  $\lambda$  to adjust for sufficient decrease throughout the simulation. We define sufficient

decrease in the steepest descent method by the Armijo rule:

$$f(\theta_c - \lambda \nabla f(\theta_c)) - f(\theta_c) < -\alpha \lambda \|\nabla f(\theta_c)\|^2, \quad (3.14)$$

where  $\alpha$  is a parameter such that  $\alpha \in (0, 1)$ .

More generally, we find a descent direction  $d = \theta - \theta_c$  that minimizes  $m(\theta)$  for the quadratic model

$$m(\theta) = f(\theta_c) + \nabla f(\theta_c)^T (\theta - \theta_c) + \frac{1}{2} (\theta - \theta_c)^T H_c (\theta - \theta_c), \quad (3.15)$$

where  $H_c$  is a symmetric positive definite matrix we call the model Hessian [61]. The descent direction  $d = \theta - \theta_c$  is then [61]

$$d = -H_c^{-1} \nabla f(\theta_c). \quad (3.16)$$

Note that if  $H_c = I$ , where  $I$  is the  $p \times p$  identity matrix,  $d$  is the steepest descent direction.

For general line search algorithms we define sufficient decrease as

$$f(\theta_c + \lambda d) - f(\theta_c) < -\alpha \lambda \nabla f(\theta_c)^T d. \quad (3.17)$$

In the context of nonlinear least squares, the Gauss-Newton direction is [61]

$$d^{GS} = -(R'(\theta_c)^T R'(\theta_c))^{-1} R'(\theta_c)^T R(\theta_c). \quad (3.18)$$

We note that  $\nabla f(\theta)^T d^{GS} < 0$  and the Gauss-Newton direction is a descent direction when  $R$  has full column rank [61]. Using the Gauss-Newton direction together with the Armijo rule results in the damped Gauss-Newton method:

$$\theta_+ = \theta_c + \lambda d^{GS} = \theta_c - \lambda (R'(\theta_c)^T R'(\theta_c))^{-1} R'(\theta_c)^T R(\theta_c). \quad (3.19)$$

In order for the convergence theory for line search methods to hold,  $R'(\theta_c)^T R'(\theta_c)$  must be uniformly bounded and well-conditioned [61]. We ensure this is the case with the use of the



Levenberg-Marquardt method:

$$\theta_+ = \theta_c - (\nu_c I + R'(\theta_c)^T R'(\theta_c))^{-1} R'(\theta_c)^T R(\theta_c), \quad (3.20)$$

where  $\nu > 0$  is a regularization parameter called the Levenberg-Marquardt parameter and  $I$  is the  $p \times p$  identity matrix [61]. The Levenberg-Marquardt method paired with a line search is:

$$\theta_+ = \theta_c - \lambda(\nu_c I + R'(\theta_c)^T R'(\theta_c))^{-1} R'(\theta_c)^T R(\theta_c). \quad (3.21)$$

We solve the nonlinear least squares problem for  $k$  using the Model-Independent Parameter Estimation (PEST) optimizer [34]. PEST uses the Gauss-Marquardt-Levenberg algorithm [34].

## 3.2 Bayesian Inference

With the Bayesian approach to parameter calibration, we consider the parameters as unknown random quantities [28] and express our knowledge about the parameters in the form of probability distributions. Bayesian inference is the use of Bayes' Theorem and data to draw conclusions about our parameters. For observations  $Y$  and parameters  $\theta$ , we form the likelihood  $f(Y|\theta)$  which expresses the probability of obtaining  $Y$  given  $\theta$ . We express our prior knowledge about the parameters, independent from  $Y$ , in the prior density  $p(\theta)$ . Using Bayes' Theorem, we form the posterior density of the parameters given our observations,  $p(\theta|Y)$ .

**Theorem 1** (*Bayes' Theorem*)

$$p(\theta|Y) = \frac{f(Y|\theta)p(\theta)}{\int f(Y|\theta)p(\theta)d\theta}. \quad (3.22)$$

This is the form of Bayes' Theorem for Bayesian inference.

### 3.2.1 Likelihood

The likelihood  $f(Y|\theta)$  expresses the probability of obtaining observations  $Y$  given model parameters  $\theta$ .

We determine the form of this density by assuming, for simplicity, a perfect model  $u(\theta)$ :

$$y_i = u(\theta)_i + \epsilon_i, \quad (3.23)$$

where  $y_i$  is the observation at spatiotemporal point  $i$ ,  $u(\theta)_i$  is the model evaluated with parameters  $\theta$  at spatiotemporal point  $i$  and  $\epsilon_i$  is the measurement error at spatiotemporal point  $i$ .

We also assume that the errors in our data are independent and normally distributed, with mean 0 and weighted variance:

$$\epsilon_i \sim N\left(0, \frac{\sigma^2}{w_i^2}\right), \quad (3.24)$$

where  $\sigma^2$  is the unweighted variance in our data and  $\epsilon_i$  is the error at data point  $i$  for  $i = 1, \dots, n$ . We weight our data because the temperature fluctuations at the surface of our domain are much greater than those at the bottom of our domain.

These two assumptions give us the form of the likelihood:

$$y_i|\theta, \sigma^2 \sim N\left(u(\theta)_i, \frac{\sigma^2}{w_i^2}\right). \quad (3.25)$$

If we were to make different assumptions about the distribution of errors in our observations or the accuracy of our model, our sampling distribution would have a different form.

We can write the likelihood as

$$f(y_i|\theta, \sigma^2) = \frac{1}{\sqrt{\frac{2\pi\sigma^2}{w_i^2}}} \exp\left\{\frac{-1}{\frac{2\sigma^2}{w_i^2}}(y_i - u(\theta)_i)^2\right\}. \quad (3.26)$$

Since we assume our errors are independent, we can write the joint likelihood:

$$f(Y|\theta, \sigma^2) = \prod_{i=1}^n \left[ \frac{1}{\sqrt{\frac{2\pi\sigma^2}{w_i^2}}} \exp\left\{ \frac{-1}{\frac{2\sigma^2}{w_i^2}} (y_i - u(\theta)_i)^2 \right\} \right] \quad (3.27)$$

$$= \prod_{i=1}^n \left[ \frac{\sqrt{w_i^2}}{\sqrt{2\pi\sigma^2}} \right] \exp\left\{ \sum_{i=1}^n \frac{-w_i^2}{2\sigma^2} (y_i - u(\theta)_i)^2 \right\} \quad (3.28)$$

### 3.2.2 Prior Density

Our prior density expresses the information we have about our parameters before we consider the data. Prior densities fall into two categories: conjugate and non-conjugate. A conjugate prior leads to a posterior density in the same family of distributions (or conjugate family of distributions) as the prior [31]. A non-conjugate prior leads to a posterior density with no known form. We choose a conjugate or non-conjugate prior by considering the form of the likelihood. If the product of the likelihood and the prior has a known form, the prior is conjugate.

Consider the following example from [31]. Suppose  $\theta = (\mu, \sigma^2)$  and we have the likelihood

$$y_1, \dots, y_n | \mu, \sigma^2 \stackrel{iid}{\sim} N(\mu, \sigma^2) \quad (3.29)$$

so,

$$f(Y|\mu, \sigma^2) = \prod_{i=1}^n \left[ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-1}{2\sigma^2} (y_i - \mu)^2 \right\} \right] \quad (3.30)$$

$$\propto \frac{1}{(\sigma^2)^{n/2}} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right\}. \quad (3.31)$$

$$(3.32)$$

In Equation (3.31) we use the proportional form as we are focused on the kernel of the product  $f(Y|\theta)p(\theta)$  [31].

Suppose also that we have the priors  $p(\mu) \sim N(a, b^2)$  and  $p(\sigma^2) \sim InvGamma(c, d)$  and

that  $\mu$  and  $\sigma^2$  are independent. Then we have the joint posterior density

$$p(\mu, \sigma^2 | Y) \propto \left( \frac{1}{(\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right\} \right) \quad (3.33)$$

$$\dots \left( \exp \left\{ -\frac{1}{2b^2} (\mu - a)^2 \right\} \right) \left( (\sigma^2)^{-(c+1)} e^{-d/\sigma^2} \right). \quad (3.34)$$

The form of  $p(\mu, \sigma^2 | Y)$  is not recognizable as a known distribution, so in the bivariate sense these priors are not conjugate. However, since  $\mu$  and  $\sigma^2$  are independent, we are able find recognizable forms for each of their full conditional densities. First we consider  $p(\mu | Y, \sigma^2)$  and let  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ :

$$p(\mu | Y, \sigma^2) \propto \left( \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right\} \right) \left( \exp \left\{ -\frac{1}{2b^2} (\mu - a)^2 \right\} \right) \quad (3.35)$$

$$\propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 - \frac{1}{2b^2} (\mu - a)^2 \right\} \quad (3.36)$$

$$\propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i^2 - 2\mu y_i + \mu^2) - \frac{1}{2b^2} (\mu^2 - 2a\mu + a^2) \right\} \quad (3.37)$$

$$\propto \exp \left\{ \frac{1}{\sigma^2} \sum_{i=1}^n (\mu y_i) - \frac{1}{2\sigma^2} n\mu^2 - \frac{1}{2b^2} \mu^2 + \frac{1}{b^2} a\mu \right\} \quad (3.38)$$

$$\propto \exp \left\{ \frac{1}{\sigma^2} \mu n\bar{y} - \frac{1}{2\sigma^2} n\mu^2 - \frac{1}{2b^2} \mu^2 + \frac{1}{b^2} a\mu \right\} \quad (3.39)$$

$$\propto \exp \left\{ \left( -\frac{1}{2\sigma^2} n - \frac{1}{2b^2} \right) \mu^2 + \left( \frac{1}{\sigma^2} n\bar{y} + \frac{1}{b^2} a \right) \mu \right\}. \quad (3.40)$$

$$(3.41)$$

In order to complete the square to find the form of a normal distribution, we let

$$A = -\frac{n}{2\sigma^2} - \frac{1}{2b^2} \quad \text{and} \quad B = \frac{\frac{n\bar{y}}{\sigma^2} + \frac{a}{b^2}}{-\frac{n}{2\sigma^2} - \frac{1}{2b^2}}. \quad (3.42)$$

Then

$$p(\mu|Y, \sigma^2) \propto \exp\left\{A\left(\mu^2 + B\mu\right)\right\} \quad (3.43)$$

$$\propto \exp\left\{A\left(\mu^2 + B\mu + \frac{B^2}{4} - \frac{B^2}{4}\right)\right\} \quad (3.44)$$

$$\propto \exp\left\{A\left(\left(\mu + \frac{B}{2}\right)^2 - \frac{B^2}{4}\right)\right\} \quad (3.45)$$

$$\propto \exp\left\{A\left(\mu + \frac{B}{2}\right)^2\right\}. \quad (3.46)$$

$$(3.47)$$

Note that if  $\mu \sim N(M, S^2)$ , then

$$A = \frac{-1}{2S^2} = \left(-\frac{n}{2\sigma^2} - \frac{1}{2b^2}\right) = \frac{-1}{2}\left(\frac{n}{\sigma^2} + \frac{1}{b^2}\right), \quad (3.48)$$

and

$$S^2 = \frac{1}{\left(\frac{n}{\sigma^2} + \frac{1}{b^2}\right)}. \quad (3.49)$$

For the posterior mean  $M$ ,

$$-M = \frac{B}{2} = \frac{1}{2} \frac{\frac{n\bar{y}}{\sigma^2} + \frac{a}{b^2}}{-\frac{n}{2\sigma^2} - \frac{1}{2b^2}} = \frac{\frac{n\bar{y}}{\sigma^2} + \frac{a}{b^2}}{-\frac{n}{\sigma^2} - \frac{1}{b^2}} = -\frac{\frac{n\bar{y}}{\sigma^2} + \frac{a}{b^2}}{\frac{n}{\sigma^2} + \frac{1}{b^2}}, \quad (3.50)$$

so

$$M = \frac{\frac{n\bar{y}}{\sigma^2} + \frac{a}{b^2}}{\frac{n}{\sigma^2} + \frac{1}{b^2}}. \quad (3.51)$$

Thus we have the full conditional density

$$\mu|\sigma^2, Y \sim N\left(\frac{\frac{n\bar{y}}{\sigma^2} + \frac{a}{b^2}}{\frac{n}{\sigma^2} + \frac{1}{b^2}}, \frac{1}{\frac{n}{\sigma^2} + \frac{1}{b^2}}\right). \quad (3.52)$$

The derivation of  $p(\sigma^2|\mu, Y)$  is simpler. We multiply the prior density and likelihood to find

the form of an inverse gamma distribution:

$$p(\sigma^2|\mu, Y) \propto \left( \frac{1}{(\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right\} \right) \left( (\sigma^2)^{-(c+1)} e^{-d/\sigma^2} \right) \quad (3.53)$$

$$\propto (\sigma^2)^{-\frac{n}{2}-(c+1)} \exp \left\{ -\frac{d}{\sigma^2} - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right\} \quad (3.54)$$

$$\propto (\sigma^2)^{-((\frac{n}{2}+c)+1)} \exp \left\{ -\frac{1}{\sigma^2} \left( d + \frac{1}{2} \sum_{i=1}^n (y_i - \mu)^2 \right) \right\}. \quad (3.55)$$

Thus,

$$\sigma^2|\mu, Y \sim \text{InvGamma} \left( \frac{n}{2} + c, d + \frac{1}{2} \sum_{i=1}^n (y_i - \mu)^2 \right). \quad (3.56)$$

So, we are able to use conjugate priors for  $\mu$  and  $\sigma^2$  to find known forms of their full conditional posterior densities.

For complicated models  $u(\theta)$ , it may impossible to choose a conjugate prior for the model parameters  $\theta$  given the likelihood (3.25). However,  $\sigma^2 \sim \text{InvGamma}(a, b)$  is a conjugate prior for  $\sigma^2$  with likelihood (3.25).

### 3.2.3 Posterior Density

The likelihood and prior density are the two terms we need to compute the posterior density using Bayes' Theorem. This computation can be done analytically or numerically by using quadrature or sampling methods like Markov Chain Monte Carlo.

We note that the denominator  $p_Y(Y) = \int f(Y|\theta)p(\theta)d\theta$  in Equation (3.22) is the marginal distribution of  $Y$  and serves as a normalizing constant for the posterior density. Since  $p_Y(Y)$  does not depend on  $\theta$ , we may obtain the same *information* about  $\theta$  using the proportional relation [16]

$$p(\theta|Y) \propto f(Y|\theta)p(\theta). \quad (3.57)$$

## Posterior Summaries

From the posterior density we may compute posterior statistics.

We often consider the expectation or posterior mean,

$$\hat{\theta} = E[\theta|Y] = \int \theta p(\theta|Y) d\theta \quad (3.58)$$

and posterior variance,

$$Var \theta = E[(\theta - \hat{\theta})^2] = \int (\theta - \hat{\theta})^2 p(\theta|Y) d\theta = \int \left( \theta - \left[ \int \theta p(\theta|Y) d\theta \right] \right)^2 p(\theta|Y) d\theta. \quad (3.59)$$

If  $\theta$  is a vector  $\theta = (\theta_1, \dots, \theta_p)$ , we take the expectation elementwise [31]:

$$E[\theta|Y] = \begin{pmatrix} E[\theta_1|Y] \\ \vdots \\ E[\theta_p|Y] \end{pmatrix} \quad (3.60)$$

and consider the covariance matrix instead of the scalar variance.

The posterior median for a scalar  $\theta$  is the value  $\tilde{\theta}$  such that

$$0.5 = \int_{-\infty}^{\tilde{\theta}} p(\theta|Y) d\theta. \quad (3.61)$$

We may find, for example, a 95% probability interval  $[a, b]$  for  $\theta$  by solving for  $a$  and  $b$ :

$$0.95 = \int_a^b p(\theta|Y) d\theta. \quad (3.62)$$

It is common [31] to choose  $a$  and  $b$  such that

$$0.025 = \int_{-\infty}^a p(\theta|Y) d\theta \quad (3.63)$$

$$0.975 = \int_{-\infty}^b p(\theta|Y)d\theta. \quad (3.64)$$

All of the above quantities involve integration over the parameter space and may be difficult or impossible if the parameter space has high dimension. In this case, simulation methods for sampling from the posterior density can be used to approximate the posterior summaries using sample statistics.



## Chapter 4

# Sampling from the Posterior

# Distribution: Markov Chain Monte Carlo

Evaluating Bayes' Theorem can be impossible for high dimensional parameter spaces. In using Markov Chain Monte Carlo (MCMC) methods, we construct a Markov chain whose stationary distribution is the posterior distribution and then sample it. With this sample we are able to approximate posterior summaries and the kernel density of the posterior without integrating over the parameter space.

We begin this chapter with a discussion of Monte Carlo integration and then consider the theory behind Markov chains. Next we explain how the two ideas are combined for Markov Chain Monte Carlo simulations and discuss Gibbs Sampling, the Metropolis-Hastings algorithm and the Delayed Rejection Adaptive Metropolis algorithm.

## 4.1 Monte Carlo Integration

Suppose we want to evaluate the expected value of a random variable  $\theta$  with density  $p(\theta)$ . That is, we want

$$E(\theta) = \int \theta p(\theta) d\theta. \quad (4.1)$$

Monte Carlo integration is an alternative to evaluating this integral analytically, which may be difficult if  $\theta$  has high dimension. With Monte Carlo integration, we use a random sample from  $p(\theta)$  to approximate  $E(\theta)$  by the sample mean,  $\bar{\theta}^m$ . This approximation is justified by the Law of Large Numbers.

**Theorem 2 Law of Large Numbers (LLN).** *Let  $\theta^1, \theta^2, \dots$  be independent and identically distributed random variables with density  $p(\theta)$ . Then,*

$$\bar{\theta}^m \equiv \frac{1}{m} \sum_{i=1}^m \theta^i \xrightarrow{p} E(\theta) \equiv \int \theta p(\theta) d\theta. \quad (4.2)$$

*The arrow superscripted with  $p$  indicates convergence in probability, or, for every  $\epsilon > 0$ , [30, 31]*

$$\lim_{m \rightarrow \infty} P\left(|\bar{\theta}^m - E(\theta)| < \epsilon\right) = 1. \quad (4.3)$$

Furthermore, we know the error,  $\bar{\theta}^m - E(\theta)$ , approaches the normal distribution with mean zero. This idea is expressed in the Central Limit Theorem.

**Theorem 3 Central Limit Theorem.** *Let  $\theta^1, \theta^2, \dots$  be independent and identically distributed random variables with  $E\theta^i = \mu$  and  $0 < \text{Var } \theta^i = \sigma^2 < \infty$ . Define  $\bar{\theta}^m = \frac{1}{m} \sum_{i=1}^m \theta^i$ . Let  $g_m(\theta)$  denote the cdf of  $\sqrt{m}(\bar{\theta}^m - \mu)/\sigma$ . Then, for any  $\theta$ ,  $-\infty < \theta < \infty$ ,*

$$\lim_{m \rightarrow \infty} g_m(\theta) = \int_{-\infty}^{\theta} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy; \quad (4.4)$$

that is,  $\sqrt{m}(\bar{\theta}^m - \mu)/\sigma$  has a limiting standard normal distribution [30].

The central limit theorem tells us that the estimate of the mean converges to the true mean at a rate  $\mathcal{O}(\frac{1}{\sqrt{m}})$  [87, 88]. That is,  $\text{Var}(\bar{\theta}^m - \mu) = \frac{\sigma^2}{m} \rightarrow 0$  as  $m \rightarrow \infty$ .

Furthermore, we can approximate other summaries of  $p(\theta)$  with a random sample from  $p(\theta)$ . For example, the (biased) sample variance,  $S_m^2$  converges in probability to  $\sigma^2$ , the variance of  $p(\theta)$  [31], where

$$S_m^2 = \frac{1}{m} \sum_{i=1}^m (\theta^i - \bar{\theta}^m)^2. \quad (4.5)$$

Similarly, the (biased) sample standard deviation converges in probability to the true standard deviation [31],

$$\sqrt{\frac{1}{m} \sum_{i=1}^m (\theta^i - \bar{\theta}^m)^2} \xrightarrow{p} \text{sd}(\theta) = \sigma. \quad (4.6)$$

We can approximate the probability that  $\theta$  is in the interval  $[1, \infty)$  by [31]

$$\frac{1}{m} \sum_{i=1}^m I_{[1, \infty)}(\theta^i) \xrightarrow{p} E(I_{[1, \infty)}(\theta)), \quad (4.7)$$

where  $I_{[a, b]}$  is an indicator function such that

$$I_{[a, b]} = \begin{cases} 1, & \text{if } \theta \in [a, b] \\ 0, & \text{if } \theta \notin [a, b]. \end{cases} \quad (4.8)$$

Similar results hold for the median and interval approximations, though they cannot be written explicitly in terms of integrals [31]. For the median,

$$\text{med}\{\theta^i : i = 1, 2, \dots, m\} \xrightarrow{p} \text{med}(\theta) \equiv \tilde{\theta}, \quad (4.9)$$

where  $\tilde{\theta}$  is the number such that

$$0.5 = \int_{-\infty}^{\tilde{\theta}} p(\theta) d\theta = \int_{\tilde{\theta}}^{\infty} p(\theta) d\theta. \quad (4.10)$$

For probability intervals, say  $\hat{\theta}_{0.025}$ ,  $\hat{\theta}_{0.975}$  are the 95% sample percentiles and  $\theta_{0.025}$ ,  $\theta_{0.975}$  are the true percentiles. We have [31],

$$0.95 = P(\theta_{0.025} < \theta < \theta_{0.975}) = \int_{\theta_{0.025}}^{\theta_{0.975}} p(\theta) d\theta \approx \int_{\hat{\theta}_{0.025}}^{\hat{\theta}_{0.975}} p(\theta) d\theta. \quad (4.11)$$

By plotting a histogram of the sample, we get an idea of the shape of the density. We can also use the sample to approximate the density itself using kernel density estimation [31, 87]. Kernel density estimation is a process of *smoothing* the histogram by using a basis of kernel densities [31, 87].

We have yet to discuss how we obtain a random sample from the density  $p(\theta)$ .

## 4.2 Markov Chains

We first define a Markov chain and discuss notation. Then we consider properties of Markov chains and convergence to stationary distributions.

### 4.2.1 Definitions

A sequence of random variables  $\theta^0, \theta^1, \theta^2, \dots$  is a Markov chain if

$$Pr(\theta^k \in A | \theta^0, \theta^1, \dots, \theta^{k-1}) = Pr(\theta^k \in A | \theta^{k-1}), \quad (4.12)$$

for an arbitrary set  $A$  [31].

This definition is also true for random vectors. The idea is that the distribution of  $\theta^k$  depends only on the distribution of  $\theta^{k-1}$  [30, 87]. This is known as the Markov property [31].

Following the notation in [31], let

$$q_{k|k-1}(\theta^k|\theta^0, \theta^1, \dots, \theta^{k-1}) = q_{k|k-1}(\theta^k|\theta^{k-1}) \quad (4.13)$$

be the conditional density of  $\theta^k$  and  $q_k(\theta^k)$  be the marginal density of  $\theta^k$ .

As in [31], we see:

$$Pr(\theta^k \in A) = \int_A q_k(\theta) d\theta \quad (4.14)$$

$$= \int_A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} q_{k|k-1}(\theta^k|\theta^{k-1}) q_{k-1|k-2}(\theta^{k-1}|\theta^{k-2}) \quad (4.15)$$

$$\dots q_{1|0}(\theta^1|\theta^0) q_0(\theta^0) d\theta^0 \dots d\theta^{k-2} d\theta^{k-1} d\theta^k, \quad (4.16)$$

where  $q_0(\theta^0)$  is the initial density of  $\theta^0$ .

If  $Pr(\theta^k \in A|\theta^{k-1})$  is independent of  $k$ , we say the Markov chain is homogeneous and drop the subscript for  $q$ :

$$q_{k|k-1}(\theta^k|\theta^{k-1}) = q(\theta^k|\theta^{k-1}). \quad (4.17)$$

We call the density  $q(\cdot|\cdot)$  the transition kernel. For example, if  $k = 2$  we can write:

$$Pr(\theta^2 \in A) = \int_A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(\theta^2|\theta^1) q(\theta^1|\theta^0) q_0(\theta^0) d\theta^0 d\theta^1 d\theta^2. \quad (4.18)$$

In this work we assume the Markov chains are homogeneous. Note that [31], even if the chain is homogeneous and the transition kernel does not depend on  $k$ , the marginal densities of the  $\theta^k$  still may depend on  $k$ .

#### 4.2.2 Properties

The distribution  $p(\cdot)$  is a stationary distribution if, for any  $k$ , [31]

$$Pr(\theta^k \in A) = \int_A p(\theta) d\theta. \quad (4.19)$$

Markov chains with certain properties have unique stationary distributions. The following definitions provide the framework for the convergence theory.

### **Irreducibility**

Let  $p$  be a stationary distribution of a Markov chain and  $A$  be a set such that  $\int_A p(\theta)d\theta > 0$ . If for any initial state the probability of reaching  $A$  at some point is positive, the Markov chain is  $p$ -irreducible [31, 91].

### **Recurrence**

Let  $A$  be any set such that  $\int_A p(\theta)d\theta > 0$ . If a chain is  $p$ -irreducible with stationary distribution  $p$  and the probability of revisiting  $A$  at some time is one, then the Markov chain is recurrent. The chain is positive recurrent if the expected value of the time it takes for the chain to return to  $A$  is finite [31, 42, 88].

If for all initial states

$$Pr(\theta^k \in A \text{ infinitely often}) = 1 \quad (4.20)$$

the Markov chain is Harris recurrent [31, 91].

### **Detailed Balance Condition**

For a distribution  $p(\cdot)$  and transition kernel  $q(\cdot|\cdot)$  we define the detailed balance condition:

$$p(\theta^*)q(\theta^k|\theta^*) = p(\theta^k)q(\theta^*|\theta^k). \quad (4.21)$$

If a Markov chain is  $p$ -irreducible and (4.21) is satisfied, then  $p(\cdot)$  is the unique stationary distribution of the Markov chain and the chain is positive recurrent [42, 91].

## Aperiodicity

If states in the Markov chain are visited only at regular intervals, the chain is said to be periodic. Otherwise, the chain is aperiodic [31, 91].

## Ergodicity

A Markov chain is ergodic if it is aperiodic and positive Harris recurrent [31, 42, 91].

### 4.2.3 Convergence Theorems

We can now state the following theorem from [31, 42, 87, 91]:

**Theorem 4** *If a Markov chain with stationary distribution  $p$  is ergodic, then  $p$  is the unique stationary distribution and*

$$\lim_{k \rightarrow \infty} Pr(\theta^k \in A) = \int_A p(\theta) d\theta. \quad (4.22)$$

The next two theorems involve a sample from an ergodic Markov chain.

**Assumption 5** *Let  $\theta^0, \theta^1, \dots, \theta^m$  be a sample from an ergodic Markov chain with stationary distribution  $p$ ,  $h(\theta)$  be a real-valued function such that*

$$\int h(\theta) p(\theta) d\theta < \infty \quad (4.23)$$

*and  $\bar{h}_m = \frac{1}{m+1} \sum_{j=0}^m h(\theta^j)$ .*

The Ergodic Theorem is similar Law of Large Numbers, Theorem 2, except the sample is no longer random.

**Theorem 6 Ergodic Theorem.** *Let Assumption 5 hold. Then for every  $\epsilon > 0$ , [30, 31, 42, 91],*

$$Pr\left(\lim_{m \rightarrow \infty} \left| \bar{h}_m - \int h(\theta) p(\theta) d\theta \right| < \epsilon\right) = 1. \quad (4.24)$$

The implication of the Ergodic Theorem is that we can use a sample from an ergodic Markov chain to approximate expectations under the stationary distribution. We also have a version of the Central Limit Theorem for Markov Chains. First, we define the notion of convergence in distribution.

**Definition 4.1** *A sequence of random variables  $\theta^0, \theta^1, \dots$ , with cumulative distribution functions  $F_{\theta^i}$ , respectively, converges in distribution to a random variable  $\theta^*$  with cumulative distribution function  $F_{\theta^*}(\theta)$  if*

$$\lim_{m \rightarrow \infty} F_{\theta^m}(\theta) = F_{\theta^*}(\theta) \quad (4.25)$$

*at all points  $\theta$  where  $F_{\theta^*}(\theta)$  is continuous. [30]*

**Theorem 7 Central Limit Theorem for Markov Chains.** *Let Assumption 5 hold. Then there exists a real number  $\sigma^2(h)$  such that*

$$\sqrt{m} \left( \bar{h}_m - \int h(\theta) p(\theta) d\theta \right) \rightarrow N(0, \sigma^2(h)), \quad (4.26)$$

*in distribution, for any initial distribution [88, 91].*

### 4.3 Markov Chain Monte Carlo

By forming a Markov chain whose stationary distribution is the posterior density  $p(\theta|Y)$  in (3.57), we can obtain a sample from  $p(\theta|Y)$ . We can then use the theory of Monte Carlo integration to approximate posterior summaries using our sample. This is the idea of MCMC methods.

We will discuss three MCMC methods: Gibbs sampling, the Metropolis-Hastings algorithm, and the Delayed Rejection Adaptive Metropolis algorithm.



### 4.3.1 Gibbs Sampling

Gibbs sampling relies on the form of the full conditional (posterior) densities of the parameters  $\theta_i$  in our parameter vector  $\theta = (\theta_1, \dots, \theta_p)$ . That is, we need to know the form of the posterior densities,  $p(\theta_i|Y, \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p)$  of the  $\theta_i$ , which are generally not the same as the marginal posterior densities,  $p(\theta_i|Y)$ . This means we need conjugate prior densities (see Section 3.2.2 for details).

The Gibbs scheme also works if we update the parameters as vectors, as long as we knew their joint full conditional densities. This idea, called blocking, is especially useful when parameters are highly correlated.

Note that in subsection 3.2.2 we give an example of deriving the full conditional posterior densities when conjugate priors are available. For the general problem with arbitrary  $\theta$  and  $Y$ , we follow the notation in [31] and assume the densities depend on the data,  $p(\theta|Y) = p(\theta)$ .

Say we have the parameters  $\theta = (\theta_1, \dots, \theta_p)$  and we know their full conditional densities:

$$\begin{aligned} & p_{1|2,\dots,p}(\theta_1|\theta_2, \dots, \theta_p) \\ & p_{2|1,3,\dots,p}(\theta_2|\theta_1, \theta_3, \dots, \theta_p) \\ & \vdots \\ & p_{p-1|1,2,\dots,p-2,p}(\theta_{p-1}|\theta_1, \dots, \theta_{p-2}, \theta_p) \\ & p_{p|1,2,\dots,p-1}(\theta_p|\theta_1, \dots, \theta_{p-1}). \end{aligned}$$

To begin the Gibbs sampler, we choose initial parameter values,  $\theta^0 = (\theta_1^0, \dots, \theta_p^0)$ , and update these values iteratively by sampling from their respective full conditional densities:

$$\begin{aligned}
\theta_1^1 &\sim p_{1|2,\dots,p}(\theta_1|\theta_2^0, \dots, \theta_p^0) \\
\theta_2^1 &\sim p_{2|1,3,\dots,p}(\theta_2|\theta_1^1, \theta_3^0, \dots, \theta_p^0) \\
\theta_3^1 &\sim p_{3|1,2,4,\dots,p}(\theta_3|\theta_1^1, \theta_2^1, \theta_4^0, \dots, \theta_p^0) \\
&\vdots \\
\theta_{p-1}^1 &\sim p_{p-1|1,2,\dots,p-2,p}(\theta_{p-1}|\theta_1^1, \dots, \theta_{p-2}^1, \theta_p^0) \\
\theta_p^1 &\sim p_{p|1,2,\dots,p-1}(\theta_p|\theta_1^1, \dots, \theta_{p-1}^1).
\end{aligned}$$

Note how each parameter is updated in all subsequent densities. The stationary transition probabilities, independent of  $k$ , are defined as:

$$\begin{aligned}
\theta_1^k &\sim p_{1|2,\dots,p}(\theta_1|\theta_2^{k-1}, \dots, \theta_p^{k-1}) \\
\theta_2^k &\sim p_{2|1,3,\dots,p}(\theta_2|\theta_1^k, \theta_3^{k-1}, \dots, \theta_p^{k-1}) \\
\theta_3^k &\sim p_{3|1,2,4,\dots,p}(\theta_3|\theta_1^k, \theta_2^k, \theta_4^{k-1}, \dots, \theta_p^{k-1}) \\
&\vdots \\
\theta_{p-1}^k &\sim p_{p-1|1,2,\dots,p-2,p}(\theta_{p-1}|\theta_1^k, \dots, \theta_{p-2}^k, \theta_p^{k-1}) \\
\theta_p^k &\sim p_{p|1,2,\dots,p-1}(\theta_p|\theta_1^k, \dots, \theta_{p-1}^k).
\end{aligned}$$

Thus, adopting the notation from [31], we obtain the stationary transition density [31]

$$q(\theta^k|\theta^{k-1}) = q(\theta_1^k, \theta_2^k, \dots, \theta_p^k|\theta_1^{k-1}, \theta_2^{k-1}, \dots, \theta_p^{k-1}) \quad (4.27)$$

$$= p_{1|2,\dots,p}(\theta_1^k|\theta_2^{k-1}, \dots, \theta_p^{k-1}) p_{2|1,3,\dots,p}(\theta_2^k|\theta_1^k, \theta_3^{k-1}, \dots, \theta_p^{k-1}) \quad (4.28)$$

$$\cdots p_{p|1,2,\dots,p-1}(\theta_p^k|\theta_1^k, \dots, \theta_{p-1}^k). \quad (4.29)$$

The Gibbs algorithm is easy to implement but requires conjugate prior distributions, which

are often difficult or impossible to obtain when considering the densities of parameters in a nonlinear model. An additional disadvantage of the Gibbs algorithm is that we are unable to update all of the parameters at once, which can lead to a slower convergence due to correlated parameters.

#### Algorithm 8 (Gibbs Sampling)

1. Choose  $N_s$ , the number of samples to draw from the posterior density, and initial parameter values  $\theta^0 = (\theta_1^0, \dots, \theta_d^0)$ , where components  $\theta_i^0$  are blocks: either scalars or vectors [31, 42].
2. For  $k = 1 : N_s$ 
  - (a) For  $i = 1 : d$ 
    - Sample  $\theta_i^k$  from  $p_{i|1,2,\dots,i-1,i+1,\dots,d}(\theta_i|Y, \theta_1^k, \dots, \theta_{i-1}^k, \theta_{i+1}^{k-1}, \dots, \theta_d^{k-1})$ .

#### 4.3.2 Metropolis-Hastings

The Metropolis-Hastings algorithm is useful for sampling the posterior density when conjugate priors are unavailable. One of the benefits of the Metropolis-Hastings algorithm is that we can update the entire parameter vector in one iteration. We could also use it to evaluate unrecognizable full conditional distributions within the Gibbs Sampler. This idea is called Metropolis-within-Gibbs [31].

Metropolis-Hastings is an accept-reject algorithm, meaning candidate parameter values are selected from a proposal density and accepted or rejected according to their value under the posterior density relative to that of the current parameter value. The Metropolis algorithm was developed by Metropolis [71] and then generalized by Hastings [49].

We begin by considering the parameters  $\theta = (\theta_1, \dots, \theta_p)$  and their posterior density  $p(\theta)$ , which in our notation we again assume as in [31] depends on the data,  $p(\theta|Y) = p(\theta)$ . We also consider the proposal density  $h(\theta^*|\theta^k)$ , which generates a sample  $\theta^*$  given the current state of the Markov chain,  $\theta^k$ . We call  $\theta^*$  the candidate parameter value.

The normal distribution is a common choice for the proposal density  $h(\theta^*|\theta^k)$ . An example of this choice is seen in [87]. In this case, the mean of the proposal is often the current parameter value,  $\theta^k$ , and the variance of the density may be tuned so the proposal density better reflects the posterior density.

We define

$$\alpha(\theta^*, \theta^k) = \min \left\{ 1, \frac{p(\theta^*)h(\theta^k|\theta^*)}{p(\theta^k)h(\theta^*|\theta^k)} \right\} \quad (4.30)$$

as the acceptance probability [31]. We then sample  $U \sim U[0, 1]$  and accept  $\theta^{k+1} = \theta^*$  if  $U \leq \alpha$ . If  $U > \alpha$ , we keep the current parameter value and set  $\theta^{k+1} = \theta^k$  [31].

This is the idea of the Metropolis-Hastings algorithm. We note that in (4.30) any constants in the posterior density will cancel in the fraction, so we are only required to know the kernel of the posterior density [31]. Also, if the proposal density is symmetric,  $h(\theta^k|\theta^*) = h(\theta^*|\theta^k)$ , then we have

$$\alpha_M(\theta^*, \theta^k) = \min \left\{ 1, \frac{p(\theta^*)}{p(\theta^k)} \right\}, \quad (4.31)$$

the acceptance probability for the Metropolis algorithm [31]. We reason as in [31] that in this case, if  $p(\theta^*) > p(\theta^k)$ , we accept  $\theta^*$ . Otherwise, we accept  $\theta^*$  with probability  $\alpha_M$ .

It is important that the proposal density  $h(\theta^*|\theta^k)$  generates values throughout the parameter space as well as parameter values with a high probability of acceptance [42]. Ideally, the proposal density will be similar to the posterior density, but this is not a requirement [31, 47]. We may tune the proposal density as the simulation progresses in order to improve the acceptance rate. An acceptance rate between 20% and 50% is desirable [42]. To accomplish this tuning, we may use a normal proposal density and adjust the variance or covariance throughout the simulation based on the acceptance rate up to that point [31].

#### **Algorithm 9 (Metropolis-Hastings)**

1. Choose  $N_s$  the number of samples to draw from the posterior density, and initial parameter values  $\theta^0$  such that  $p(\theta^0) > 0$  [87].

2. For  $k = 1 : N_s$

(a) Sample  $\theta^*$  from proposal density  $h(\theta^*|\theta^{k-1})$ .

(b) Compute

$$\alpha(\theta^*, \theta^{k-1}) = \min \left\{ 1, \frac{p(\theta^*)h(\theta^{k-1}|\theta^*)}{p(\theta^{k-1})h(\theta^*|\theta^{k-1})} \right\}. \quad (4.32)$$

(c) Sample  $U$  from  $U[0, 1]$ .

(d) If  $U \leq \alpha$  set  $\theta^k = \theta^*$ , Else set  $\theta^k = \theta^{k-1}$ .

### 4.3.3 Delayed Rejection Adaptive Metropolis

The Delayed Rejection Adaptive Metropolis (DRAM) algorithm [47] builds upon the Metropolis-Hastings algorithm by using multiple proposal densities and tuning the proposal covariance with a recursive formula. Matlab code written by Marko Laine for this algorithm is available here: <http://helios.fmi.fi/~lainema/mcmc>.

#### Delayed Rejection

The delayed rejection (DR) aspect of DRAM is implemented when a candidate value  $\theta^{1*}$  at step  $k$  is rejected. Instead of retaining the current value  $\theta^k$  as the new value  $\theta^{k+1}$ , a new candidate value  $\theta^{2*}$  is selected from a second stage proposal density [47]. This delayed rejection scheme may be repeated for an arbitrary number of stages,  $i$ , ultimately drawing a candidate value from the  $i^{th}$ -stage proposal density and deciding to accept  $\theta^{k+1} = \theta^{i*}$  or retain our current value,  $\theta^{k+1} = \theta^k$ .

We explain the scheme in terms of the acceptance probabilities,  $\alpha$ . First, consider the acceptance probability  $\alpha_1$  of accepting  $\theta^{1*}$ , simulated from proposal density  $h_1(\theta^{1*}|\theta^k)$ , given the current state of the chain  $\theta^k$ :

$$\alpha_1(\theta^{1*}, \theta^k) = \min \left\{ 1, \frac{p(\theta^{1*})h_1(\theta^k|\theta^{1*})}{p(\theta^k)h_1(\theta^{1*}|\theta^k)} \right\} = \min \left\{ 1, \frac{N_1}{D_1} \right\}. \quad (4.33)$$

If  $\theta^{1*}$  is rejected, we propose another value  $\theta^{2*}$  from the second stage proposal density  $h_2(\theta^{2*}|\theta^k, \theta^{1*})$  and accept it with probability [47]:

$$\alpha_2(\theta^{2*}, \theta^{1*}, \theta^k) = \min \left\{ 1, \frac{p(\theta^{2*})h_1(\theta^{1*}|\theta^{2*})h_2(\theta^k|\theta^{2*}, \theta^{1*})[1 - \alpha_1(\theta^{1*}, \theta^{2*})]}{p(\theta^k)h_1(\theta^{1*}|\theta^k)h_2(\theta^{2*}|\theta^k, \theta^{1*})[1 - \alpha_1(\theta^{1*}, \theta^k)]} \right\} \quad (4.34)$$

$$= \min \left\{ 1, \frac{N_2}{D_2} \right\}. \quad (4.35)$$

We can consider any number of proposal densities, making this an iterative process [47]. The acceptance probability at stage  $i$  is [47] [72]:

$$\begin{aligned} \alpha_i(\theta^{i*}, \dots, \theta^{1*}, \theta^k) &= \\ & \min \left\{ 1, \frac{p(\theta^{i*})h_1(\theta^{i-1*}|\theta^{i*})h_2(\theta^{i-2*}|\theta^{i*}, \theta^{i-1*}) \dots h_i(\theta^k|\theta^{i*}, \theta^{i-1*}, \dots)}{p(\theta^k)h_1(\theta^{1*}|\theta^k)h_2(\theta^{2*}|\theta^k, \theta^{1*}) \dots h_i(\theta^{i*}|\theta^k, \theta^{1*}, \dots)} \right. \\ & \quad \left. \frac{[1 - \alpha_1(\theta^{i-1*}, \theta^{i*})][1 - \alpha_2(\theta^{i-2*}, \theta^{i-1*}, \theta^{i*})] \dots [1 - \alpha_{i-1}(\theta^{1*}, \dots, \theta^{i*})]}{[1 - \alpha_1(\theta^{1*}, \theta^k)][1 - \alpha_2(\theta^{2*}, \theta^{1*}, \theta^k)] \dots [1 - \alpha_{i-1}(\theta^{i-1*}, \dots, \theta^{1*}, \theta^k)]} \right\} \\ &= \min \left\{ 1, \frac{N_i}{D_i} \right\}, \end{aligned}$$

where  $q_i$  is the  $i^{th}$ -stage proposal density. Each time that a candidate value  $\theta^{i*}$  is rejected, it means that  $\alpha_i = \frac{N_i}{D_i}$  as well as that  $\alpha_1, \dots, \alpha_{i-1}$  are  $\frac{N_1}{D_1}, \dots, \frac{N_{i-1}}{D_{i-1}}$ , respectively. Thus we have the formulas [47]

$$D_i = h_i(\theta^{i*}|\theta^k, \dots, \theta^{i-1*})(D_{i-1} - N_{i-1}), \quad (4.36)$$

and

$$D_i = h_i(\theta^{i*}|\theta^k, \dots, \theta^{i-1*})[h_{i-1}(\theta^{i-1*}|\theta^k, \dots, \theta^{i-2*})[h_{i-2}(\theta^{i-2*}|\theta^k, \dots, \theta^{i-3*}) \dots \quad (4.37)$$

$$[h_3[h_2(\theta^{2*}|\theta^k, \theta^{1*})[p(\theta^k)h_1(\theta^{1*}|\theta^k) - N_1] - N_2] - N_3] - N_4] \dots - N_{i-1}]. \quad (4.38)$$

We explain the benefit of the delayed rejection scheme in terms of asymptotic variance. Consider an estimate computed using the sample from DRAM. As the sample size grows, the variance of this estimate decreases. The benefit of using the delayed rejection scheme is that the variance of the estimate decreases more quickly than the variance of an estimate computed with a Metropolis-Hastings generated sample. In this sense, the delayed rejection scheme is more efficient [47, 88].

### Adaptive Metropolis

The adaptive aspect of DRAM uses the previous sampled values from the Markov chain to compute the covariance matrix of a normal proposal density. After a certain non-adaptation period in which we use the initial covariance, we begin the adaptation phase [47]. We fix the length of the non-adaptation period as  $k_0$ , meaning we draw samples  $\theta^0, \dots, \theta^{k_0}$  using the initial proposal covariance  $C_0$ . For samples after this point, we update the proposal covariance matrix  $C_k$  using the sample up to the current point. For small  $\epsilon > 0$ , scaling parameter  $s_p$  which depends only on the dimension of the parameter space and  $p$ -dimensional identity matrix  $I_p$  [47],

$$C_k = \begin{cases} C_0, & k \leq k_0 \\ s_p \text{Cov}(\theta^0, \dots, \theta^{k-1}) + s_p \epsilon I_p, & k > k_0 \end{cases}. \quad (4.39)$$

Letting  $\bar{\theta}^k = \frac{1}{k+1} \sum_{i=0}^k \theta^i$ , the formula for the covariance is [47]:

$$\text{Cov}(\theta^0, \dots, \theta^k) = \frac{1}{k} \left( \sum_{i=0}^k \theta^i (\theta^i)^T - (k+1) \bar{\theta}^k (\bar{\theta}^k)^T \right). \quad (4.40)$$

The proposal covariance can be updated with a recursive formula [47]:

$$C_{k+1} = \frac{k-1}{k} C_k + \frac{s_p}{k} \left( k \bar{\theta}^{k-1} (\bar{\theta}^{k-1})^T - (k+1) \bar{\theta}^k (\bar{\theta}^k)^T + \theta^k (\theta^k)^T + \epsilon I_p \right). \quad (4.41)$$

The adaptive scheme makes the overall simulation more efficient [47] by increasing the ability of the proposal density to reflect the variances of and covariances between the parameters.

We would like for our proposal density to be similar to the posterior density. The delayed rejection strategy of DRAM uses multiple proposal densities, some of which may be similar to the posterior. Meanwhile, the adaptive strategy tunes the proposal densities based on the previous samples. Together, these two strategies lead the proposal to become more like the posterior so that we obtain samples from the posterior more efficiently.

As described in [47], one choice for the first stage proposal density is the normal distribution. Thus, the covariance is tuned following the adaptive Metropolis theory. The higher stage proposal densities are of the same form, but the covariance of each density is a constant multiple of the stage one covariance [47].

Below is the DRAM algorithm. We follow [87] and show only the case for a second stage proposal.

**Algorithm 10 (DRAM)**

1. Choose  $N_s$  the number of samples to draw from the posterior density, initial covariance  $C_0$ , initial parameter values  $\theta^0$  and the length of the adaptation interval,  $k_0$ .
2. For  $k = 1 : N_s$ 
  - (a) Sample  $\theta^{1*}$  from proposal density  $h_1(\theta^{1*}|\theta^{k-1}) \sim N(\theta^{k-1}, C_k)$ .
  - (b) Compute
$$\alpha_1(\theta^{1*}, \theta^{k-1}) = \min \left\{ 1, \frac{p(\theta^{1*})h_1(\theta^{k-1}|\theta^{1*})}{p(\theta^{k-1})h_1(\theta^{1*}|\theta^{k-1})} \right\}$$
  - (c) Sample  $U$  from  $U[0, 1]$ .
  - (d) If  $U \leq \alpha_1$  set  $\theta^k = \theta^{1*}$ .
  - (e) Else sample  $\theta^{2*}$  from  $h_2(\theta^{2*}|\theta^{k-1}, \theta^{1*})$ 
    - Compute  $\alpha_2(\theta^{2*}, \theta^{1*}, \theta^{k-1})$  (See Equation (4.34)).



- *Sample  $U$  from  $U[0, 1]$ .*
  - *If  $U \leq \alpha_2$  set  $\theta^k = \theta^{2*}$*
  - *Else set  $\theta^k = \theta^{k-1}$*
3. *If  $k \geq k_0$ , update  $C_k$  by (4.41).*

We note that this algorithm could be adapted slightly to allow for non-adaptation periods throughout the simulation, instead of just at the beginning [47].

### 4.3.4 Convergence Diagnostics

The theory for MCMC methods establishes that the distribution from which we sample will converge to the posterior distribution, eventually. At the beginning of our simulation, our samples may be from a distribution far from the desired posterior distribution. Thus, we discard the first portion of samples before we perform our analysis. The first part of the simulation which we discard is called the burn-in period [31]. The problem is deciding the length of the burn-in period. There are many convergence diagnostics to determine if we have performed a sufficient number of iterations in order to consider our sample an approximate random sample from the posterior density [31, 42]. Here we consider diagnostics based on plots of the chain histories, the autocorrelation function and the Monte Carlo error.

#### Chain Histories

A simple way to assess the convergence of the Markov chain is to plot the history of the chain. The plot of a chain that has settled down [31] to the stationary distribution has noticeable characteristics. For one, it will not contain long flat periods in which several consecutive candidate values have been rejected. The chain will no longer follow trajectories exploring the parameter space but will be focused on an interval. It will appear to be rapidly and randomly oscillating between values with a consistent amplitude.

To guard against sampling from local maxima, a test is to run multiple chains with different initial values. The plots of the chains should all converge to the same distribution [31].

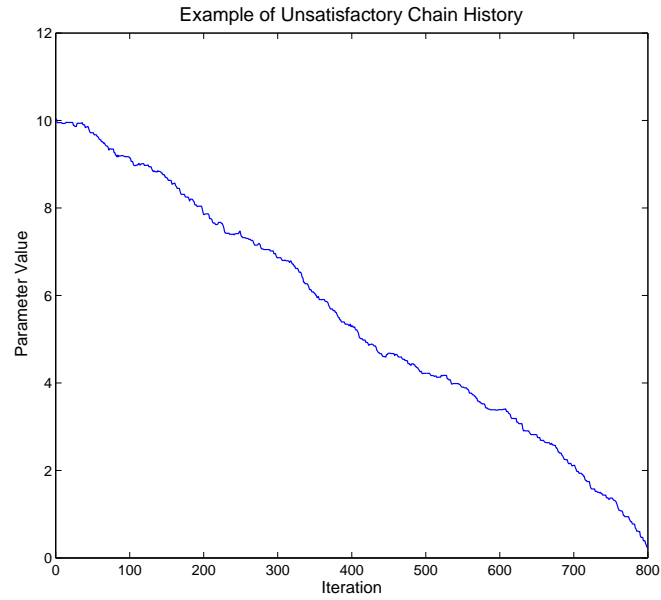


Figure 4.1: Example of Markov chain following a trajectory through the parameter space.

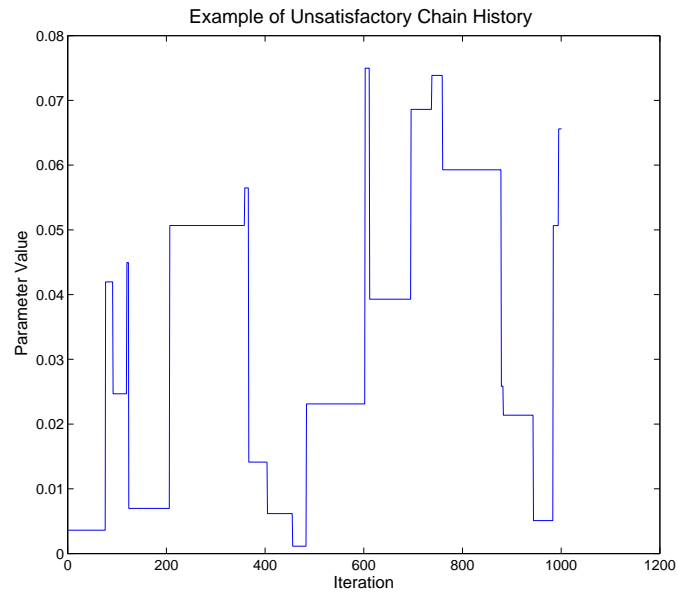


Figure 4.2: Example of Markov chain rejecting too many values.

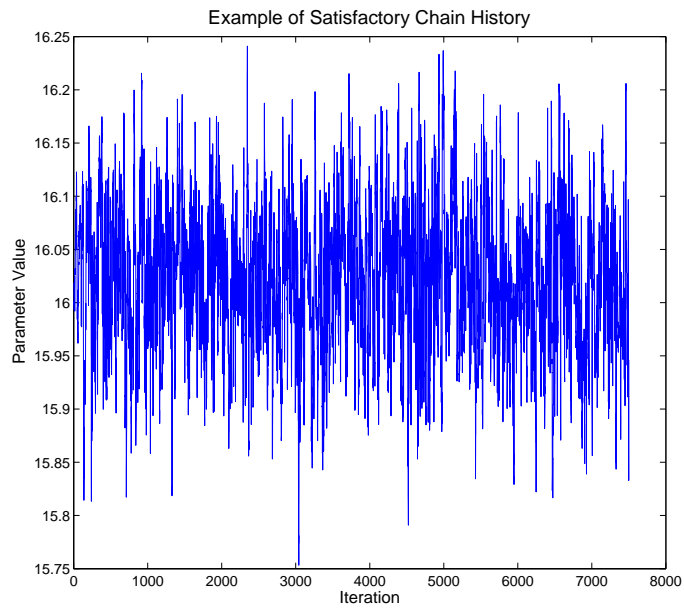


Figure 4.3: Example of Markov chain convergence to stationary distribution.

## Autocorrelation Function

Because we use a Markov chain to draw samples from the posterior distribution, consecutive samples are not independent. It is important to sample enough values to accommodate for this inherent correlation [31]. We use the autocorrelation function to estimate the correlation between draws that are  $m$  integers apart.

If we let  $\overline{\theta^k} = \frac{1}{k+1} \sum_{i=0}^k \theta^i$ , the empirical lag  $m \geq 0$  autocovariance for scalar  $\theta$  is [40, 42, 44]:

$$\gamma(m) = \frac{1}{k+1} \sum_{i=0}^{k-m} (\theta^i - \overline{\theta^k})(\theta^{i+m} - \overline{\theta^k}). \quad (4.42)$$

We use the sample variance  $\gamma(0) = \frac{1}{k} \sum_{i=0}^k (\theta^i - \overline{\theta^k})^2$  [30] to compute the lag  $m$  autocorrelation [16, 42]:

$$\rho(m) = \frac{\gamma(m)}{\gamma(0)} = \frac{\frac{1}{k+1} \sum_{i=0}^{k-m} (\theta^i - \overline{\theta^k})(\theta^{i+m} - \overline{\theta^k})}{\frac{1}{k} \sum_{i=0}^k (\theta^i - \overline{\theta^k})^2}. \quad (4.43)$$

The integrated autocorrelation time,  $\tau$ , is a scalar summary of the autocorrelation [42]:

$$\tau = 1 + 2 \sum_{m=1}^{\infty} \rho(m). \quad (4.44)$$

The autocorrelation function should decrease as  $m$  increases. A satisfactory plot of  $\rho(m)$  is in Figure 4.4, while an unsatisfactory plot is in Figure 4.5. If the autocorrelation does not decrease rapidly, we must draw more samples or use thinning to create a random sample [31].

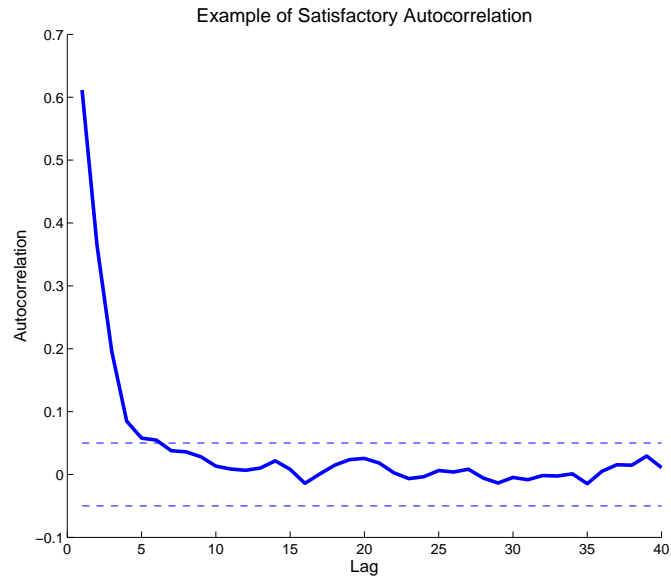


Figure 4.4: Example of autocorrelation plot with satisfactory rate of decrease.

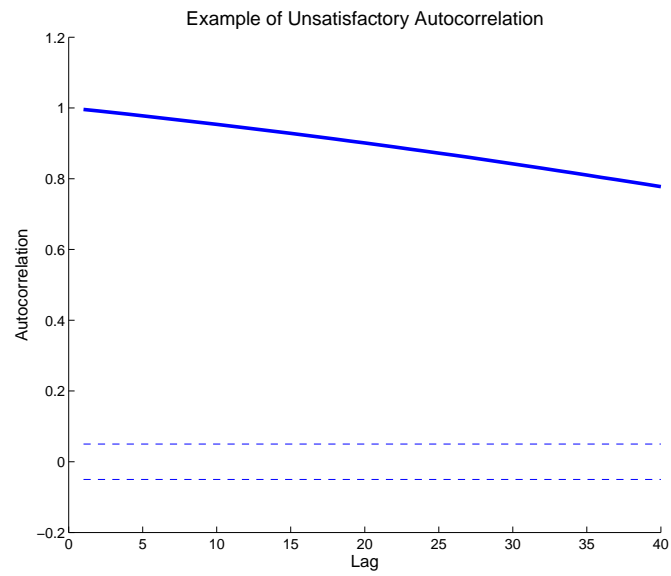


Figure 4.5: Example of autocorrelation plot with high autocorrelation (unsatisfactory).

## Monte Carlo Standard Error

From the Central Limit Theorem, Theorem 3, and the Ergodic Theorem, Theorem 6, we know that

$$\sqrt{m}(\overline{\theta^m} - \mu) \rightarrow N(0, \sigma^2) \text{ as } m \rightarrow \infty. \quad (4.45)$$

The Monte Carlo Standard Error,  $\frac{\hat{\sigma}}{\sqrt{m}}$ , relates the sample size  $m$  to the Monte Carlo Error  $\overline{\theta^m} - \mu$  [38, 40]. We compute the estimate  $\hat{\sigma}$  using batch means [40, 65]. That is [40], for  $n = a_n b_n$  and

$$\overline{\theta^k} = \frac{1}{b_n} \sum_{i=1}^{b_n} \theta^{k b_n + i}, \quad (4.46)$$

$$\hat{\sigma}^2 = \frac{b_n}{a_n - 1} \sum_{k=0}^{a_n-1} (\overline{\theta^k} - \overline{\theta^m})^2. \quad (4.47)$$

The number of batches,  $a_n$ , and the number of samples in each batch,  $b_n$ , may depend on  $n$  [38]. The default values in R are  $\sqrt{n} = a_n = b_n$  and in DRAM are  $b_n = \max(10, \text{fix}(n/20))$  [39, 40, 65, 82].

If the Monte Carlo Standard Error is large, we should continue to sample from the posterior [31]. Consider the interval  $\overline{\theta^m} \pm 2 \frac{\hat{\sigma}}{\sqrt{m}}$ . We use an example from [31] to illustrate satisfactory and unsatisfactory Monte Carlo Standard Errors: A satisfactory interval is

$$\overline{\theta^m} \pm 2 \frac{\hat{\sigma}}{\sqrt{m}} = 10 \pm 0.00001, \quad (4.48)$$

while unsatisfactory intervals are  $10 \pm 0.5$  and  $0.37 \pm 0.1$ .

### 4.3.5 Prediction Intervals

One of the advantages of using Bayesian inference is the theoretical framework for propagating uncertainty in parameters through the model. In this section we describe the use of prediction intervals for future observations.

Prediction intervals are a type of interval estimate, as are confidence and credible intervals.

However, each of these types of interval estimates is interpreted differently. Confidence intervals are common in the frequentist approach to statistical inference and quantify variability in repeating the experiment. A 95% confidence interval means that if an experiment is repeated many times, 95% of the confidence intervals will contain the true value of the quantity of interest. In this sense, the quantity of interest is fixed and unknown, but not random.

We find credible, or probability, intervals using the posterior density of parameters or predictions, whichever is our quantity of interest. A 95% credible interval means that we believe the quantity of interest is in the interval with 95% probability. Here, we consider the quantity of interest a random variable, making this a product of Bayesian inference.

When we discuss credible intervals in the context of predictions, again a Bayesian idea, we consider the expected value of future observations  $E(y_{new})$ , not the observations themselves,  $y_i$  [87]. Prediction intervals, however, quantify the value and variability of future observations  $y_i$  and incorporate measurement error. A 95% prediction interval means we believe that future observations are within the interval with a probability of 95%.

In practice, we compute prediction intervals by sampling  $\theta_1, \dots, \theta_m$  from the posterior density  $p(\theta|Y)$  and computing model values at each spatiotemporal point  $u(\theta_i) = u_i$  at the sample values. That is, for each spatiotemporal point we have  $m$  function values. We also sample the error variance  $\sigma^2$  at the same points in the chain so that we have  $m$  values for the error variance. We perturb the  $u_i$  by adding  $p_i\sigma_i$ , where  $p_i \sim N(0, 1)$ , to compute [65]

$$\hat{u} = u + p_i\sigma_i = u_i + \hat{\epsilon}_i.$$

Thus we have a sample of  $m$  values:  $\hat{u} = u + \hat{\epsilon}$  for each spatiotemporal point [88]. We compute the sample percentile values to approximate the distribution percentile values [31, 47, 65], or prediction intervals

## Chapter 5

# Results: Calibration of the Apparent Thermal Diffusivity

In this chapter we discuss the results of using ADH with PEST to determine a scalar estimate for  $k$  as well as the analytical model with DRAM to sample the posterior density of  $k$ . We first describe the data we use in the calibration process.

### 5.1 Data

We use data collected by the US Army Corps of Engineers to calibrate our model and parameters. The data is from three locations in North America representing three different climatic zones under the Köppen climate classification system [36], [35]. The Köppen system classifies the earth’s landmass into five regions depending on temperature and precipitation, Types A-E [77]. An additional high-altitude region is specified Type H [77]. We consider data from tropical, Type A; desert, Type B; and undifferentiated highlands, Type H; climactic zones. We refer to these datasets as Dataset A, Dataset B and Dataset H, respectively. A meteorological (met) station at each location was used to collect the data over 5-minute intervals for a period of 40 days. Since we are considering data with a constant water content, we only consider



data for consecutive days when there was no precipitation. Thus, each dataset has a different temporal length [36].

The met stations served to collect subsurface, surface and air temperatures, shortwave and longwave radiation measurements, wind speed and direction, soil moisture content, relative humidity, barometric pressure, subsurface heat flux and precipitation. For our calibration efforts we only use data from the temperature probes buried below the surface but note that additional data is used in the ADH boundary condition. Thermocouples located on the probes at depths of 1, 5, 10, 15, 20, 25 and 30cm are accurate to  $\pm 0.1$  degrees Celsius [35]. Additional details regarding the data are in Appendix A.

In addition to the met station data, we have soil samples from each location in order to perform lab measurements. The lab analysis leads to the empirical estimates for  $k$  [36] of Lu and Johansen [57, 66]

Our measurements for the shortwave and longwave radiation give us insight into variations in the surface and subsurface temperatures. We collect the shortwave data using the Eppley Laboratory Precision Spectral Pyranometer (PSP) which corresponds to a wavelength range of 0.285-2.8  $\mu\text{m}$ , including ultraviolet, visible and near-infrared radiation [35] [13]. This shortwave data is the amount of shortwave solar radiation reaching the earth's surface in  $\text{W}/\text{m}^2$ . Radiation emitted by the earth and atmosphere is terrestrial radiation. This longwave or infrared radiation falls into a range of 4-100 $\mu\text{m}$  [13], the far infrared spectrum [56]. We use the Eppley Laboratory Precision Infrared Radiometer to measure longwave radiation in a wavelength range of 3.5-50 $\mu\text{m}$  [35].

We illustrate the effects of solar and terrestrial radiation on ground temperature in Figures 5.1 and 5.2, where we plot the shortwave and longwave radiation as well as the temperature at 1cm depth for Datasets A and H.

In Figure 5.1, we see spikes in the shortwave radiation up to nearly 1000  $\text{W}/\text{m}^2$  in the first two days and the fourth day of our data. However, on days 2 and 5 shortwave radiation peaks closer to 200  $\text{W}/\text{m}^2$ . Longwave radiation remains near 400  $\text{W}/\text{m}^2$  both day and night. Since

Dataset A is from a tropical rainforest, both cloud cover and vegetation prevent much of the shortwave radiation from reaching the ground surface. Small breaks in cloud cover allowing the shortwave radiation to reach the surface cause spikes in the ground temperature, as seen in the temperature plot in Figure 5.1.

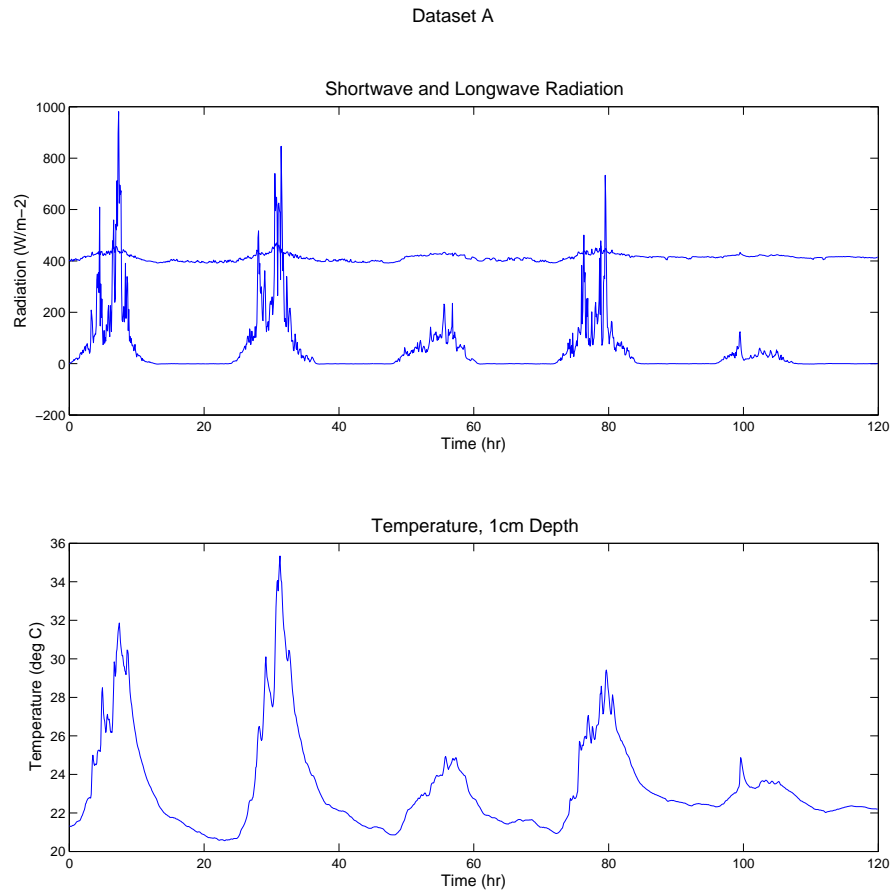


Figure 5.1: Shortwave radiation, longwave radiation and temperature at 1cm depth for Dataset A. The effects of spikes in shortwave radiation are visible in the plot of temperature at 1cm depth.

In Figure 5.2 we see daily peaks in shortwave radiation around 700-900 W/m-2, a contrast

to the large variations seen in Figure 5.1. The ground temperature again reflects the amount of shortwave radiation reaching the surface. The relative difference between temperature peaks is less for Dataset H than for Dataset A.

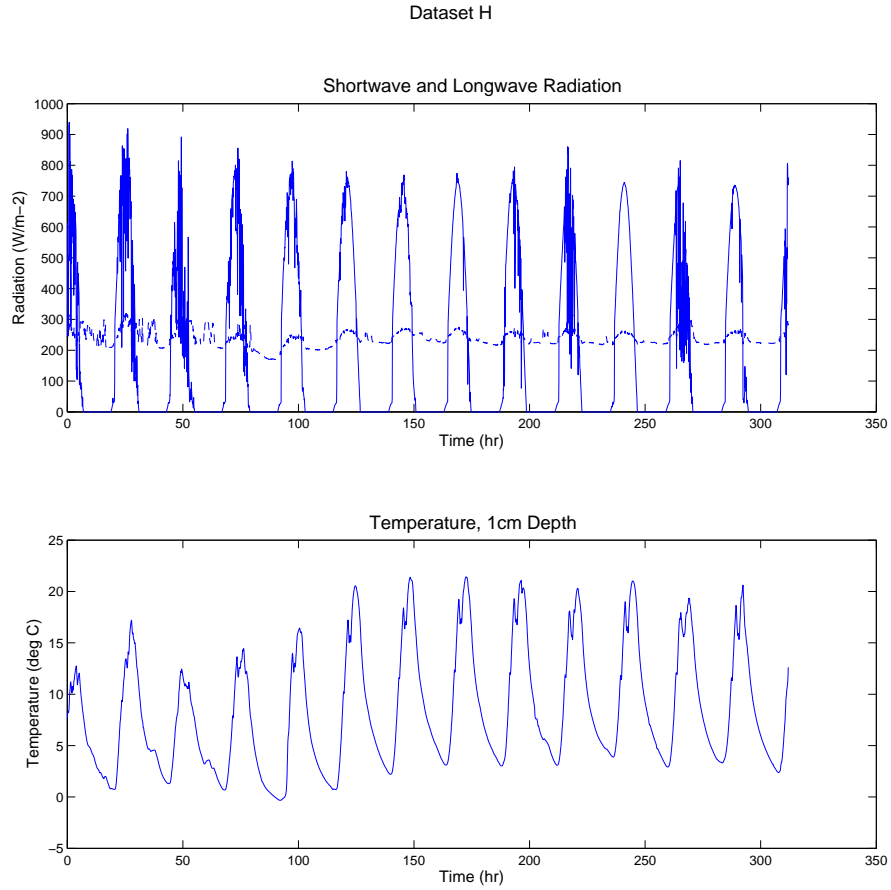


Figure 5.2: Shortwave radiation, longwave radiation and temperature at 1cm depth for Dataset H.

We will see the effects discussed here when we discuss our results.

For further details on the met station, see Appendix A.

## 5.2 Least Squares (PEST) with Numerical Model (ADH)

We optimize over soil parameters using PEST with ADH as the model. Specifically, we are interested in the six thermal soil material parameters in Table 5.1.

Table 5.1: Six thermal material parameters of interest.

Parameter	Symbol
Specific Heat of Solid	$c_{p,s}$
Dry Thermal Conductivity	$\kappa_d$
Saturated Thermal Conductivity	$\kappa_s$
Emissivity	$\epsilon_g$
Porosity	$\phi$
Albedo	$\alpha_g$

For Dataset H, we narrow down a subset of the parameters as detailed in [35] and Subsection 5.2.3. For Datasets A and B, we use subset selection to reduce the number of parameters. Doing so reduces the computation time for our calibration problem as well as the condition number of the Jacobian matrix. We select a subset of columns corresponding to a well-conditioned Jacobian matrix [55] using the strong rank revealing QR factorization algorithm [46] using the code implemented in [55]. These columns correspond to an identifiable subset of our parameters which we will keep for the calibration process [55, 81]. We set the rest of the parameters to nominal values [55, 81].

We use the constants in Table 5.2 in our simulation and to calculate the diffusivity  $k$  from the PEST results.

Table 5.2: Additional values used in ADH or to calculate  $k$ .

Parameter (ADH Card, Units)	Symbol	Value/Source
Thermal Conductivity of Water (TKW, W/m-K)	$\kappa_w$	(0.58) [25]
Specific Heat of Water (SHW, W-hr/g-K)	$c_{p,w}$	(0.001161472) [48]
Specific Gravity (Density) of Water (SGW, g/cm <sup>3</sup> )	$\rho_w$	(10 <sup>6</sup> ) [33]
Specific Heat of Gas (SHG, W-hr/g-K)	$c_{p,g}$	(0.000281111) [9]
Specific Gravity (Density) of Gas (SGG, g/cm <sup>3</sup> )	$\rho_g$	(1000) Value from [33] rounded to one significant figure
Thermal Conductivity of Gas (TKG, W/m-K)	$\kappa_g$	(0.025) [93]

The start and end times of the precipitation-free data from each location are detailed in Table 5.3, as well as the initial time specified for the ADH simulation. We compare the output of ADH to the data every 15 minutes and set the maximum time step at 15 minutes.

Table 5.3: Date of ADH simulation for each dataset.

<b>Dataset</b>	<b>Initial Time (Julian day, hour)</b>	<b>Start, End Time of Data (Julian day, hour)</b>
A	125 0:00	126 6:00 - 130 6:00
B	219 0:00	220 1:00 - 230 1:00
H	260 0:00	261 12:00 - 274 12:00

We must run the groundwater portion of ADH, even though we have a constant saturation, so we must specify an initial hydraulic head. A vertical column of water exerts a pressure directly related to its height [98]. The hydraulic head is described by the height of the column in meters, relative to some reference point. In our case, setting an initial hydraulic head of -25 meters means the water table is 25 meters below the bottom of our domain, which is 1m depth.

We determine the initial temperature by computing the average temperature at 30cm depth over the necessary ramp-up period. We either round the temperature to three decimal places or to the nearest whole number. We choose the initial total head by first determining the average saturation over the ramp-up period. To do this we divide the average water content by the initial porosity. Once we have the average saturation over the ramp-up period, we set the initial total head so that the initial saturation values ADH computes are as close as possible to the values we have determined. We use only increments of meters to fix the initial total head.

We perform our simulations on a 4x4x200 mesh. We use ADH svn revision number 8832M, PEST version 11.3. (Watermark Numerical Computing), a MacPro with two Quad-Core Intel Xeon processors and OS X 10.6.8.

### 5.2.1 Results: Dataset A

#### Initial Conditions

The optimization routine requires a set of subsurface data, the calibration data, to measure against the simulation (ADH) output. We also must inform ADH of the surface data, the simulation data, in order to implement the Neumann boundary condition. For Datasets B and H, our simulation data begins at least 24 hours before our calibration data, allowing constant initial conditions [35]. However, for Dataset A, our simulation data begins just 6 hours before our calibration data. We choose to use the output of a previous simulation as an initial condition since we do not have the necessary time interval (24 hours) to accommodate a constant initial condition.

We choose the initial condition in the following way, noting that the results here are for the post-subset selection initial case. The precipitation-free probe data begins at 6:00 on Julian day 125 while the boundary condition data for ADH begins at 0:00 on Julian day 125. We begin a simulation at 0:00 on day 125 with a constant initial condition and use the output after 48 hours (at 24:00 on day 126) as our new initial condition. We compare the following two simulations to evaluate this choice:

Table 5.4: Test cases for Dataset A initial conditions.

Case	Description
A	Begin simulation at 0:00 Julian day 125 with constant initial condition. Measure against probe data for hours 30-54 (allowing 30 hours for the initial condition to be incorporated).
B	Begin simulation at 0:00 Julian day 126 with initial condition from previous simulation. Measure against probe data for hours 6-30 (same comparison window as A).

Case A will be the standard of comparison as we compare ADH output to data after running ADH for 24 hours. Case B is the proposed initial condition for our simulations. We set the initial conditions for both hydraulic head and temperature, obtaining them

The weighted relative residual sum-of-squares for Case A is 114.5 and for Case B is 123.1. Plot 5.3 is Case A, Case B and the data at 6:00 on Julian day 125. Plot 5.4 is Case A, Case B and the data at 12:00 on Julian day 125. We find that the initial condition used in Case B is sufficient and proceed with the simulation.

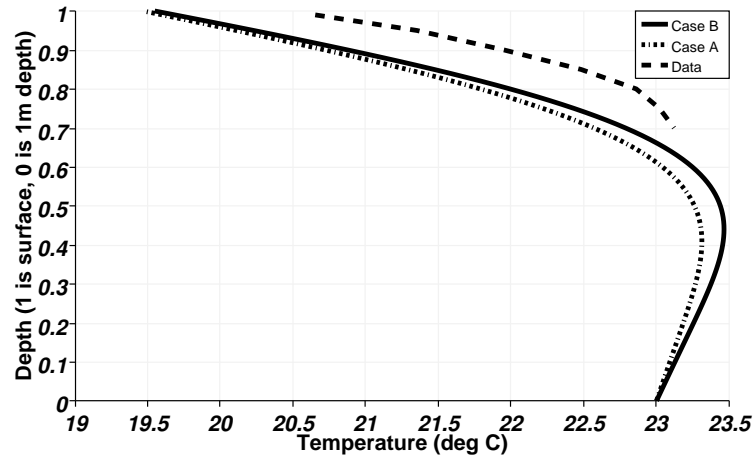


Figure 5.3: Temperature data and ADH output for Cases A and B, 6:00 on Julian day 125.



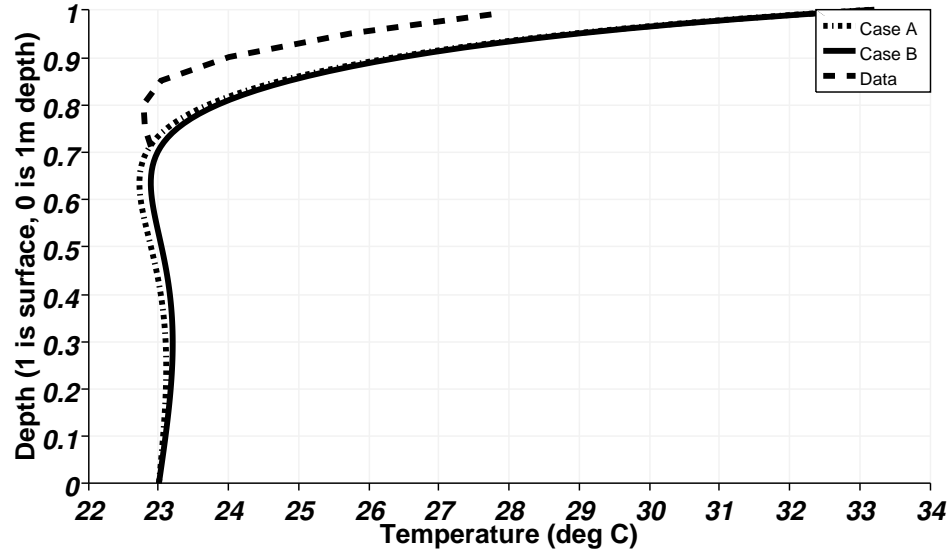


Figure 5.4: Temperature data and ADH output for Cases A and B, 12:00 on Julian day 125.

### Optimization Details

We weight the data for Dataset A in the following way:

Table 5.5: Weights for Dataset A.

Weight	Depth
0.0800	1 cm
0.1000	5 cm
0.1250	10 cm
0.2100	15 cm
0.3125	20 cm
0.5000	25 cm
0.7140	30 cm

We first compute the Jacobian using all six parameters in Table 5.1 using a ramped-up initial condition similar to the one described above. The singular values of the Jacobian matrix computed with the initial parameter values are in Table 5.6. These values indicate the Jacobian has a condition number on the order of  $10^7$ .

Table 5.6: Singular values of Jacobian computed with all initial parameters for Dataset A.

3.6493e+05
2.5037e+02
3.2584e+01
6.5330e+00
2.0259e+00
5.5485e-02

We perform subset selection and fix the dry thermal conductivity  $\kappa_d$  at 0.2, porosity  $\phi$  at 0.5 and saturated thermal conductivity  $\kappa_s$  at 1.4. Table 5.7 contains the fixed and initial values used in our simulations.

Table 5.7: Fixed values and initial values for PEST optimization, Dataset A constant saturation.

Parameter (ADH Card)	Symbol	Value	Source
Porosity (POR)	$\phi$	0.5	DBS&A or ERDC lab
Quartz Fraction (QTZ)	$q$	0.45	Expert
Percentage of Sand and Gravel (FSG)		0.875	ERDC lab
Specific Gravity (Density) of Solid (SGS)	$\rho_s$	2.6	Calculated using DBS&A data
Dry Thermal Conductivity (TKD)	$\kappa_d$	0.2	Depends on bulk soil density $\rho_b$ [57] or porosity $\phi$ [66]
Saturated Thermal Conductivity (TKS)	$\kappa_s$	1.4	Depends on $\phi$ , $q$
Specific Heat of Solid (SHS)	$c_{p,s}$	3.240 E-06	(Initial value)
Emissivity (EMS)	$\epsilon_g$	0.941152	(Initial value)
Albedo (ALB)	$\alpha_g$	0.399762	(Initial value)

The results of the PEST optimization for the remaining three parameters are in Table 5.8. The weighted residual sum-of-squares for the optimized values was 10.03.

Table 5.8: Results of PEST optimization for Dataset A.

Parameter	Symbol	Value
Specific Heat of Solid	$c_{p,s}$	1.99 E-04
Emissivity	$\epsilon_g$	0.380492
Albedo	$\alpha_g$	0.701515

ADH calculates a final saturation of 38.7% at a depth of 15cm. We use the values in Tables 5.2, 5.7, and 5.8 to calculate  $k$ : 2.071E-03.

We plot ADH temperature output using the optimized parameters against our data to evaluate the PEST results. The ADH simulations are run on a column corresponding to 2m depth. Figure 5.5 is the output at 1cm and 30cm depth over the calibration period.

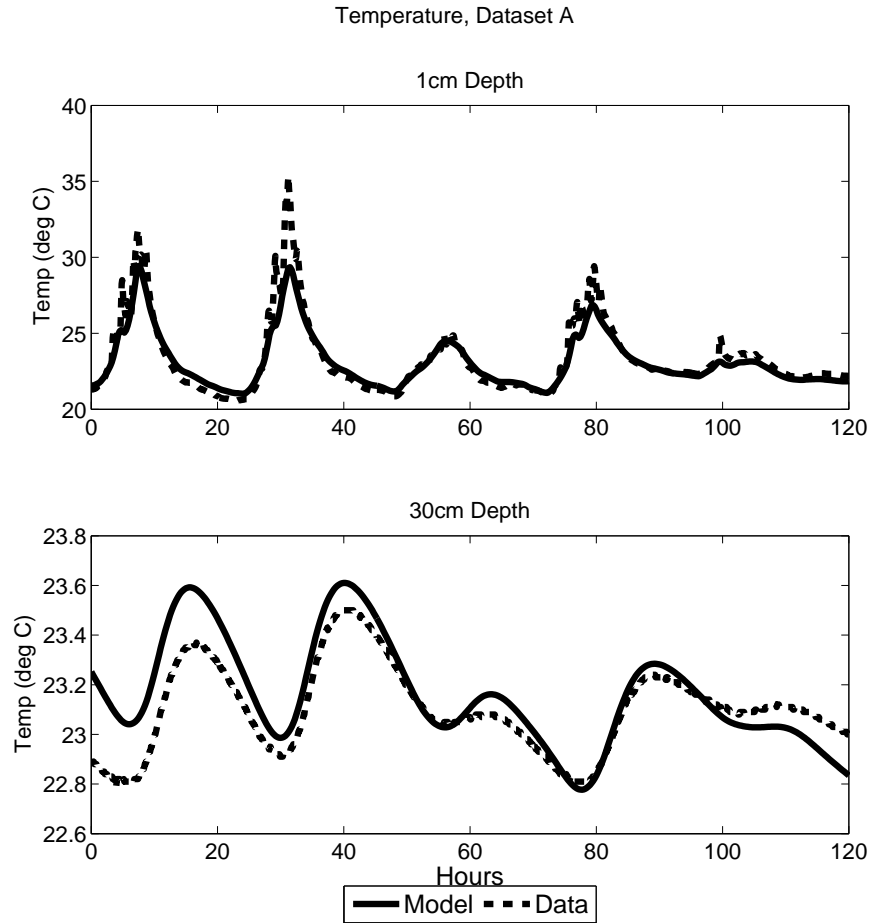


Figure 5.5: Dataset A model and temperature data at 1cm and 30cm depth over calibration period.

For a plot of the temperature output at all depths, see Appendix D.

### 5.2.2 Results: Dataset B

For Dataset B, we set a constant initial temperature of 32.837 degrees Celsius for all depths. This is the average temperature at 30cm depth using the available data from the ramp-up period. We set a constant initial hydraulic head of -8m, corresponding to an average saturation of about 14% at 15cm depth over the ramp-up period. Please note that we only have probe data beginning Julian day 219 at time 16:05 through Julian day 220 0:55 to use to determine these initial conditions.

#### Optimization Details

We weight the data for Dataset B in the same way as Dataset A. We again compute the Jacobian of the parameters in Table 5.1. The singular values of the Jacobian matrix computed with initial parameter values are in Table 5.9. These values indicate the Jacobian has a condition number on the order of  $10^8$ .

Table 5.9: Singular values of Jacobian computed with all initial parameters for Dataset B.

1.0229e+07
1.4038e+03
2.1444e+02
5.6277e+01
1.3927e+01
6.0764e-01

We perform subset selection and fix the dry thermal conductivity  $\kappa_d$  at 0.33, porosity  $\phi$  at 0.311 and saturated thermal conductivity  $\kappa_s$  at 1.7. The values used in ADH are listed in Table 5.10

Table 5.10: Fixed values and initial values for PEST optimization, Dataset B constant saturation.

Parameter (ADH Card)	Symbol	Value	Source
Porosity (POR)	$\phi$	0.311	Average from DBS&A over all samples and depths
Quartz Fraction (QTZ)	$q$	0.3	Expert
Percentage of Sand and Gravel (FSG)		0.95	DBS&A
Specific Gravity (Density) of Solid (SGS)	$\rho_s$	2.675	Average from ERDC or DBS&A
Dry Thermal Conductivity (TKD)	$\kappa_d$	0.33	Depends on bulk soil density $\rho_b$ [57] or porosity $\phi$ [66]
Saturated Thermal Conductivity (TKS)	$\kappa_s$	1.7	Depends on $\phi$ , $q$
Emissivity (EMS)	$\epsilon_g$	0.694223	(Initial value)
Albedo (ALB)	$\alpha_g$	0.306232	(Initial value)
Specific Heat of Solid (SHS)	$c_{p,s}$	3.240E-06	(Initial value)

The results PEST optimization for the remaining three parameters are in Table 5.11. The weighted residual sum-of-squares for the optimized values was 214.5.

Table 5.11: Results of PEST optimization for Dataset B.

Parameter	Symbol	Value
Specific Heat of Solid	$c_{p,s}$	4.52 E-04
Emissivity	$\epsilon_g$	0.952783
Albedo	$\alpha_g$	0.142241

## Simulation Details and Results

ADH calculates a final saturation of 13.1% at a depth of 15 cm. We use the values in Tables 5.2, 5.10, and 5.11 to calculate  $k$ : 9.277E-04.

Figure 5.6 contains the ADH output and data at 1cm and 30cm depth.

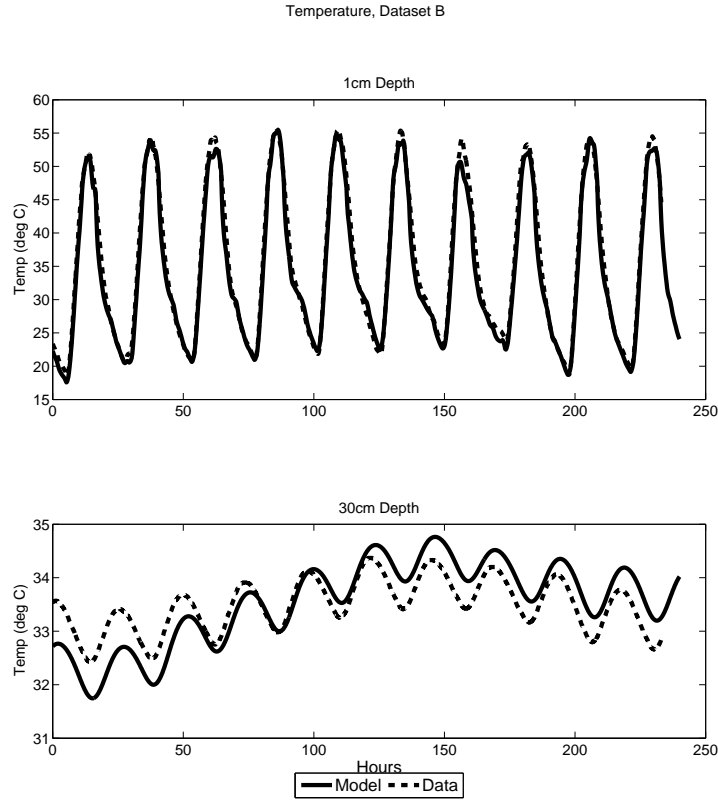


Figure 5.6: Dataset B model and temperature data at 1cm and 30cm depth over calibration period.

For a plot of the temperature output at all depths, see Appendix D.

### 5.2.3 Results: Dataset H

In [35], extensive care is taken to calibrate ADH to match Dataset H. We refer to [35] for many details justifying the fit but make a few points here. The domain in Dataset H is treated as two materials because simulations using one material did not fit the data for the entire domain, but only for portions [35]. We specify a 1cm insulating layer at the surface of the soil and estimate the diffusivity of the deeper soil below the gravel layer [35]. At the suggestion of our collaborators, we shift the ADH output up 2cm before comparing with the data.

For Dataset H, we set a constant initial temperature of 5.830 degrees Celsius for all depths and a constant initial hydraulic head of -6m, corresponding to a saturation of  $\approx 37\%$ .

### Optimization Details

We use the following weights for this dataset:

Table 5.12: Weights for Dataset H [35].

Weight	Depth
0.0750	1 cm
0.1000	5 cm
0.1250	10 cm
0.2100	15 cm
0.3125	20 cm
0.5000	25 cm
0.7140	30 cm

Instead of using subset selection for this dataset, we follow [35] in fixing and varying parameters. We begin with the six thermal parameters in Table 5.1, but we actually have twelve parameters because we divide the domain into two materials. We fix porosity at 37.7% in both



materials using DBS&A lab results. The dry and saturated thermal conductivities of the deeper soil are fixed at 0.28 and 1.9, respectively, at the suggestion of our collaborators, based on lab measurements. We follow [35] in setting the dry thermal conductivity equal to the saturated thermal conductivity of the gravel layer. In [35], a consistent value of 0.2 W/m-K is found for the dry/saturated thermal conductivity in the gravel layer, and we use this value in our simulations. We fix the albedo and emissivity of the deeper soil because these properties only relate to temperature changes at the surface. We use arbitrary vales of 0.625 and 0.575 for the deeper soil albedo and emissivity, respectively. Finally, we fix the albedo in the gravel layer at 0.35, a value determined using calibration with the radiometer data in [35].

The values used in ADH are listed in Table 5.13.

Table 5.13: Fixed values and initial values for PEST optimization, Dataset H constant saturation.

Parameter (ADH Card)	Symbol	Value	Source
Porosity (POR 1, POR 2)	$\phi$	0.377	DBS&A
Quartz Fraction (QTZ 1, QTZ 2)	$q$	0.45	Median value of range given by expert [35]
Percentage of Sand and Gravel (FSG 1, FSG 2)		0.925	Median value of range given by expert [35]
Specific Gravity (Density) of Soilds (SGS 1, SGS 2)	$\rho_s$	2.69	Calibrated using DBS&A data
Dry Thermal Conductivity, Deeper Soil (TKD 1)	$\kappa_d$	0.28	Depends on bulk soil density $\rho_b$ [57] or porosity $\phi$ [66]
Saturated Thermal Conductivity, Deeper Soil (TKS 1)	$\kappa_s$	1.9	Depends on $\phi, q$
Dry Thermal Conductivity, Gravel Layer (TKD 2)	$\kappa_d$	0.2	Depends on bulk soil density $\rho_b$ [57] or porosity $\phi$ [66]
Saturated Thermal Conductivity, Gravel Layer (TKS 2)	$\kappa_s$	0.2	Depends on $\phi, q$
Albedo, Deeper Soil (ALB 1)	$\alpha_g$	0.625	Arbitrary
Emissivity, Deeper Soil (EMS 1)	$\epsilon_g$	0.575	Arbitrary
Albedo, Gravel Layer (ALB 2)	$\alpha_g$	0.350	Determined using calibration with the radiometer data in [35]
Specific Heat of Solid, Deep Soil (SHS 1)	$c_{p,s}$	2.084 E-4	(Initial value)
Specific Heat of Solid, Gravel Layer (SHS 2)	$c_{p,s}$	9.861 E-5	(Initial value)
Emissivity, Gravel Layer (EMS 2)	$\epsilon_g$	0.92	(Initial value)

## Simulation Details and Results

The results of the PEST optimization for the three parameters are in Table 5.14. The final weighted residual sum-of-squares was 818.20.

Table 5.14: Results of PEST optimization for Dataset H.

Parameter	Symbol	Value
Specific Heat of Solid, Deep Soil	$c_{p,s}$	1.890000E-04
Specific Heat of Solid, Gravel Layer	$c_{p,s}$	5.332000E-04
Emissivity, Gravel Layer	$\epsilon_g$	0.808947

ADH calculates a final saturation of 37.1% at a depth of 15cm. We calculate  $k$ : 2.793E-03.

Figure 5.7 contains the ADH output and temperature data at 1cm and 30cm depth.

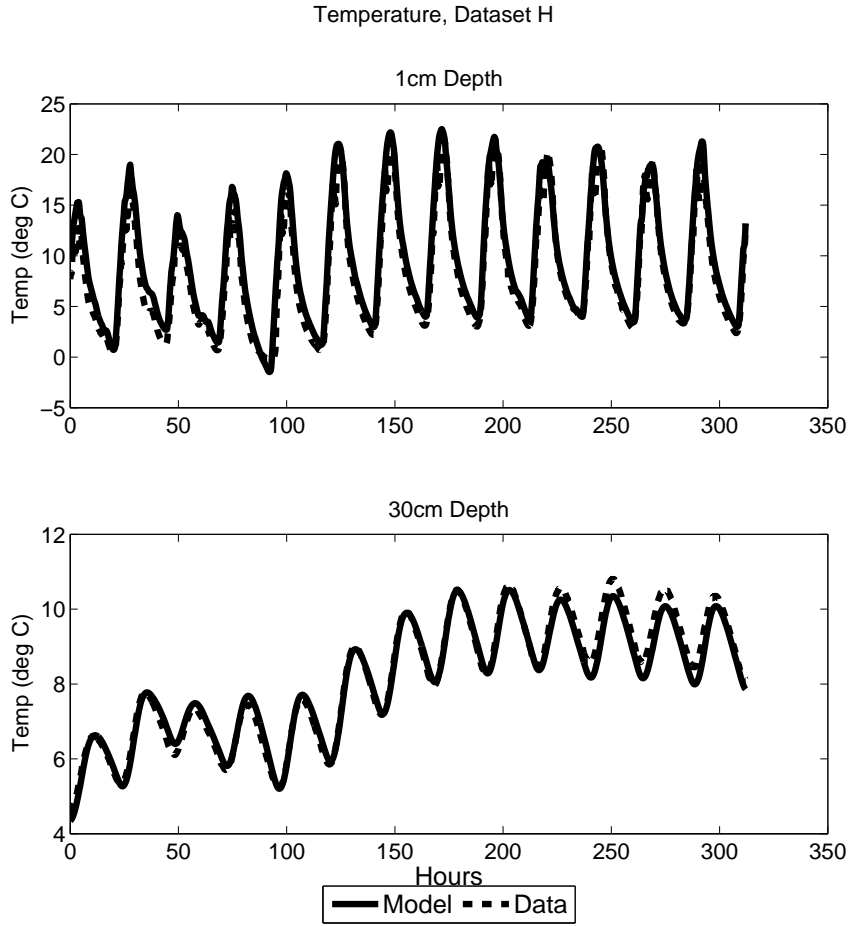


Figure 5.7: Dataset H model and temperature data at 1cm and 30cm depth over calibration period.

For a plot of the temperature output at all depths, see Appendix D.

### 5.3 Bayesian Inference (DRAM) with Analytic Model

For calibration using Bayesian inference, we use DRAM with the analytic model described in Chapter 2, Equation (2.8). We begin by presenting diagnostic results from our simulation and conclude that we achieve convergence. We then present results for the apparent thermal

diffusivity and prediction intervals for future observations.

### 5.3.1 Diagnostics

We use several diagnostics to determine if the Markov chains have converged to their stationary distributions. We choose to show the diagnostics for Dataset A since it is representative of all three datasets. For diagnostics for Datasets B and H, see Appendix D. Figures 5.8-5.10 contain the chain history for three different sets of samples. Figure 5.8 contains the first 3000 samples from the chain. The first 500 samples appear to have some direction, but after that the path of the chain is centered around a mean value.

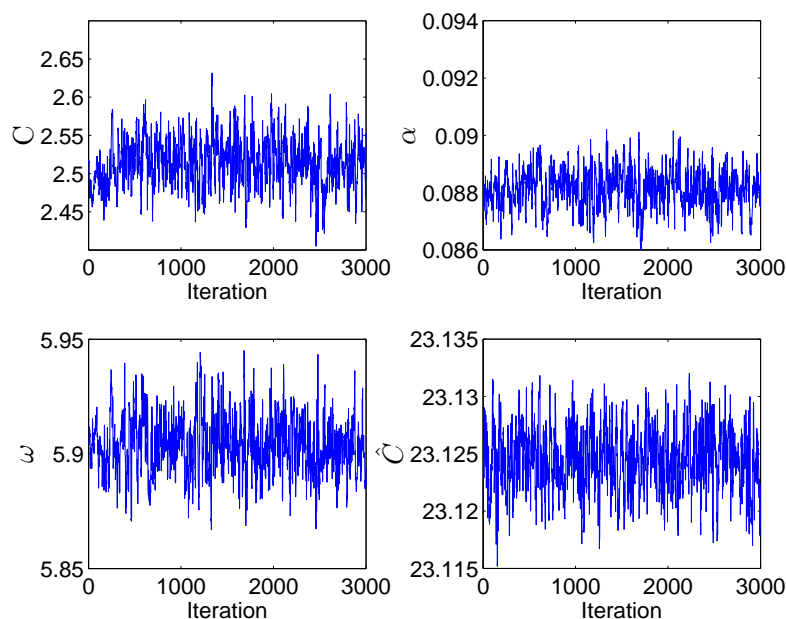


Figure 5.8: Chain history from 1 to 3,000 samples.

Figure 5.9 contains the first 10,000 samples from the Markov chain. The burn-in period is barely visible in these plots, indicating that the chains have converged in the first 1000-3000 iterations.

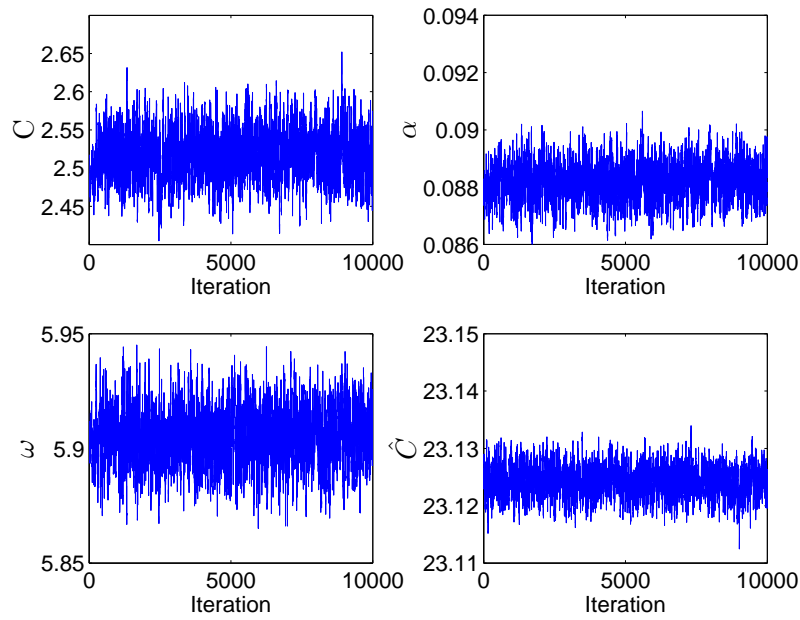


Figure 5.9: Chain history from 1 to 10,000 samples.

Figure 5.10 contains the history from 10,000 to 100,000 samples. It is clear that the chains have converged to their stationary distributions from these plots. For the purposes of this work, we consider the sample from iteration 3000 to iteration 10,000.

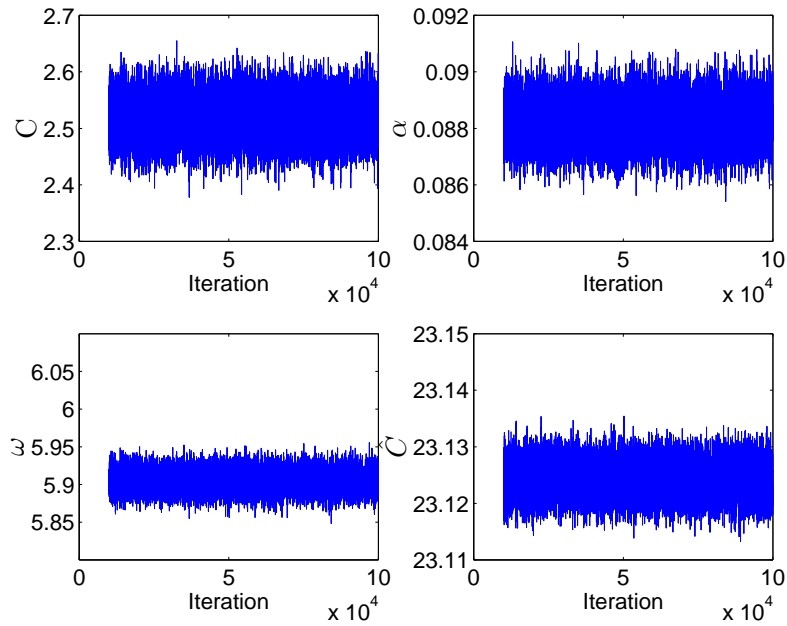


Figure 5.10: Chain history from 10,000 to 100,000 samples.

We plot the autocorrelation function for Dataset A in Figure 5.11 for the sample from iteration 3000 to iteration 10,000. This property indicates a sufficiently large sample size.

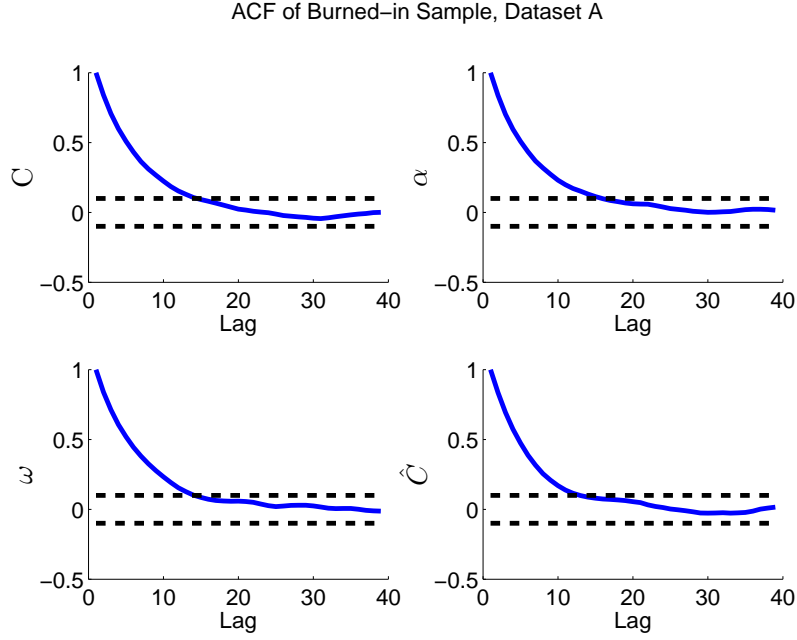


Figure 5.11: Autocorrelation for burned-in portion of chain.

In Table 5.15 we report the mean, standard deviation and Monte Carlo standard error (MCSE) for parameter  $\alpha$  for Dataset A. The MCSE is several orders of magnitude less than the mean of the sample, so each of these is a sufficiently large sample from the posterior density.

Table 5.15: Mean, standard deviation and Monte Carlo standard error (MCSE) for  $\alpha$  from Dataset A.

Sample	Mean	St. Dev.	MCSE
1:3000	8.813 E-2	6.635 E-4	4.538 E-5
1:10,000	8.819 E-2	6.623 E-4	2.528 E-5
10,000:100,000	8.823 E-2	6.871 E-4	7.576 E-6
3000:10,000	8.822 E-2	6.597 E-4	2.626 E-5



### 5.3.2 Values for Apparent Thermal Diffusivity

We summarize our results for each dataset in Tables 5.16-5.18. The values for  $k$  in Section 5.2 are included as well as results for  $k$  using the methods of Lu, Johansen, DFT, amplitude and phase methods. See Section 2.3 for details on these methods including the resampling strategy.

We choose a sample of size 7001 for each dataset using MCMC. The statistics here are with respect to the marginal posterior density for  $k$  [47, 87].

In Table 5.17, we see that the Lu estimate is larger than the Johansen estimate. Eslinger, in [36], attributes this difference in the empirical estimates of Lu and Johansen to the dry soil in the desert region. These estimates depend on the water content of the soil in different ways and this difference is especially apparent for low water content.

The statistics for the amplitude, phase and DFT-based methods are the sample statistics from the estimates computed with resampled data.

The values for  $k$  found using MCMC are consistent with the values found using the experiments, optimization, amplitude, phase and DFT methods. Furthermore, we obtain statistics about the variance of  $k$  using MCMC.

Table 5.16: Statistics of  $k$  estimates for Dataset A. Units are in  $m^2/hr$  [36].

Method	Estimates	Median	Mean	St. Dev.
Lu	1	1.250 E-3	-	-
Johansen	1	1.292 E-3	-	-
FEM-PEST	1	2.071 E-3	-	-
DFT	1800	1.511 E-3	1.480 E-3	3.826 E-4
Amplitude	105	1.173 E-3	1.284 E-3	4.362 E-4
Phase	105	2.668 E-3	5.838 E-3	1.724 E-3
MCMC	7001	1.682 E-3	1.682 E-3	2.515 E-5

Table 5.17: Statistics of  $k$  estimates for Dataset B. Units are in  $m^2/hr$  [36].

Method	Estimates	Median	Mean	St. Dev.
Lu	1	1.239 E-3	-	-
Johansen	1	9.308 E-4	-	-
FEM-PEST	1	9.277 E-4	-	-
DFT	6600	9.437 E-4	9.518 E-4	6.669 E-5
Amplitude	210	9.271 E-4	9.248 E-4	1.077 E-4
Phase	209	1.100 E-3	1.099 E-3	2.555 E-4
MCMC	7001	1.010 E-3	1.011 E-3	5.135 E-6

Table 5.18: Statistics of  $k$  estimates for Dataset H. Units are in  $m^2/hr$  [36].

Method	Estimates	Median	Mean	St. Dev.
Lu	1	2.770 E-3	-	-
Johansen	1	2.911 E-3	-	-
FEM-PEST	1	2.793 E-3	-	-
DFT	10920	3.075 E-3	3.120 E-3	3.648 E-4
Amplitude	273	2.568 E-3	2.681 E-3	6.635 E-4
Phase	273	4.400 E-3	5.069 E-3	3.007 E-3
MCMC	7001	3.059 E-3	3.060 E-3	5.647 E-5

### 5.3.3 Pairwise Scatterplots

In Figure 5.12, we display pairwise joint scatterplots of the four model parameters for Dataset A. The shape of these joint posterior densities indicates the correlation between parameters. Parameter  $\hat{C}$  has little correlation with the other parameters, while the correlations for pairs  $(\alpha, C)$  and  $(\alpha, \omega)$  are 0.6687 and 0.6352, respectively. We also observe a correlation of 0.424 between  $C$  and  $\omega$ .

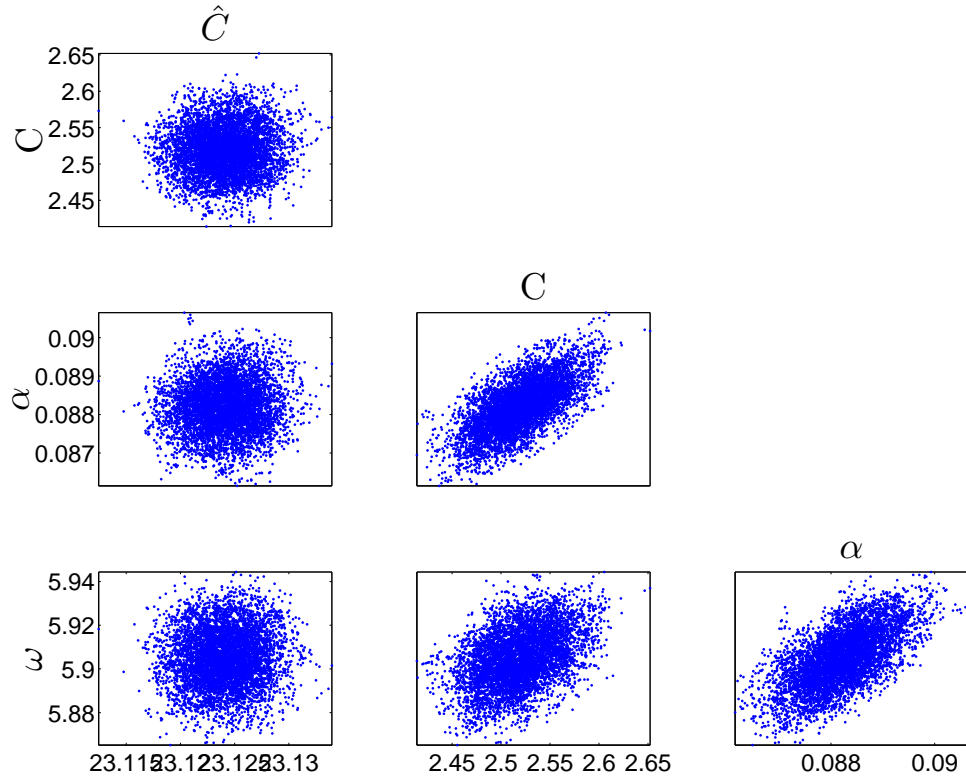


Figure 5.12: Pairwise marginal posterior densities for the model parameters, Dataset A.

Figures 5.13 and 5.14 are the pairwise scatterplots for Datasets B and H, respectively. We again see little correlation between  $\hat{C}$  and the other parameters. The correlations for the pairs  $(\alpha, C)$  and  $(\alpha, \omega)$  are 0.5988 and 0.5815, respectively, for Dataset B and 0.7883 and 0.7927, respectively, for Dataset H. The correlations between  $C$  and  $\omega$  are 0.3744 and 0.618 for Datasets B and H, respectively.

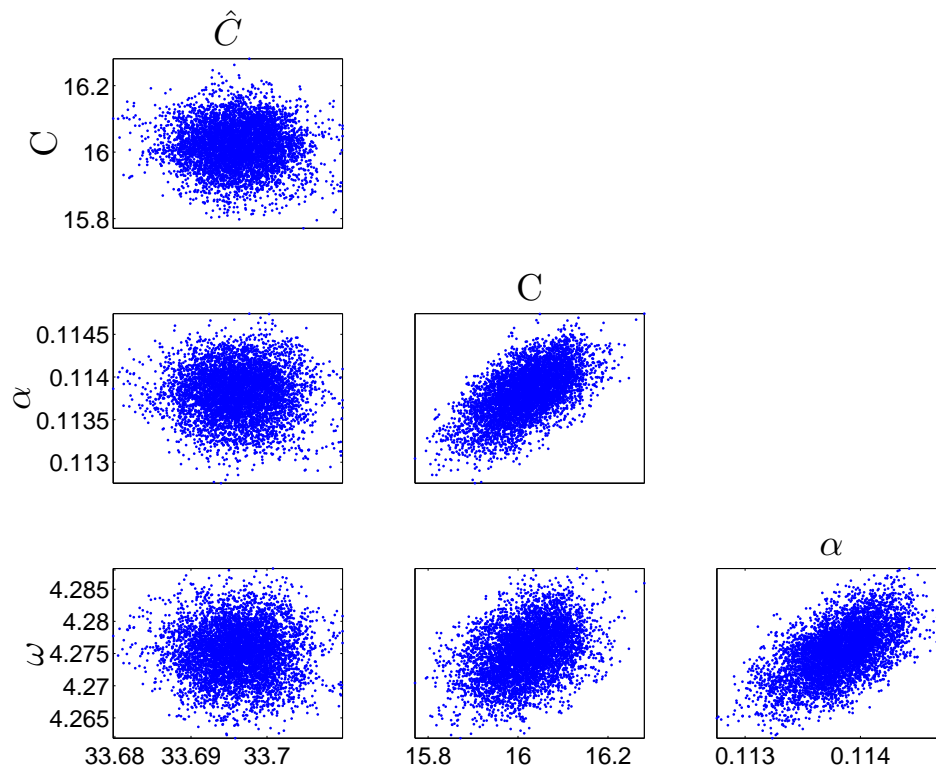


Figure 5.13: Pairwise marginal posterior densities for the model parameters, Dataset B.

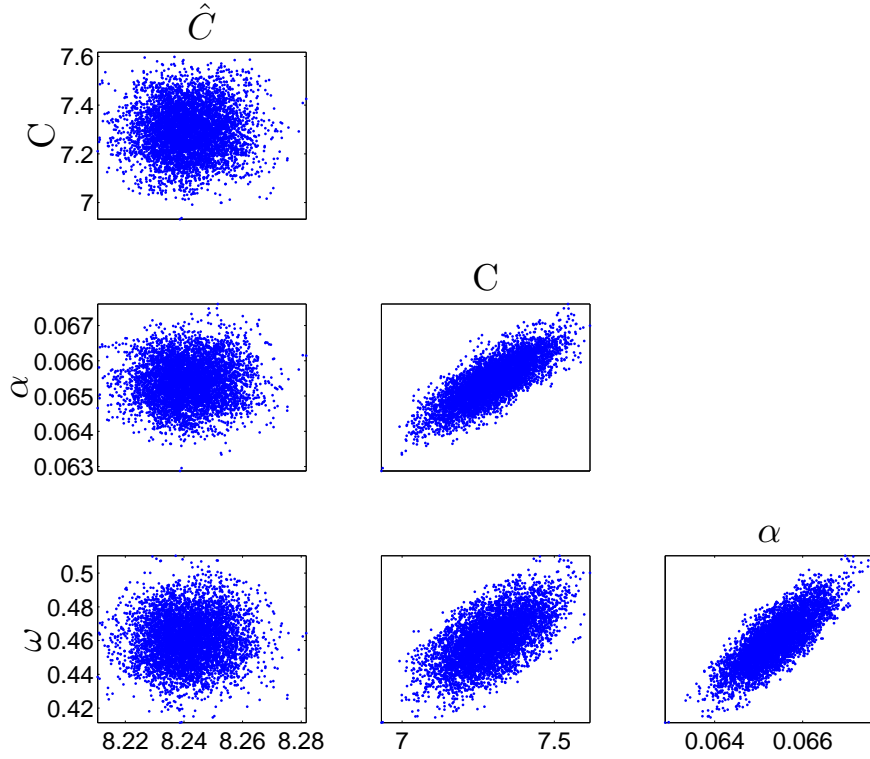


Figure 5.14: Pairwise marginal posterior densities for the model parameters, Dataset H.

### 5.3.4 Prediction Intervals

For further analysis, we produce prediction intervals for future observations  $y_{new}$  given our posterior density.

In Figures 5.15-5.17, we plot the prediction intervals over time for each dataset at depths 1cm and 30cm. The computations and plots are created using DRAM [65] with a random sample of 1000 from the 7001 samples from the posterior.

Figure 5.15 contains our model, data and 95% prediction intervals for Dataset A. Despite the data points outside of the prediction interval seen in the 1cm depth, over 95% of our data is within the prediction intervals for the entire dataset. The temperature spikes seen in our data

near the surface during the first two days is likely the result of a break in cloud cover in the rainforest. Measurements of the shortwave and longwave radiation support this conclusion, as discussed in Section 5.1. Even though our model does not capture this phenomenon near the surface, we are able to generate prediction intervals that contain the correct percentage of data and estimates for  $k$  that are consistent with estimates found using other methods.

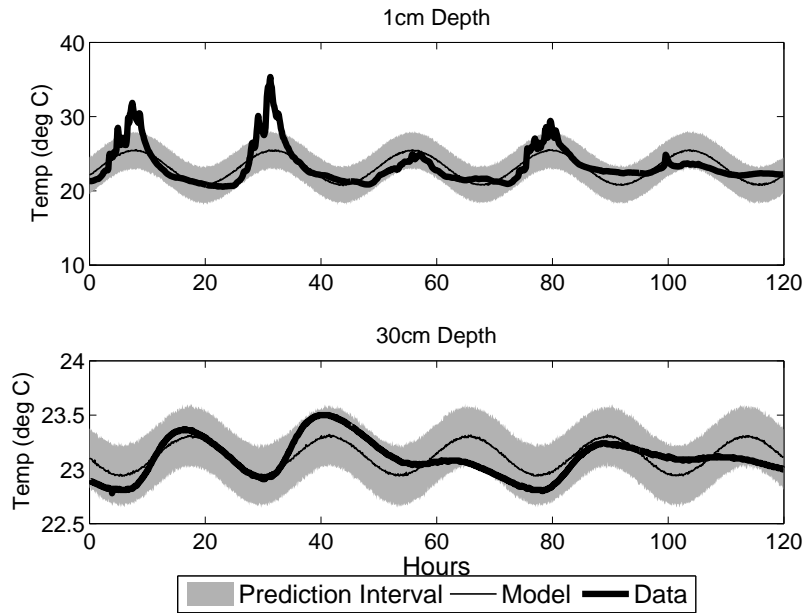


Figure 5.15: Model, data and prediction intervals for Dataset A.

Figure 5.16 contains the same information for Dataset B. The data appears to fall generally within the prediction intervals, though the actual percentage of data points within the intervals is 94.75%. Since the computation of the prediction intervals is not deterministic, we consider the prediction intervals for Dataset B satisfactory. The model seems unable to capture a trend visible at 30cm depth but still captures over 86% of the data at 30cm depth.

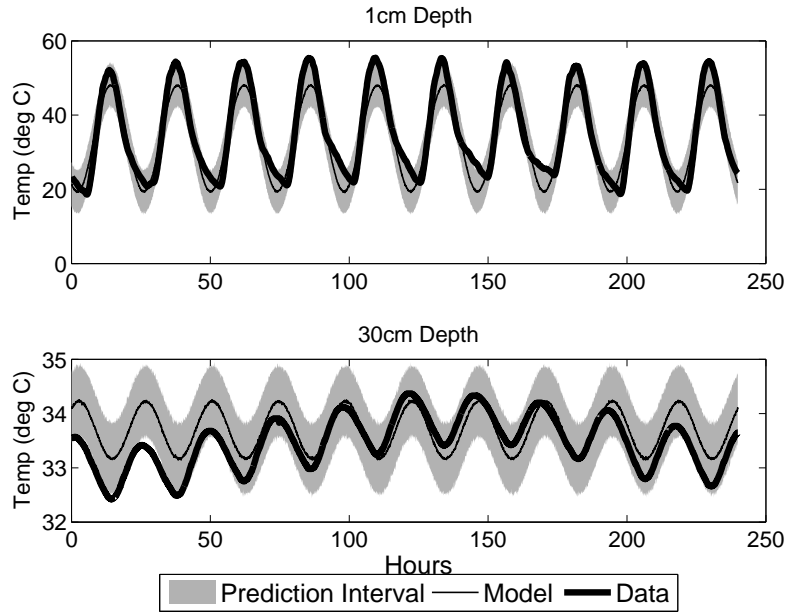


Figure 5.16: Model, data and prediction intervals for Dataset B.

The results for Dataset H are shown in Figure 5.17. Here, nearly 95% (94.9%) of the data is within the prediction intervals. The least amount of data falls within the intervals at 30 cm depth, again apparently the result of an inadequate model.

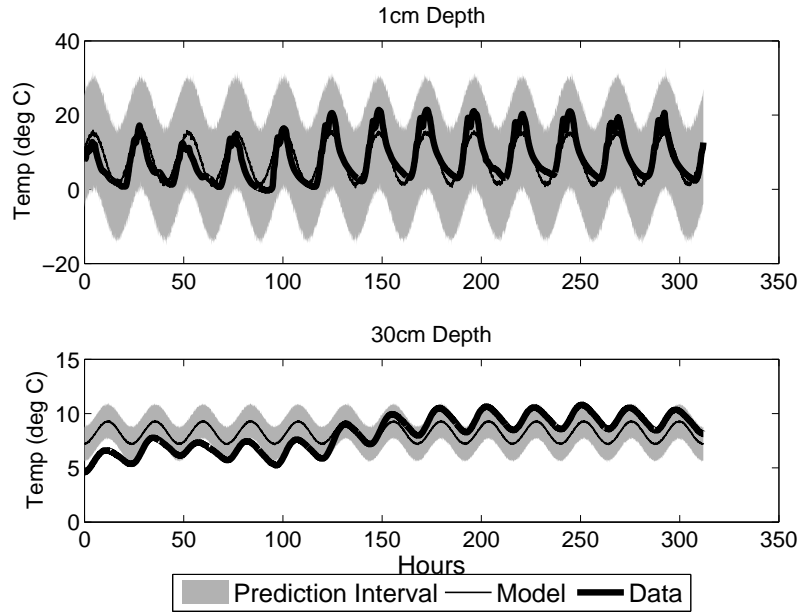


Figure 5.17: Model, data and prediction intervals for Dataset H.

We superimpose the least squares results using ADH as a model in Figures 5.18-5.20. We note that the results using PEST with ADH as a model are a much better match to the data.



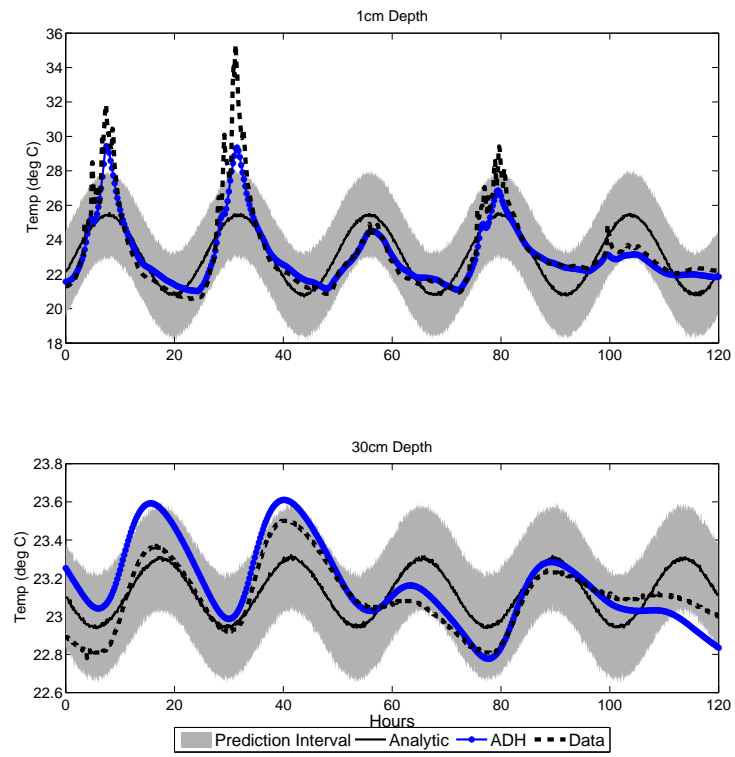


Figure 5.18: Model, data, prediction intervals, and PEST+ADH results for Dataset A.

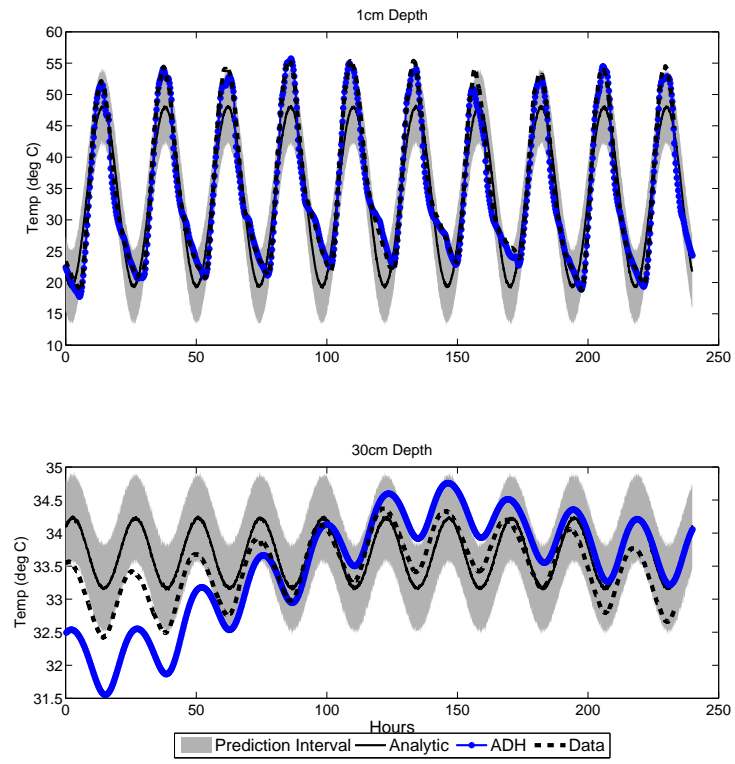


Figure 5.19: Model, data, prediction intervals, and PEST+ADH results for Dataset B.

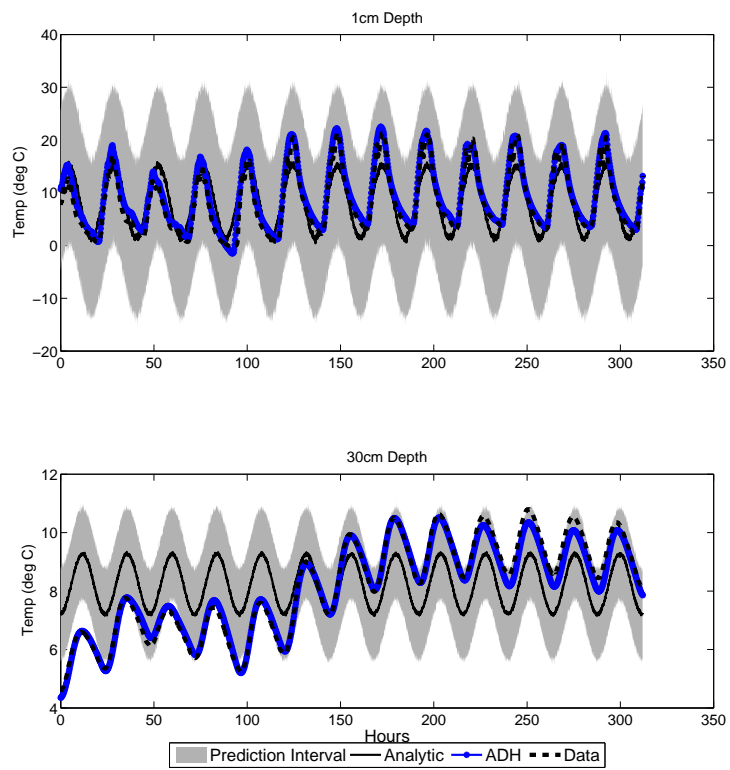


Figure 5.20: Model, data, prediction intervals, and PEST+ADH results for Dataset H.

## Chapter 6

# Results: Calibration of the Thermal Soil Parameters in a Variably Saturated Domain

In this chapter we describe our use of Bayesian inference to calibrate thermal properties of variably saturated soil with the Adaptive Hydrology simulation. Recall the model for a variably saturated domain from Section 2.2.2. We solve the equations

$$(c_p\rho)u_t - \nabla \cdot (\kappa \nabla u) = 0, \quad (6.1)$$

and

$$S_s S_w(\psi) \frac{\partial \psi}{\partial t} + \phi \frac{\partial S_w(\psi)}{\partial t} = \nabla [K_s k_r(\psi) \nabla (\psi + z)], \quad (6.2)$$

which are coupled by the dependence of the volumetric heat capacity  $c_p\rho$  in Equation (6.1) upon the saturation  $S_w$  through Equation (6.3):

$$(c_p\rho) = (1 - \phi)c_{p,s}\rho_s + \phi(S_w c_{p,w}\rho_w + (1 - S_w)c_{p,g}\rho_g). \quad (6.3)$$

For the domains with a constant saturation, we calibrated the apparent thermal diffusivity,  $k$ . However,  $k$  is no longer constant for variably saturated domains. Thus, in these unsaturated domains with rain events, we study the specific heat of the soil,  $c_{p,s}$ .

We first describe the variably saturated datasets, then we discuss the results of using a traditional least squares method to calibrate the thermal parameters. For Bayesian inference, we first perform our analysis on the data from Chapter 5 in order to test our methods. Then, in Section 6.3.2, we perform and describe the analysis with the variably saturated data.

## 6.1 Data

The variably saturated data is also from the data collection described in Section 5.1. We now select a 24-hour period during the data collection containing a rain event for our calibration. We also ensure that the 48-hour period preceding the calibration period does not contain a rain event. We find a 48-hour ramp-up period necessary for the variably saturated ADH simulations. Dataset B is from the desert and contains no precipitation, so we do not perform the variably saturated analysis for this dataset.

We do not have saturation measurements, so we take the measured water content and divide it by the porosity. For Dataset A, we note that the water content plateaus as a result of the rain event, so we assume the soil is fully saturated during this time of constant saturation. Thus, we are able to determine a porosity of 0.576 for Dataset A. For Dataset H, we use the porosity from the DBS&A lab results: 0.377.

### 6.1.1 Dataset A

For Dataset A, we ramp up our simulations over Julian days 128 and 129 and use day 130 as our calibration data. This dataset contained missing values. We selected this portion of the data as it was complete.

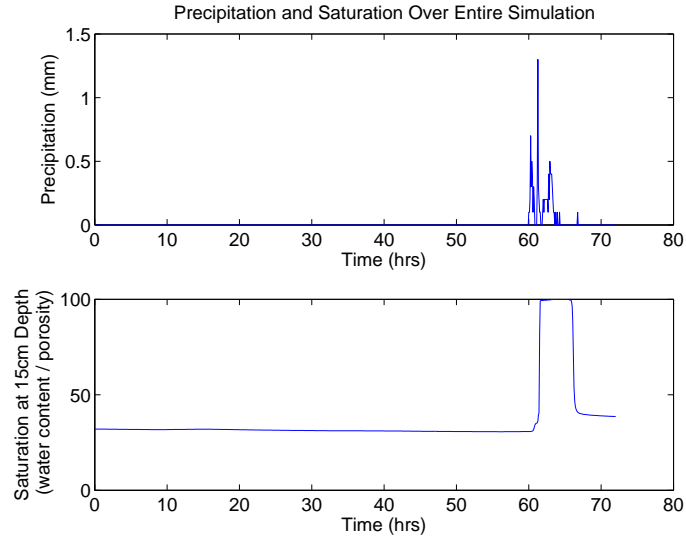


Figure 6.1: Dataset A precipitation and saturation over the ramp-up and calibration periods.

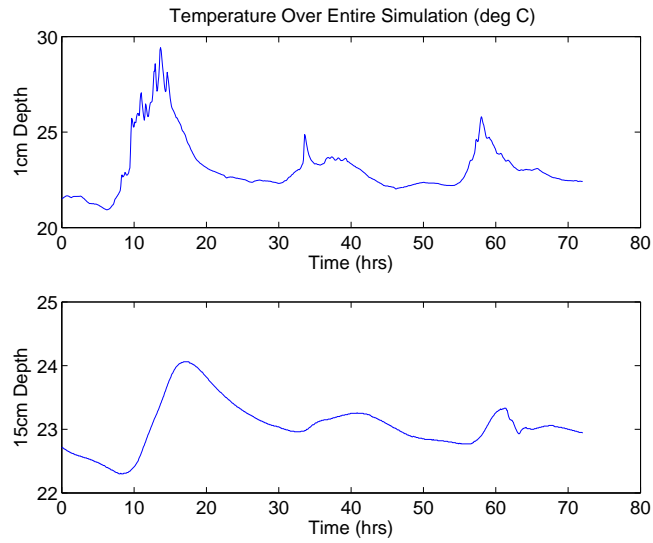


Figure 6.2: Dataset A temperature at 1cm and 15cm depth over the ramp-up and calibration periods.

### 6.1.2 Dataset H

For Dataset H, we ramp up our simulations over Julian days 247 and 248 and then use day 249 as the calibration data.

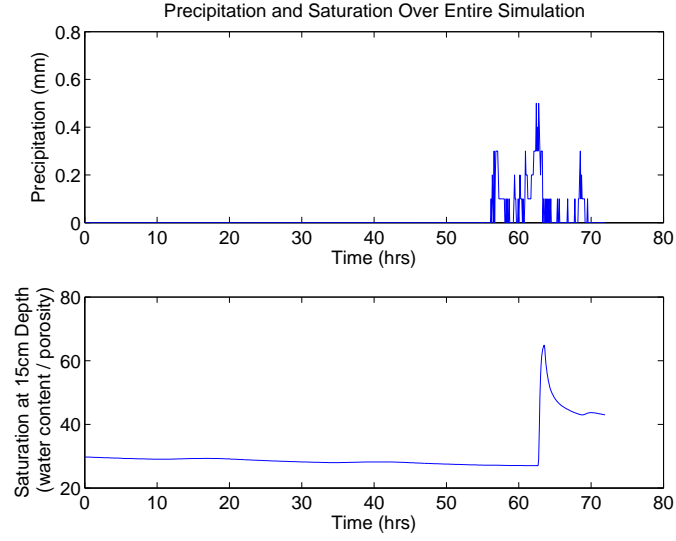


Figure 6.3: Dataset H precipitation and saturation over the ramp-up and calibration periods.

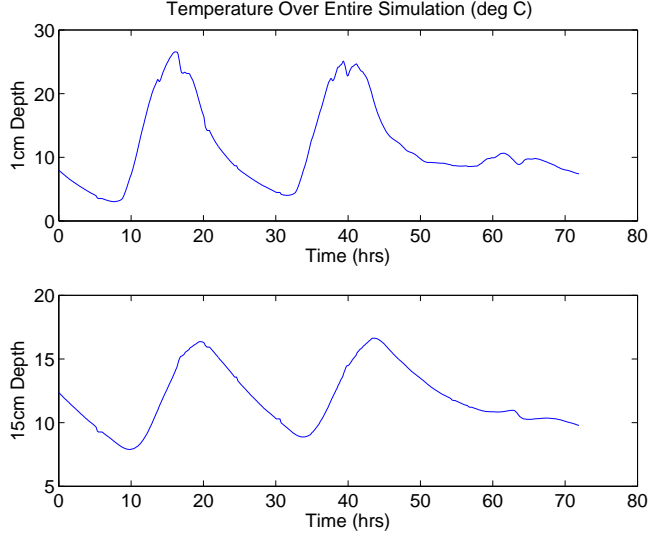


Figure 6.4: Dataset H temperature at 1cm and 15cm depth over the ramp-up and calibration periods.

## 6.2 Least Squares (PEST) with Numerical Model (ADH)

We first determine the least squares solution to our problem using PEST and ADH. We must supply ADH with initial conditions for temperature and total head. We determine the initial temperature by computing the average temperature at 30cm depth over the 48-hour ramp-up period.

We choose the initial total head by first determining the average saturation over the 48-hour ramp-up period. We divide the average water content by the porosity, thus we fix the porosity throughout our simulations. For Dataset A, we notice that the water content plateaus over a period of time during the data comparison period so that the soil must fully saturated during this interval. We divide the maximum measured water content by 100 to determine the porosity for Dataset A. Then we compute the average saturation over the ramp-up period using this value for porosity and fix porosity throughout our analysis. For Dataset H, we fix the porosity at 0.377, the value determined by the DBS&A lab. Once we have the average saturation over



the ramp-up period, we set the initial total head so that the initial saturation values ADH computes are as close as possible to the values we have determined. We use only increments of meters to fix the initial total head.

We then compute the Jacobian matrix using all of the heat and groundwater parameters and use subset selection to determine an appropriate set of parameters to fit.

### **6.2.1 Dataset A**

For this dataset, the soil measurements were inadequate to perform the intended analysis. We tried fixing and fitting numerous combinations of parameters, fitting only the saturation, and fitting the initial head and were unable to make progress. Thus, we have determined fitting Dataset A in the unsaturated case is beyond the scope of this work.

### **6.2.2 Dataset H**

For Dataset H, we use an initial temperature of 11.746 degrees Celsius and an initial total head of -10.0 meters, corresponding to a saturation of  $\approx 29\%$ . We use the same temperature weights for Dataset H as in Section 5.2.3 and weight each saturation value 1000. We use the same constant values as in Table 5.2.

The seven largest singular values are listed in Table 6.1.

Table 6.1: Singular values of Jacobian computed with all initial parameters for Dataset H, unsaturated case.

1.1908e+05
5.2852e+03
1.3640e+02
9.1177e+01
2.7599e+01
1.7573e+01
1.1366e+01

We again use the subset selection algorithm [46] and code implemented in [55] to determine three parameters to fit: the specific heats of the gravel layer and deep soil as well as the albedo of the gravel layer. We find the specific heat in the gravel layer is difficult to fit using PEST and determine that our model is not adequate in modeling some physics at the surface, likely due to radiation. Thus, we fix the specific heat in the gravel layer. We also note that the remaining two parameters are heat parameters only and have no affect on the groundwater physics. Thus, we only use the temperature probe data to fit the specific heat of the deep soil and the albedo of the gravel layer.

We use the same initial values for the specific heat of the deep soil and the albedo of the gravel layer that we used in the dry case. We fix the specific heat in the gravel layer at 5.332 E-4, the value determined using PEST and ADH in the dry case, see Section 5.2.3. We also fix the emissivity in the gravel layer at 0.808947, the value determined using PEST and ADH in the dry case, see Section 5.2.3. We fix all other heat parameters as before, listed in Table 5.13. Additionally, we fix the same groundwater parameters as before using the values in Table 6.2.

Table 6.2: Fixed groundwater parameters for unsaturated case, Dataset H. These values are for both the gravel and deep soil layers.

Parameter (ADH Card, symbol)	Value
Hydraulic conductivity (K, $K_s$ )	0.0657
Residual saturation (RSD, $S_r$ )	0.01
van Genuchten alpha (VGA, $\alpha$ )	0.75
van Genuchten N (VGN, $n$ )	1.6
van Genuchten max capillary pressure for curve (VGP)	40
Specific storage (SS, $S_s$ )	0.0001
Tortuosity (TOR)	0.5
Longitudinal dispersivity (DPL)	1
Transverse dispersivity (DPT)	0.1

The optimization results are in Table 6.3.

Table 6.3: Results of PEST optimization for Dataset H, unsaturated case.

Parameter	Symbol	Value
Specific Heat of Solid, Deep Soil	$c_{p,s}$	1.696 E-4
Albedo, Gravel Layer	$\alpha_g$	0.378328

We are most interested in fitting the specific heat of the deep soil. We see that our values for the specific heat of the deep soil are consistent with those found in the constant saturation (dry) case. These values are in Table 6.4.

Table 6.4: Specific heat optimization results for Dataset H (Fraser/Highlands), unsaturated and dry cases (PEST + ADH).

Parameter	Unsat Value	Dry Value
Specific Heat of Solid, Deep Soil	1.696 E-4	1.890 E-4

Figure 6.5 is ADH evaluated with the optimal parameter values plotted with the temperature data at two depths. We see good agreement between the model output and the data. A plot with all depths is included in Appendix D

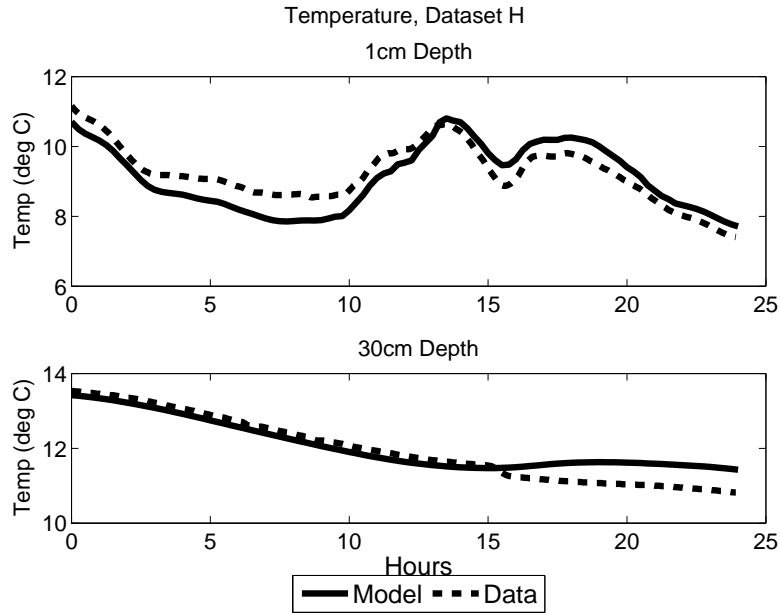


Figure 6.5: Dataset H model and temperature data at 1cm and 30cm depth over calibration period.

Figure 6.6 is the ADH evaluated with the optimal parameter values plotted with the saturation data at 15cm depth. We see that ADH does not capture the spike in saturation around 15 hours, but this does not significantly affect fit to the temperature data.

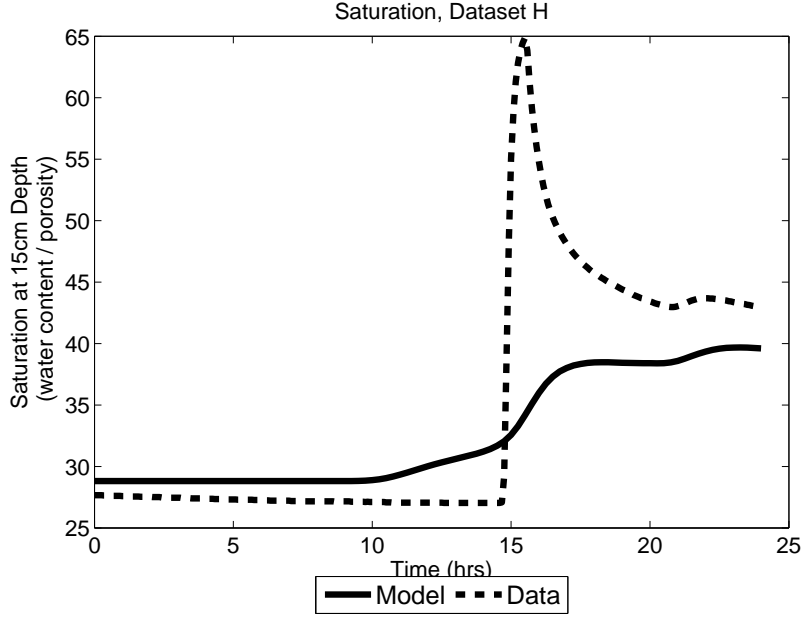


Figure 6.6: Dataset H saturation 15cm depth over calibration period.

### 6.3 Bayesian Inference (DRAM) with Numerical Model (ADH)

We now perform Bayesian inference with ADH as our model for heat transport. First we use the data with constant saturation and then move on to the unsaturated case. For both cases, we only use a 24-hour calibration period as the DRAM simulations require many function evaluations and a smaller time interval will reduce computational time.

We compute the least squares solution using PEST and use the results to generate initial values for the DRAM simulations as well as to validate our results.

#### 6.3.1 Simulations with Constant Saturation

We choose a 24-hour calibration period to reduce ADH computation time. The days we use are Julian days 126, 222, and 268 for Datasets A, B, and H, respectively. A 24-hour ramp-up period is sufficient for the datasets with no precipitation, so our total simulations are for a

48-hour time period.

The initial conditions we use are in Table 6.5.

Table 6.5: Initial conditions for constant saturation 24-hour simulations.

Dataset	Initial Temp (deg C)	Initial Head	Corresponding Saturation
A	23.013	-215.0	$\approx 39\%$
B	33.017	-8.0	$\approx 14\%$
H	8.396	-8.0	$\approx 32\%$

## PEST Results

We optimize over the same parameters as in Chapter 5 for each dataset.

For Dataset A, we calibrate the specific heat of the solid  $c_{p,s}$ , emissivity  $\epsilon_g$ , and albedo  $\alpha_g$ . The remaining fixed thermal parameters and initial values for the calibrated parameters are the same as in Table 5.7.

For Dataset B, we calibrate the specific heat of the solid  $c_{p,s}$ , emissivity  $\epsilon_g$ , and albedo  $\alpha_g$ . The remaining fixed thermal parameters and initial values for the calibrated parameters are as in Table 5.10.

For Dataset H, we calibrate the specific heats of both the gravel and deep soil layers as well as the emissivity of the gravel layer. The remaining fixed thermal parameters and initial values for the calibrated parameters are as in Table 5.13.

The PEST optimization results for each dataset are in Tables 6.6-6.8, and temperature plots for the results at two depths are in Figures 6.7-6.9. Plots of the results at all depths are in Appendix D.

Table 6.6: Results from PEST optimization, Dataset A constant saturation 24-hour case.

Parameter	Symbol	Value
Specific Heat of Solid	$c_{p,s}$	2.650 E-4
Emissivity	$\epsilon_g$	0.840188
Albedo	$\alpha_g$	0.509577

Table 6.7: Results from PEST optimization, Dataset B constant saturation 24-hour case.

Parameter	Symbol	Value
Specific Heat of Solid	$c_{p,s}$	4.260 E-04
Emissivity	$\epsilon_g$	0.937750
Albedo	$\alpha_g$	0.137470

Table 6.8: Results from PEST optimization, Dataset H constant saturation 24-hour case.

Parameter	Symbol	Value
Specific Heat of Solid, Deep Soil	$c_{p,s}$	1.693 E-4
Specific Heat of Solid, Gravel Layer	$c_{p,s}$	8.508 E-4
Emissivity, Gravel Layer	$\epsilon_g$	0.979113

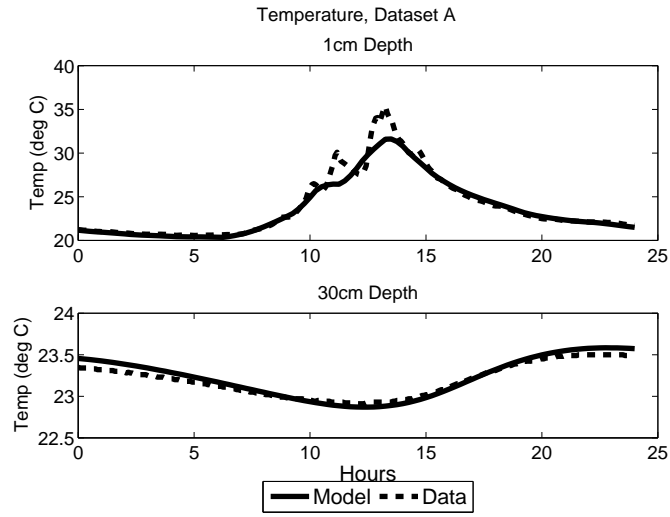


Figure 6.7: Dataset A model and temperature data at 1cm and 30cm depth over calibration period.

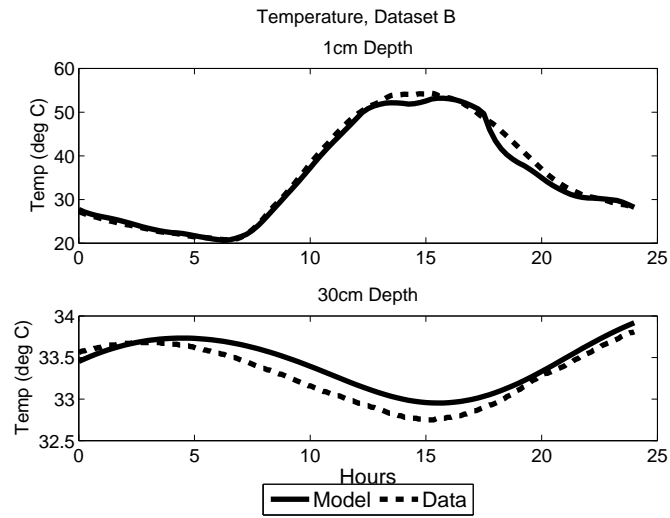


Figure 6.8: Dataset B model and temperature data at 1cm and 30cm depth over calibration period.



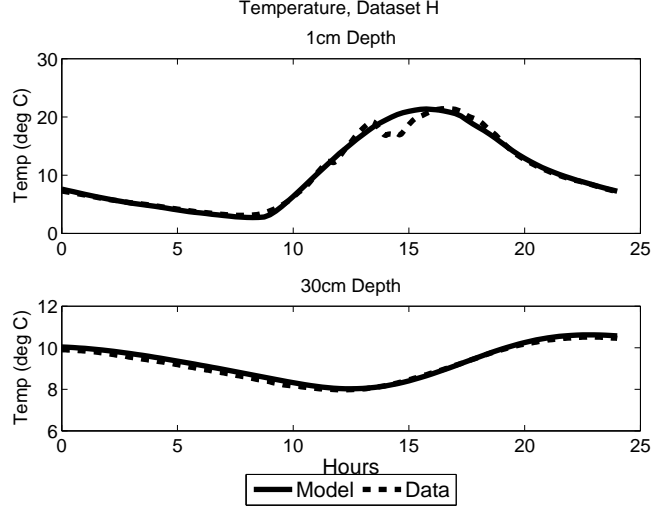


Figure 6.9: Dataset H model and temperature data at 1cm and 30cm depth over calibration period.

## DRAM

Because of the substantial computation time of the ADH simulations, we must run our DRAM simulations in parallel. We choose to run 8 independent, parallel MCMC chains each of length 2500 samples [101] because we have 8 processors available. We discard the first 999 samples of each chain as burn-in and thus have a sample of  $8 \times 1501 = 12008$  from each posterior density.

We use the results from the least squares PEST optimizations to generate random initial values for each chain. For the model parameters, we generate initial values from the multivariate normal distribution with the PEST optimal values as the mean and the PEST output covariance matrix as the covariance. If the initial values generated are negative, we multiply the covariance matrix by 0.01 to encourage the choice of positive values.

For the error variance, we generate random initial values from the gamma distribution with shape parameter 0.1 and scale parameter 0.1. This distribution has mean 0.01 and variance 0.001 [31]. Since the specifications of the thermocouples give a variance of 0.01, we believe this is an appropriate distribution for selecting initial values of the error variance.

We are careful to create an independent random stream of numbers on each processor to ensure independence between chains. We use the Matlab Multiplicative lagged Fibonacci generator, keyword `mlfg6331_64`, as it is appropriate for parallel random number generation [67,68]. Details for the implementation are in Appendix C.

**Diagnostics** As in Chapter 5, we show diagnostics for Dataset A as they are representative of all three datasets. Again, diagnostics for the other two datasets are in Appendix C. Figure 6.10 is the first 10 samples from each chain. Here we can see the different initial values. Figure 6.11 is the entire history of all of the chains plotted together for each parameter. It appears that the burn-in period is very small but it is difficult to discern because the chains are layered on top of one another. We plot the history for an individual chain in Figure 6.12 to show that there is indeed little burn in period. We take a conservative approach and use samples 1000 to 2500 from each chain to ensure we obtain a sample from the stationary, or posterior, distribution.

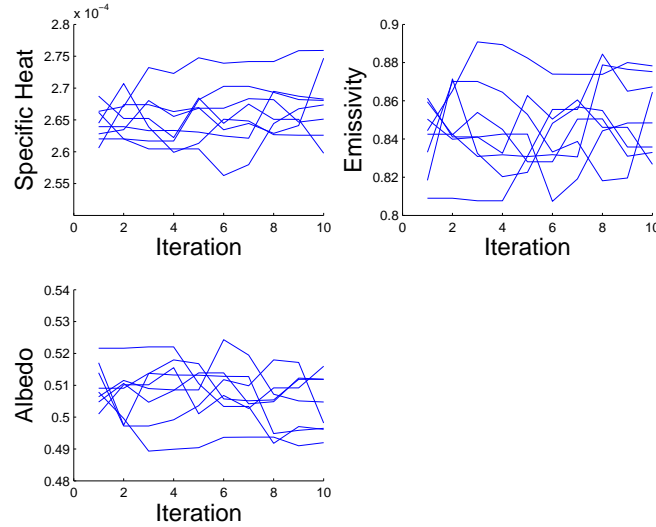


Figure 6.10: Dataset A chain histories for the first 10 samples of each individual chain.

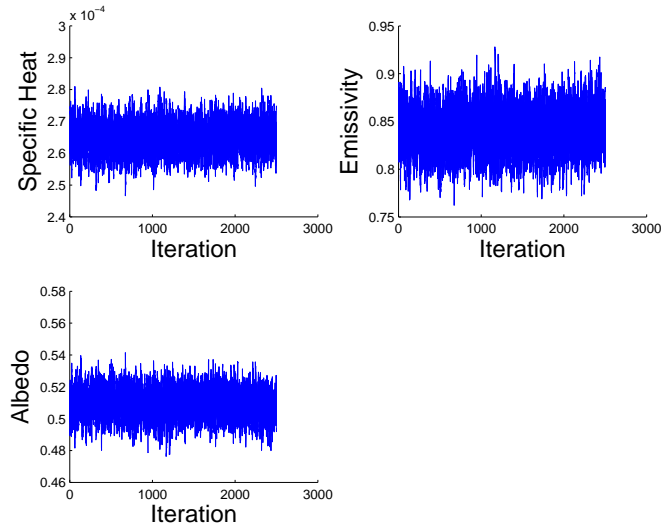


Figure 6.11: Dataset A chain histories for all eight chains.

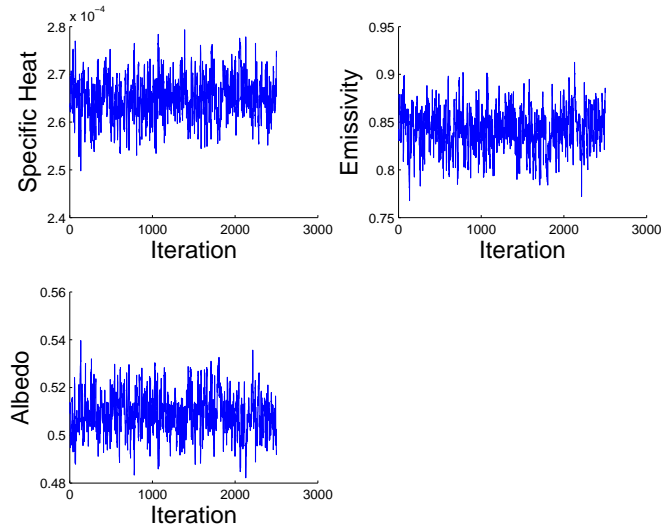


Figure 6.12: Dataset A chain history for one individual chain.

Figure 6.13 is the autocorrelation function showing the correlation between successive sam-

ples of the burned-in portion of the chain. We note that the autocorrelation decreases rapidly as the lag between samples increases, indicative of a sufficiently independent sample size.

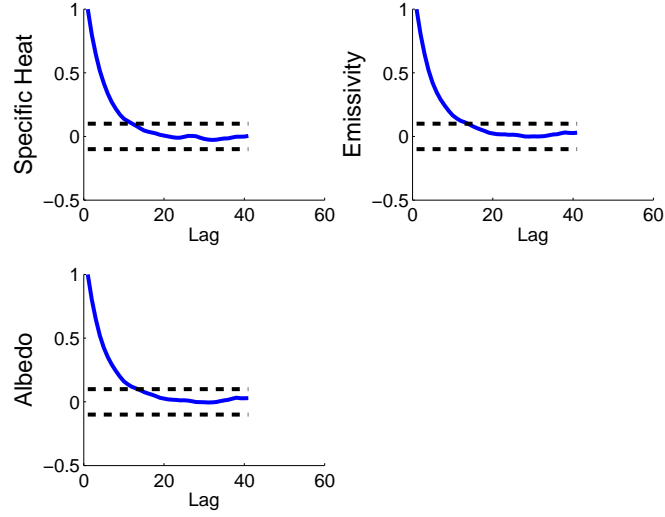


Figure 6.13: Autocorrelation function for Dataset A.

In Table 6.9, we provide the mean, standard deviation, and Monte Carlo standard error for each parameter from the sample from the posterior. The Monte Carlo standard errors are several orders of magnitude less than the sample means, which indicates a sufficiently large sample size.

Table 6.9: Mean, standard deviation and Monte Carlo standard error (MCSE) for each parameter from Dataset A.

Parameter	Mean	St. Dev.	MCSE
Specific Heat	2.650 E-04	4.400 E-06	1.416 E-07
Emissivity	8.415 E-01	2.142 E-02	8.287 E-04
Albedo	5.091 E-01	8.457 E-03	3.371 E-04

**Results** In Figures 6.14-6.16 we plot the model, data, and prediction intervals at depths of 1cm and 30cm for each dataset. Prediction intervals for all depths are in Appendix C. Overall, the model evaluated with values from the posterior density for our parameters matches the data. At 1cm depth in Datasets A and H, we see jagged peaks that the model does not capture. These are likely the result of shadows or cloud cover. This fit is an improvement over the fit in Chapter 5 with the analytic solution to the heat equation as a model.

Despite the data outside the prediction intervals at 1cm depth in Figure 6.14, over 97% of the data is within the prediction intervals over all depths.

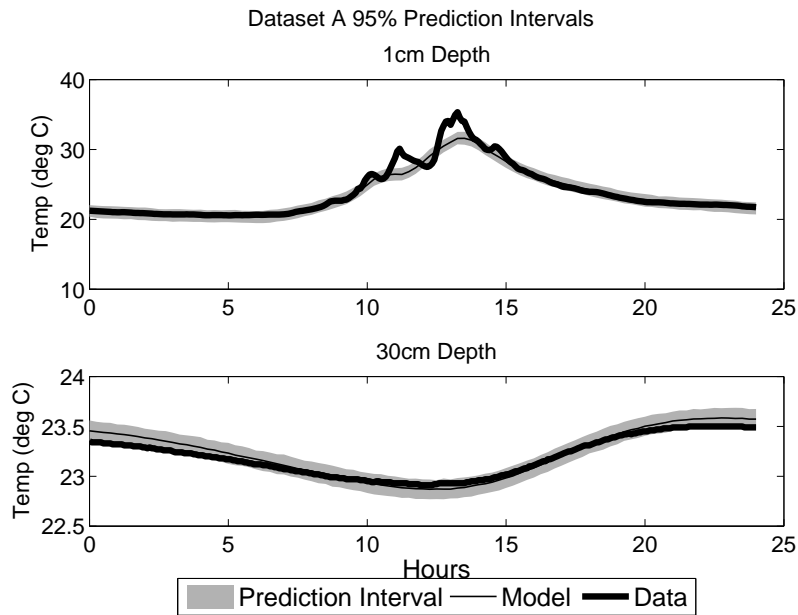


Figure 6.14: 95% Prediction intervals for Dataset A.

In Figure 6.15 we have the prediction intervals at two depths for Dataset B. Over all depths, 91.31% of the data is within the prediction intervals.

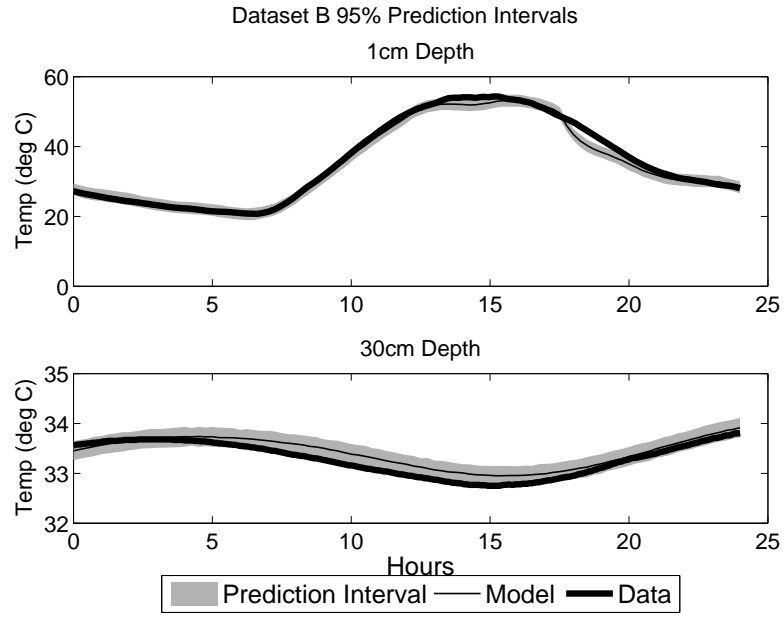


Figure 6.15: 95% Prediction intervals for Dataset B.

The prediction intervals for Dataset H, seen in Figure 6.16, contain over 92% of the data.

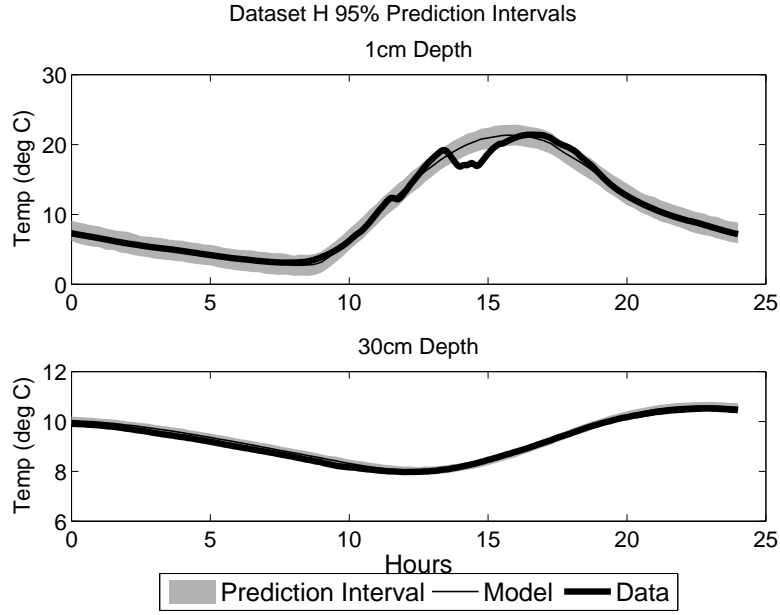


Figure 6.16: 95% Prediction intervals for Dataset H.

Figures 6.17 and 6.19 contain the pairwise scatterplots of the (burned-in) results for Datasets A and H, respectively. We see the strongest correlation between the albedo and emissivity for Datasets A and B, likely because they both influence the heat flux at the surface of our domain.

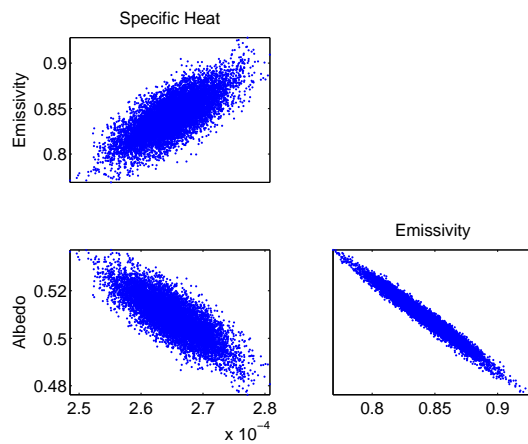


Figure 6.17: Pairwise scatterplots from burned-in portion of chain for Dataset A.

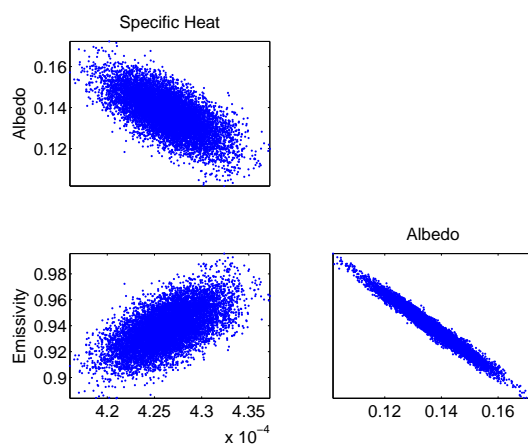


Figure 6.18: Pairwise scatterplots from burned-in portion of chain for Dataset B.



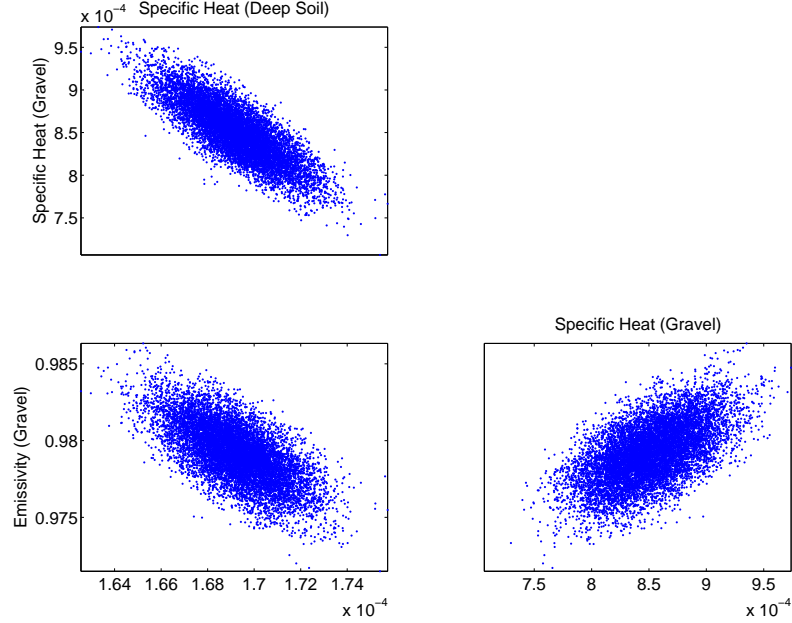


Figure 6.19: Dataset H pairwise plots, constant 24-hour case.

Finally, in Tables 6.10-6.12 we display the results statistics of our analysis. Here we include some results from the tables from Chapter 5 with our new results using a 24-hour time period for the constant saturation case. We see that our results are again consistent with those from the empirical methods as well as our previous results. The results on the 24-hour dataset using ADH with PEST and DRAM are very similar, seen in the last two rows of each table.

Table 6.10: Statistics of  $k$  estimates for Dataset A. Units are in  $m^2/hr$  [36].

<b>Method</b>	<b>Estimates</b>	<b>Median</b>	<b>Mean</b>	<b>St. Dev.</b>
Lu	1	1.250 E-3	-	-
Johansen	1	1.292 E-3	-	-
FEM-PEST (all dry data)	1	2.071 E-3	-	-
MCMC (all dry data)	7001	1.682 E-3	1.682 E-3	2.515 E-5
FEM-PEST (24-hour dry data)	1	1.759 E-3	-	-
MCMC - ADH (24-hour dry data)	12008	1.759 E-3	1.759 E-3	1.761 E-5

Table 6.11: Statistics of  $k$  estimates for Dataset B. Units are in  $m^2/hr$  [36].

<b>Method</b>	<b>Estimates</b>	<b>Median</b>	<b>Mean</b>	<b>St. Dev.</b>
Lu	1	1.239 E-3	-	-
Johansen	1	9.308 E-4	-	-
FEM-PEST (all dry data)	1	9.277 E-4	-	-
MCMC (all dry data)	7001	1.010 E-3	1.011 E-3	5.135 E-6
FEM-PEST (24-hour dry data)	1	9.809 E-4	-	-
MCMC - ADH (24-hour dry data)	12008	9.801 E-04	9.801 E-04	6.539 E-06

Table 6.12: Statistics of  $k$  estimates for Dataset H. Units are in  $m^2/hr$  [36].

Method	Estimates	Median	Mean	St. Dev.
Lu	1	2.770 E-3	-	-
Johansen	1	2.911 E-3	-	-
FEM-PEST (all dry data)	1	2.793 E-3	-	-
MCMC (all dry data)	7001	3.059 E-3	3.060 E-3	5.647 E-5
FEM-PEST (24-hour dry data)	1	2.974 E-3	-	-
MCMC - ADH (24-hour dry data)	12008	2.973 E-3	2.974 E-3	1.927 E-5

In Figure 6.20 we plot all values of  $k$  found in Tables 6.10-6.12. We are able to distinguish three different groups of values corresponding to the three datasets. There is some overlap between Datasets A and B due to the high estimate for  $k$  from Lu. As stated in Chapter 5, we believe this is because of the dependence of this estimate on the saturation, which is very low in this case.

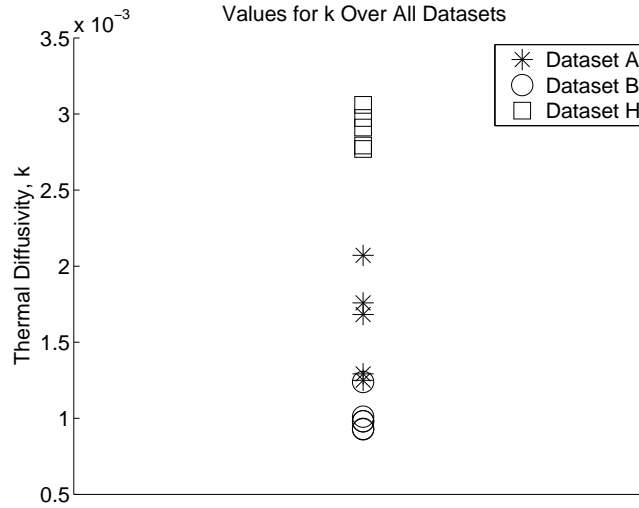


Figure 6.20: Estimates for  $k$  from all methods for all datasets.

Figure 6.21 is the same plot but with the log scale on the y-axis. The three groups are perhaps even more distinct in this plot.

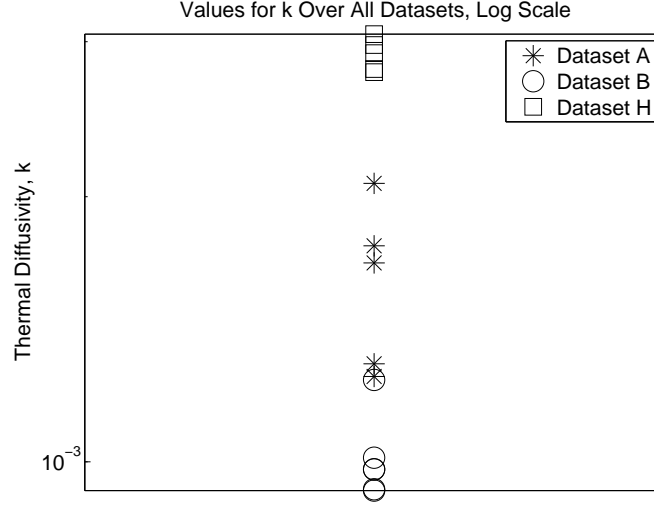


Figure 6.21: Estimates for  $k$  from all methods for all datasets with log scale on y-axis.

### 6.3.2 Variable Saturation

We now discuss the DRAM results for Dataset H in the unsaturated case. Recall from Section 6.2.2 the implementation details and results of using PEST to optimize over the specific heat  $c_{p,s}$  of the deep soil and the albedo  $\alpha_g$  of the gravel layer. Here we repeat the process in Section 6.3.1 of using ADH as a model for heat transport in the DRAM simulations.

#### Diagnostics

We plot the first 5 samples from each chain in Figure 6.22. Here we show that each chain had a different initial value, though they are close together.

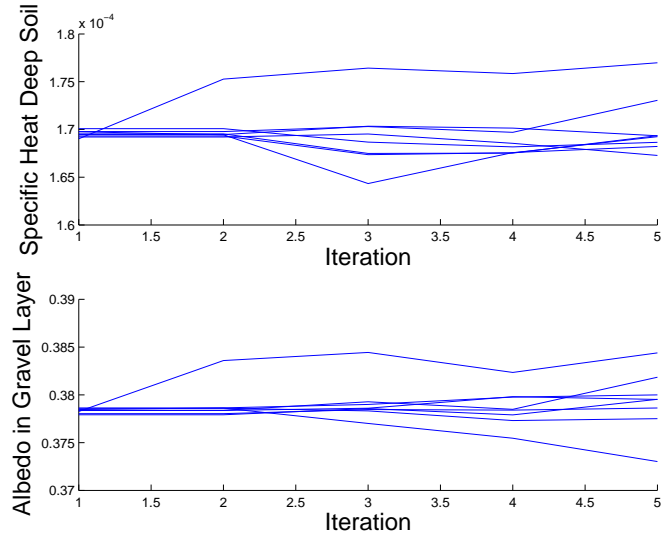


Figure 6.22: Dataset H chain histories for the first 5 samples of each individual chain.

In Figure 6.23 we plot the entire chain histories of all eight chains. See Figure 6.24 for the form of a chain history for an individual chain.

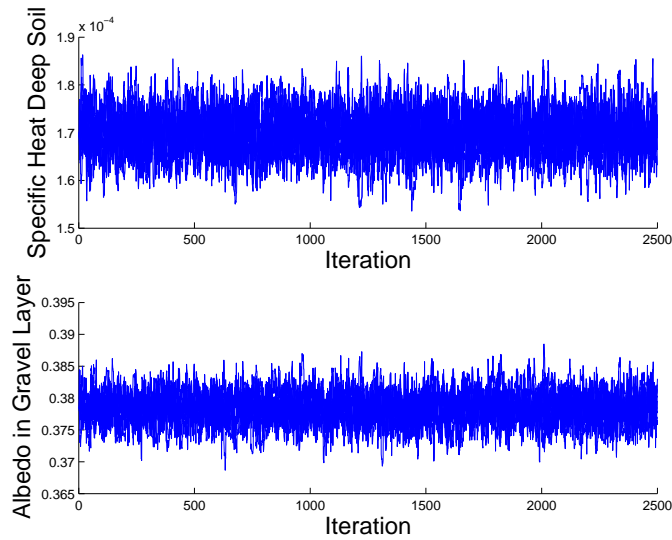


Figure 6.23: Dataset H chain histories for all eight chains.

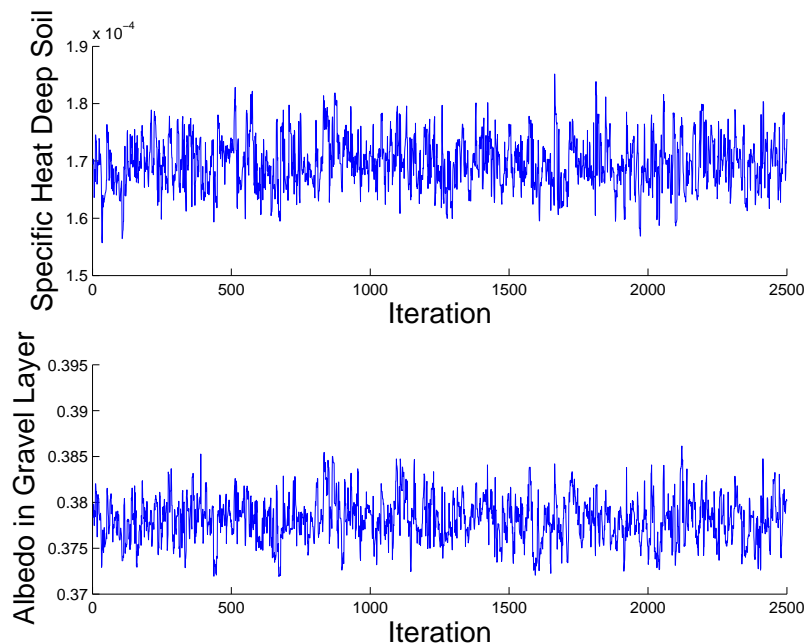


Figure 6.24: Dataset H chain history for one individual chain.

Figure 6.25 is the autocorrelation function showing the correlation between successive samples of the burned-in portion of the chain. We note that the autocorrelation decreases rapidly as the lag between samples increases, indicative of a sufficiently independent sample size.

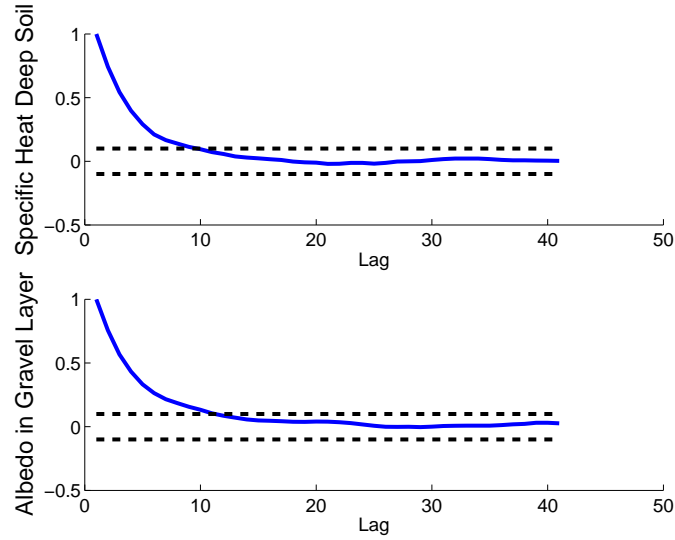


Figure 6.25: Autocorrelation function for Dataset H.

In Table 6.13 we show the mean, standard deviation, and Monte Carlo standard error for the samples for each parameter. Again, because the MCSE is several magnitudes less than the mean, we conclude we have obtained a sufficiently large sample from the posterior density.

Table 6.13: Mean, standard deviation and Monte Carlo standard error (MCSE) for each parameter from Dataset H.

Parameter	Mean	St. Dev.	MCSE
Specific Heat (Deep soil)	1.699 E-4	4.597 E-6	9.604 E-8
Albedo (Gravel layer)	3.783 E-1	2.513 E-3	7.598 E-5

### Prediction intervals

In Figures 6.26 and 6.27 we have the prediction intervals at depths of 1cm and 30cm and all depths, respectively. Near the surface, the prediction intervals contain 100% of the data. However, as the depth increases the model was unable to capture the drop in temperature due

to the rain event. Since the prediction intervals are centered around the model, they likewise fail to capture the data at increasing depth. This is a violation of the assumption that the errors are normally distributed with mean zero. Overall, however, these prediction intervals capture over 93% of the data.

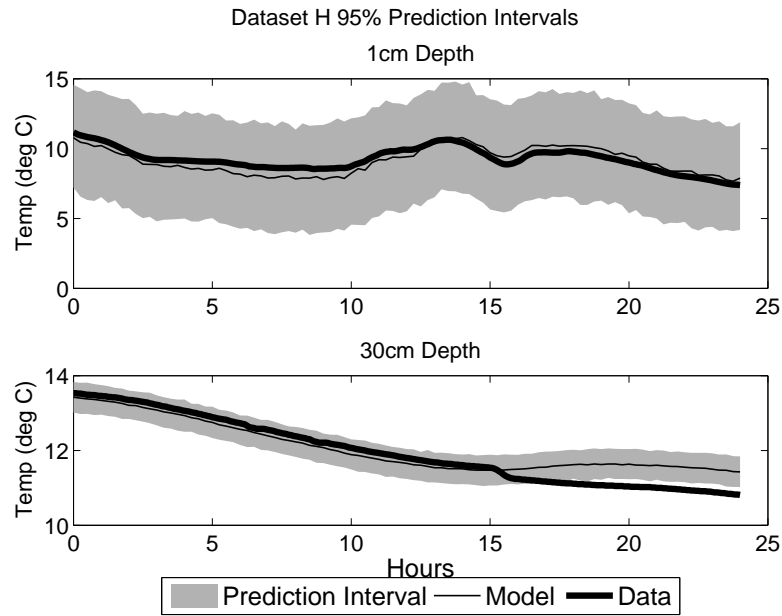


Figure 6.26: 95% Prediction intervals for Dataset H. Depths of 1 and 30cm.

Figure 6.27 is the prediction intervals at all depths.



Dataset H 95% Prediction Intervals

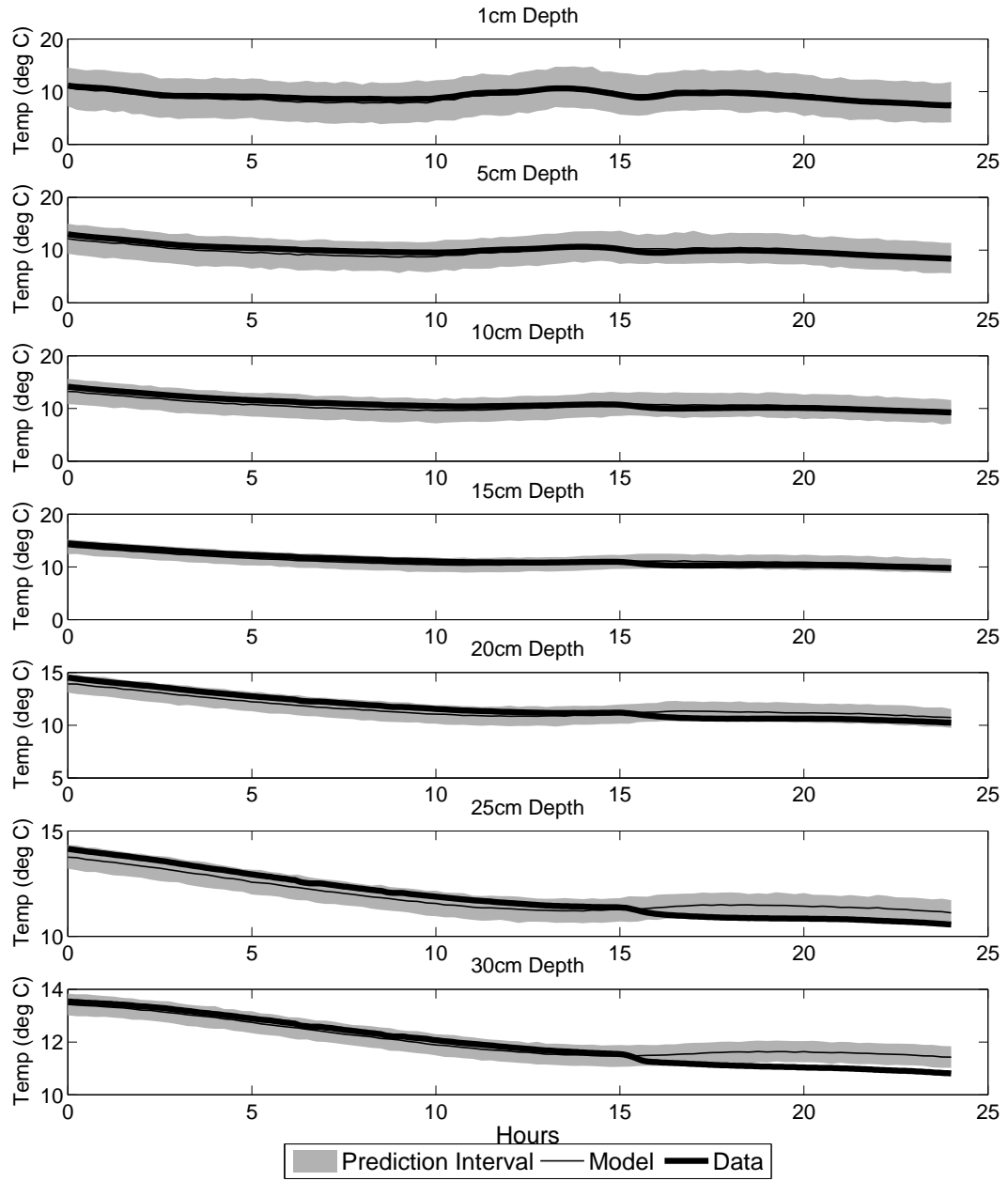


Figure 6.27: 95% Prediction intervals for Dataset H. All depths shown.

## Pairwise scatterplots

In Figure 6.28 we have the pairwise scatterplot of the specific heat of the deep soil and the albedo of the gravel layer. The correlation between these parameters is 0.636.

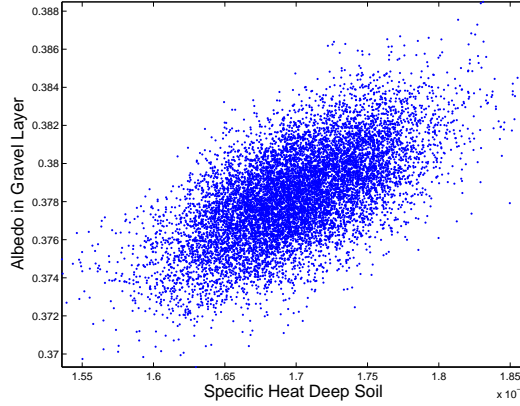


Figure 6.28: Pairwise scatterplots from burned-in portion of chain for Dataset H.

## Results for Parameter Values

In Table 6.14 we show estimates for the specific heat of the deep soil from several sources. First, we include the estimate derived from lab measurements [35]. We include the estimates for the specific heat from using ADH as our model and PEST optimization software for the entire set of dry data (see Chapter 5), PEST and DRAM for the 24-hour dry dataset (see Sections 6.3.1 and 6.3.1 for PEST and DRAM results, respectively), and the new results using PEST and DRAM for the unsaturated dataset. We see good agreement between the estimates for all of the calibration methods as well as a reasonable match to the lab-derived estimate. Thus, we are satisfied with these results.

Table 6.14: Statistics of specific heat estimates for Dataset H. Units are in (W-hr/g-K) [35].

<b>Method</b>	<b>Estimates</b>	<b>Median</b>	<b>Mean</b>	<b>St. Dev.</b>
Estimate derived from lab measurements	1	2.23 E-4	-	-
FEM-PEST (all dry data)	1	1.890 E-4	-	-
FEM-PEST (24-hour dry data)	1	1.693 E-4	-	-
MCMC-ADH (24-hour dry data)	16008	1.692 E-4	1.692 E-4	1.642 E-6
FEM-PEST (24-hour unsaturated data)	1	1.696 E-4	-	-
MCMC - ADH (24-hour unsaturated data)	16008	1.698 E-4	1.699 E-4	4.597 E-6

## Chapter 7

# Conclusions

We consider the soil thermal parameters for three datasets using both optimization and Bayesian inference. For the optimization method, we use PEST optimization software with the ADH model for heat transport. These simulations give us scalar values for parameters for each dataset. For the Bayesian approach, we use the DRAM simulation and both an analytic model derived from the heat equation in 1-D and ADH. The DRAM simulations give us a sample from the posterior density of parameters given each dataset.

For both approaches, we find values for the parameters that are consistent with values estimated for the same dataset using other methods. Moreover, the output of DRAM allows us to quantify uncertainty by constructing prediction intervals for future observations.

One of the benefits of the least squares optimization approach is the convergence theory. The benefits of the Bayesian approach are that it allows us to estimate the densities of parameters, predict future observations and incorporate prior knowledge about the problem into our analysis. The limitations of the Bayesian approach are the proper use of DRAM in determining convergence of the simulation and incorporating prior knowledge, and the computational burden to obtain a sufficiently large sample from the posterior density.

ADH models many physical processes but increases the computational expense of the simulation. The analytic model is simple but fails to capture physical processes such as variable

saturation and some effects of heat flux at the surface.

## REFERENCES

- [1] *Duplex Insulated Copper-Constantan Duplex ANSI Type T*. OMEGA Engineering.
- [2] *HFP01 Heat Flux Plate/Heat Flux Sensor*. Hukseflux Thermal Sensors.
- [3] *Precision Infrared Radiometer, Model PIR*. The Eppley Laboratory, Inc.
- [4] *Vaisala BAROCAP Barometer PTB110 Series User's Guide*. Vaisala.
- [5] *Model HMP45C Temperature and Relative Humidity Probe*, 2002. Campbell Scientific.
- [6] *TE525 Tipping bucket rain gage instruction manual*, 2010. Campbell Scientific, Inc.
- [7] *WindSonic Ultrasonic Wind Sensor*, 2010. Campbell Scientific.
- [8] *CS616 and CS625 water content reflectometers*, 2012. Campbell Scientific.
- [9] I. A. ABBUD, G. O. G. LÖF, AND D. C. HITTLE, *Simulation of solar air heating at constant temperature*, Solar Energy, 54 (1995), pp. 75–83.
- [10] R. G. ALLEN, L. S. PEREIRA, D. RAES, AND M. SMITH, *Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56*, FAO, (1998).
- [11] J. B. J. BARON FOURIER, *The Analytical Theory of Heat*, Cambridge University Press, London, 1878.
- [12] A. G. BARR, K. M. KING, T. J. GILLESPIE, G. D. HARTOG, AND H. H. NEUMANN, *A comparison of Bowen ratio and eddy correlation sensible and latent heat flux measurements above deciduous forest*, Boundary-Layer Meteorology, 71 (1994), pp. 21–41.
- [13] R. G. BARRY AND R. J. CHORLEY, *Atmosphere, Weather and Climate, 8th ed.*, Routledge, London, 2003.

- [14] R. BERGER AND S. HOWINGTON, *Discrete fluxes and mass balance in finite elements*, ASCE J. of Hydraulic Engrg., 128 (2002), pp. 87–92.
- [15] R. BERGER, J. TATE, G. BROWN, AND G. SAVANT, *Adaptive Hydraulics*, 2011.
- [16] J. M. BERNARDO AND A. F. M. SMITH, *Bayesian Theory*, John Wiley & Sons, LTD, Chichester, 2004.
- [17] J. R. BILSKIE, *Dual probe methods for determining soil thermal properties: numerical and laboratory study*, PhD thesis, Iowa State University, 1994.
- [18] R. B. BIRD, W. E. STEWART, AND E. N. LIGHTFOOT, *Transport Phenomena*, John Wiley & Sons, New York, 1960.
- [19] S. A. BOWERS AND R. J. HANKS, *Specific heat capacity of soils and minerals as determined with a radiation calorimeter*, Soil Science, 94 (1962), pp. 392–396.
- [20] K. L. BRISTOW, *Measurement of thermal properties and water content of unsaturated sandy soil using dual-probe heat-pulse probes*, Agricultural and Forest Meteorology, 89 (1998), pp. 75–84.
- [21] K. L. BRISTOW, J. R. BILSKIE, G. J. KLUITENBERG, AND R. HORTON, *Comparison of techniques for extracting soil thermal properties from dual-probe heat-pulse data*, Soil Science, 160 (1995), pp. 1–7.
- [22] K. L. BRISTOW, G. J. KLUITENBERG, AND R. HORTON, *Measurement of Soil Thermal Properties with a Dual-Probe Heat-Pulse Technique*, Soil Science Society of America Journal, 58 (1994), p. 1288.
- [23] K. L. BRISTOW, R. WHITE, AND G. KLUITENBERG, *Comparison of single and dual probes for measuring soil thermal properties with transient heating*, Australian Journal of Soil Research, 32 (1994), pp. 447–464.

- [24] R. H. BROOKS AND A. T. COREY, *Hydraulic Properties of Porous Media*, in Hydrology Papers, A. T. Corey, R. E. Dils, and V. M. Yevdjovich, eds., no. March, Colorado State University, Fort Collins, 1964.
- [25] M. BURZO, P. RAAD, P. KOMAROV, C. WICAKSONO, AND T. CHOI, *Measurement of thermal conductivity of nanofluids and thermal interface materials using the laser-based transient thermorefectance method*, in 29th IEEE SEMI-THERM Symposium, San Jose, CA, 2013, IEEE, pp. 194–199.
- [26] G. S. CAMPBELL, K. L. BRISTOW, J. D. J. JUNGBAUER, AND R. D. HUNGERFORD, *Soil temperature and water content beneath a surface fire*, Soil Science, 159 (1995).
- [27] G. S. CAMPBELL, C. CALISSENDORFF, AND J. H. WILLIAMS, *Probe for Measuring Soil Specific Heat Using A Heat-Pulse Method*, Soil Science Society of America Journal, 55 (1991), pp. 291–293.
- [28] B. P. CARLIN AND T. A. LOUIS, *Bayes and Empirical Bayes Methods for Data Analysis*, Chapman & Hall, London, 1996.
- [29] J. E. CARSON, *Analysis of soil and air temperatures by Fourier techniques*, J. of Geophys. Res., 68 (1963), pp. 2217–2232.
- [30] G. CASELLA AND R. L. BERGER, *Statistical Inference, 2nd ed.*, Thomson Learning, Pacific Grove, CA, 2002.
- [31] R. CHRISTENSEN, W. JOHNSON, A. BRANSCUM, AND T. E. HANSON, *Bayesian Ideas and Data Analysis*, Taylor & Francis Group, Boca Raton, 2011.
- [32] A. M. DE SILANS, B. A. MONTENY, AND J. P. LHOMME, *Apparent soil thermal diffusivity, a case study: HAPEX-Sahel experiment*, Agricultural and Forest Meteorology, 81 (1996), pp. 201–216.



- [33] D. A. DE VRIES, *Thermal properties of soils*, in Physics of Plant Environment, 2nd Ed., W. R. Van Wijk, ed., North-Holland Publishing Co., Amsterdam, 1966, pp. 210–235.
- [34] J. DOHERTY, *PEST: Model-Independent Parameter Estimation User Manual: 5th Edition*, Watermark Numerical Computing, Corinda, Queensland, Australia, 2004. Available online from <http://www.pesthomepage.org>.
- [35] O. J. ESLINGER, A. M. HINES, C. WINTON, S. E. HOWINGTON, R. GOODSON, R. KALA, J. R. FAIRLEY, S. PRICE, AND K. ELDER, *GEOTACS : Countermining computational testbed results for an undifferentiated highlands climatic zone, part 1, ERDC Technical report 12-13*, (2012).
- [36] O. J. ESLINGER, C. WINTON, J. R. BALLARD JR., S. E. HOWINGTON, A. FREGOSI, K. WARD, AND C. T. KELLEY, *Estimating sampling distributions of apparent thermal diffusivity for partially saturated soils*, In revision, (2015).
- [37] J. R. FANCHI, *Rock Properties*, in Principles of Applied Reservoir Simulation, Gulf Professional Publishing, Amsterdam, 3rd ed., 2005, pp. 255–281.
- [38] J. M. FLEGAL, *Monte Carlo Standard Errors for Markov Chain Monte Carlo*, PhD thesis, University of Minnesota, 2008.
- [39] J. M. FLEGAL AND J. HUGHES, *mcmcse: Monte Carlo Standard Errors for MCMC*, 2012.
- [40] J. M. FLEGAL AND G. L. JONES, *Batch means and spectral variance estimators in Markov chain Monte Carlo*, The Annals of Statistics, 38 (2010), pp. 1034–1070.
- [41] D. FRITTON, *Evaluation of experimental procedures by error analysis*, Journal of Agronomic Education, 3 (1974), pp. 43–48.
- [42] D. GAMERMAN AND H. F. LOPES, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2nd ed.*, Chapman and Hall/CRC, Boca Raton, 2006.

- [43] Z. GAO, L. WANG, AND R. HORTON, *Comparison of six algorithms to determine the soil thermal diffusivity at a site in the Loess Plateau of China*, Hydrology and Earth System Sciences Discussions, 6 (2009), pp. 2247–2274.
- [44] C. J. GEYER, *Practical Markov Chain Monte Carlo*, Institute of Mathematical Statistics, 7 (1992), pp. 473–483.
- [45] M. S. GOCKENBACH, *Understanding and Implementing the Finite Element Method*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2006.
- [46] M. GU AND S. C. EISENSTAT, *Efficient algorithms for computing a strong rank-revealing QR factorization*, SIAM J. Sci. Comput., 17 (1996), pp. 848–869.
- [47] H. HAARIO, M. LAINE, A. MIRA, AND E. SAKSMAN, *DRAM: Efficient adaptive MCMC*, Statistics and Computing, 16 (2006), pp. 339–354.
- [48] A. HANSLMEIER, *Water on Earth, Properties of Water*, in Water in the Universe, vol. 368 of Astrophysics and Space Science Library, Springer Netherlands, Dordrecht, 2011, pp. 1–24.
- [49] W. K. HASTINGS, *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, 57 (1970), pp. 97–109.
- [50] D. HILLEL, *Soil Temperature and Heat Flow*, in Introduction to Soil Physics, Academic Press, New York, 1982, ch. 9, pp. 155–175.
- [51] ———, *Introduction to Environmental Soil Physics*, Elsevier, San Diego, 2004.
- [52] ———, *Water Flow in Saturated Soil*, in Introduction to Environmental Soil Physics, Elsevier Academic Press, Amsterdam, 2004, ch. 7, pp. 127–148.
- [53] R. HORTON, P. J. WIERENGA, AND D. R. NIELSEN, *Evaluation of Methods for Determining the Apparent Thermal Diffusivity of Soil Near the Surface*, Soil Science Society of America Journal, 47 (1983), pp. 25–32.

- [54] S. E. HOWINGTON, J. F. PETERS, J. R. BALLARD JR., L. N. LYNCH, J. P. HALLBERG, AND C. E. KEES, *A suite of models for producing synthetic, small-scale thermal imagery of vegetated soil surfaces*, in Proceedings of the XVIth International Conference on Computational Methods in Water Resources, 2006, pp. 1–8.
- [55] I. C. F. IPSEN, C. T. KELLEY, AND S. R. POPE, *Rank-Deficient Nonlinear Least Squares Problems and Subset Selection*, SIAM Journal on Numerical Analysis, 49 (2011), pp. 1244–1266.
- [56] M. Z. JACOBSON, *Fundamentals of Atmospheric Modeling*, Cambridge University Press, Cambridge, 1999.
- [57] O. JOHANSEN, *Thermal Conductivity of Soils*, PhD thesis, Norwegian University of Science and Technology, 1975.
- [58] R. JORDAN, *A One-Dimensional Temperature Model for a Snow Cover: Technical Documentation for SNTHERM.89*, Unbc.Ca, (1991), p. 49.
- [59] K. R. KAVANAGH, *Nonsmooth Nonlinearities in Applications From Hydrology*, PhD thesis, North Carolina State University, 2003.
- [60] B. D. KAY AND J. B. GOIT, *Temperature-dependent Specific Heats of Dry Soil Materials*, Canadian Geotechnical Journal, 12 (1975), pp. 209–212.
- [61] C. T. KELLEY, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.
- [62] M. S. KERSTEN, *Thermal properties of soils*, in Engineering Experiment Station Bulletins, no. 28, Minneapolis, 1949, University of Minnesota.
- [63] G. J. KLUITENBERG, K. L. BRISTOW, AND B. S. DAS, *Error analysis of heat pulse method for measuring soil heat capacity, diffusivity, and conductivity*, Soil Science Society of America Journal, 59 (1995), pp. 719–726.

- [64] G. J. KLUITENBERG, J. M. HAM, AND K. L. BRISTOW, *Error Analysis of the Heat Pulse Method for Measuring Soil Volumetric Heat Capacity*, Soil Science Society of America Journal, 57 (1993), p. 1444.
- [65] M. LAINE, *MCMC Toolbox for Matlab*. <http://helios.fmi.fi/~lainema/mcmc>, 2013.
- [66] S. LU, T. REN, Y. GONG, AND R. HORTON, *An Improved Model for Predicting Soil Thermal Conductivity from Water Content at Room Temperature*, Soil Science Society of America Journal, 71 (2007), pp. 8–14.
- [67] G. MARSAGLIA AND W. W. TSANG, *The Ziggurat method for generating random variables*, Journal of Statistical Software, 5 (2000), pp. 1–7.
- [68] M. MASCAGNI AND A. SRINIVASAN, *Parameterizing parallel multiplicative lagged-Fibonacci generators*, Parallel Computing, 30 (2004), pp. 899–916.
- [69] W. J. MASSMAN, *Modeling soil heating and moisture transport under extreme conditions: Forest fires and slash pile burns*, Water Resources Research, 48 (2012), pp. 1–12.
- [70] W. J. MASSMAN, J. M. FRANK, AND N. B. REISCH, *Long-term impacts of prescribed burns on soil thermal conductivity and soil heating at a Colorado Rocky Mountain site: A data/model fusion study*, International Journal of Wildland Fire, 17 (2008), pp. 131–146.
- [71] N. METROPOLIS, A. ROSENBLUTH, AND M. ROSENBLUTH, *Equation of state calculations by fast computing machines*, The Journal of Chemical Physics, (1953).
- [72] A. MIRA, *On Metropolis-Hastings algorithms with delayed rejection*, Metron, LIX (2001), pp. 231–241.
- [73] K. NOBORIO, K. J. MCINNES, AND J. L. HEILMAN, *Measurements of soil water content, heat capacity, and thermal conductivity with a single TDR probe*, Soil Science, 161 (1996), pp. 22–28.

- [74] T. E. OCHSNER, R. HORTON, AND T. REN, *A New Perspective on Soil Thermal Properties*, Soil Science Society of America Journal, 65 (2001), p. 1641.
- [75] T. R. OKE, *Boundary Layer Climates*, Methuen & Co Ltd, London, 1978.
- [76] D. C. PEARCE AND L. W. GOLD, *Observations of ground temperature and heat flow at Ottawa, Canada*, Journal of Geophysical Research, 64 (1959), pp. 1293–1298.
- [77] M. C. PEEL, B. L. FINLAYSON, AND T. A. MCMAHON, *Updated world map of the Köppen-Geiger climate classification*, Hydrology and Earth System Sciences Discussions, 11 (2007), pp. 1633–1644.
- [78] C. PETERS-LIDARD, E. BLACKBURN, X. LIANG, AND E. WOOD, *The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures*, Journal of the Atmospheric Sciences, 55 (1998), pp. 1209–1224.
- [79] J. S. PETTWAY, J. H. SCHMIDT, AND A. K. STAGG, *Adaptive meshing in a mixed regime hydrologic simulation model*, Computational Geosciences, 14 (2010), pp. 665–674.
- [80] R. A. PIELKE, *Mesoscale Meteorological Modeling, 2nd ed.*, Academic Press, San Diego, 2002.
- [81] S. POPE, L. ELLWEIN, C. ZAPATA, V. NOVAK, C. KELLEY, AND M. OLUFSEN, *Estimation and identification of parameters in a lumped cerebrovascular model*, Mathematical Biosciences and Engineering, 6 (2008), pp. 93–115.
- [82] R CORE TEAM, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014.
- [83] T. REN, K. NOBORIO, AND R. HORTON, *Measuring Soil Water Content, Electrical Conductivity, and Thermal Properties with a Thermo-Time Domain Reflectometry Probe*, Soil Science Society of America Journal, 63 (1999), p. 450.

- [84] G. SAVANT, C. BERGER, T. O. MCALPIN, AND J. N. TATE, *Intelligent Adaptive Time-Step Control for Modeling Rapidly-Evolving Hydrodynamic Flows in Adaptive Hydraulics (ADH)*, Tech. Rep. ERDC TN-SWWRP-10-6, U.S. Army Engineer Research and Development Center, Vicksburg, MS, 2010.
- [85] D. P. SHOEMAKER AND C. W. GARLAND, *Experiments in physical chemistry*, McGraw-Hill, New York, 1962.
- [86] S. R. SINGH AND B. K. SINHA, *Soil Thermal Diffusivity Determination from Overspecification of Boundary Data*, Soil Science Society of America Journal, 41 (1977), pp. 831–834.
- [87] R. C. SMITH, *Uncertainty Quantification: Theory, Implementation and Applications*, SIAM, Philadelphia, PA, 2013.
- [88] A. SOLONEN, *Monte Carlo methods in parameter estimation of nonlinear models*, PhD thesis, Lappeenranta University of Technology, 2006.
- [89] J. M. TARARA AND J. M. HAM, *Measuring soil water content in the laboratory and field with dual-probe heat-capacity sensors*, Agronomy Journal, 89 (1997), pp. 535–542.
- [90] S. TAYLOR AND R. JACKSON, *Heat capacity and specific heat*, in Methods of Soil Analysis, Part 1., Agronomy Monograph 9.1, C. A. Black, ed., American Society of Agronomy, Inc., Madison, Wis., 1965, pp. 345–348.
- [91] L. TIERNEY, *Markov chains for exploring posterior distributions*, The Annals of Statistics, 22 (1994), pp. 1701–1762.
- [92] M. D. TOCCI, C. T. KELLEY, AND C. T. MILLER, *Accurate and economical solution of the pressure-head form of Richards equation by the method of lines*, Advances in Water Resources, 20 (1997), pp. 1–14.
- [93] B. USOWICZ, *Statistical-Physical Model of Thermal Conductivity in Soil*, Polish Journal of Soil Science, XXV (1992), pp. 25–34.

- [94] M. T. VAN GENUCHTEN, *A closed-form equation for predicting the hydraulic conductivity of unsaturated soils*, Soil Science Society of America Journal, 44 (1980), pp. 892–898.
- [95] W. R. VAN WIJK, *General temperature variations in a homogeneous soil*, in Physics of Plant Environment, 2nd Ed., W. R. Van Wijk, ed., North-Holland Publishing Co., Amsterdam, 1966, pp. 144–170.
- [96] ———, *Physics of Plant Environment*, 2nd ed., North-Holland Publishing Co., Amsterdam, 1966.
- [97] W. R. VAN WIJK AND D. A. DE VRIES, *Periodic temperature variations in a homogeneous soil*, in Physics of Plant Environment, 2nd Ed., W. R. Van Wijk, ed., North-Holland Publishing Co., Amsterdam, 1966, pp. 102–143.
- [98] ———, *The atmosphere and the soil*, in Physics of Plant Environment, 2nd Ed., W. R. Van Wijk, ed., North-Holland Publishing Co., Amsterdam, 1966, pp. 17–61.
- [99] S. M. WELCH, G. J. KLUITENBERG, AND K. L. BRISTOW, *Rapid numerical estimation of soil thermal properties for a broad class of heat-pulse emitter geometries*, Measurement Science and Technology, 7 (1999), pp. 932–938.
- [100] P. J. WIERENGA, D. R. NIELSEN, AND R. M. HAGAN, *Thermal Properties of a Soil Based Upon Field and Laboratory Measurements*, Soil Science Society of America Journal, 33 (1969), pp. 354–360.
- [101] D. J. WILKINSON, *Parallel Bayesian Computation*, in Handbook of Parallel Computing and Statistics, E. J. Kontoghiorghes, ed., 2005.
- [102] C. W. WINTON, *Parameter Estimation in Groundwater Models Using Proper Orthogonal Decomposition*, PhD thesis, North Carolina State University, 2012.
- [103] M. R. YADAV AND G. S. SAXENA, *Effect of Compaction and Moisture Content on Specific Heat and Thermal Capacity of Soils*, J. Indian Soc. Soil Sci., 21 (1973), pp. 129–132.

## APPENDICES



# Appendix A

## Data

In this appendix we include pictures of a meteorological (met) station, pictures of instruments on the met station, and details about the instruments.

### A.1 Instruments

Below is a picture of an entire met station located in Vicksburg, Mississippi.



Figure A.1: Met Station in Vicksburg

### A.1.1 Subsurface Instruments

Three instruments are buried near the met station to measure quantities below the surface: the temperature probe, flux plate, and soil moisture probe. The soil moisture probe is a Campbell

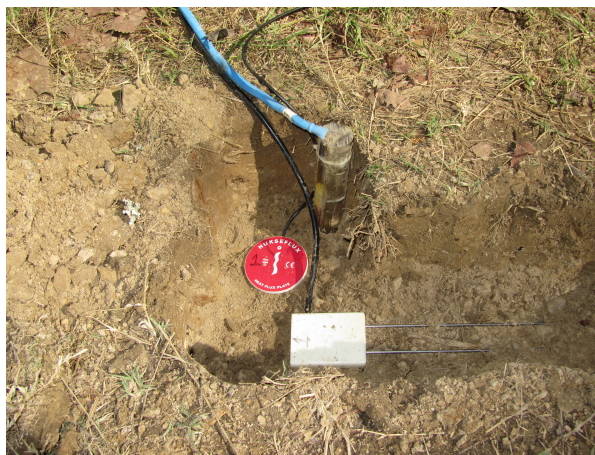


Figure A.2: Layout of Instruments

Scientific, Model CS-616. To measure the water content of the soil, a signal is sent down the two 30 cm stainless rods [8]. The dielectric properties of the soil determine the deviation of the return signal, which is measured every second and averaged over the five minute output interval [35]. The units of output are per cent of volumetric water content and the accuracy of this device is  $\pm 2\%$  [35]. The temperature probe gives the temperature of the soil at locations 1, 5, 10, 15, 20, 25, and 30 cm below the surface. We use a Type T Thermocouple Junction manufactured by Omega Engineering, Inc [1]. The output is the average temperature in degrees Celsius of one second samples over the output interval of five minutes [35]. The accuracy of this device is less than  $\pm 0.1$  degrees Celsius [35]. There are two temperature probes on each met station. The soil thermal flux is measured by the Hukseflux Model HFP01 heat flux plate [2]. This sensor is buried at a depth of 15 cm and gives the average heat flux for samples every second over an interval of five minutes. The output is in units of Watts per square meter [35].



Figure A.3: Temperature Probe



Figure A.4: Flux Plate

Because flux carries a positive or negative sign based on the direction of the movement, the plate is labelled to indicate which side must be placed up. There are two flux plates on each met station.

### A.1.2 Radiometer

The radiometer reports the radiated surface temperature of the soil. It is located above the surface and gives the temperature in degrees Celsius of the soil at a desired location [35]. The radiometer is pointed at this location. It is not reasonable to assume that the radiometer is

pointed at the location of the buried temperature probe, but in some cases this may be true. In Figure A.5 we have a picture where the radiometer is indeed pointed at the temperature probe. The radiometer and the surface location of the temperature probe are circled. There are two radiometers on each met station, giving the target temperature, or temperature at a desired location on the surface of the ground.



Figure A.5: Radiometer

### A.1.3 Air Temperature and Relative Humidity

The Type 1 Vaisala HMP45C temperature and relative humidity probe records the temperature in degrees Celsius and per cent relative humidity, respectively, at 0.5 M and 2.3 M above the surface. The temperature accuracy is  $\pm 0.2$  degrees C from -40 to 60 degrees [35]. The relative humidity accuracy is  $\pm 2\%$  in the range 0 to 90 % RH and  $\pm 3\%$  in the range 90 to 100% RH [5]. As before, the data is collected every second and the average over five minutes is recorded.

#### **A.1.4 Barometric Pressure**

The barometric pressure is measured by the Vaisala Model PTB110 1BoAA Pressure Transmitter at about 0.25 M above the surface [4]. With units of milliBars, its accuracy is  $\pm 6$  milliBars in the range -40 C to 60 C [35]. Again, the output is the average of measurements every second for a period of five minutes.

#### **A.1.5 Radiation**

We have data for both shortwave and longwave radiation. The Eppley Laboratory Model Precision Spectral Pyranometer [3] measures the shortwave radiation while the Model Precision Infrared Radiometer measures the longwave, or infrared, radiation [35]. These devices are about 2 M above the surface and give output in units of Watts per square meter [35].

#### **A.1.6 Wind Speed and Direction**

The Windsonic 2D Sonic Anemometers measure wind speed and wind direction at 0.5 M and 2.3 M above the surface. The wind speed is accurate to  $\pm 2\%$  of the reading, in a range of 0 to 60 meters per second [7]. The output is the mean horizontal wind speed over five minutes in units of meters per second [35]. The wind direction is accurate to  $\pm 3$  Degrees with a resolution of 1 Degree [7]. The wind direction output is the unit vector mean wind direction over 5 minutes, in units of Degrees from North [35].

#### **A.1.7 Precipitation**

We use the Texas Electronics, Model TE525MM rain gage to measure precipitation. It is calibrated for millimeter output and accurate for rainfall rates [6]:

Up to 10 mm/hr	$\pm 1 \%$
10 to 20 mm/hr	+0, -3 %
20 to 30 mm/hr	+0, -5 %

The output is the total millimeters of rainfall over a 5 minute interval [35].

#### **A.1.8 Storage of Data**

A Garmin GPS on the met station maintains accurate datalogger time. Power supply for the station comes from a Campbell Scientific, Inc. Model CH100 Charging Regulator, Yuasa NP24-12 battery and BP Solar Panel Model SX10M. Data is stored on a Silicon Systems Compact Flash Card with capacity 1 GB, allowing over 4 months of data collection. This data is retrieved with Campbell Scientific software [35].

## Appendix B

# Adaptive Hydrology

In Chapter 2 we mention using ADH to simulate heat transport in the soil. We now describe the model in more detail and give instructions for implementing ADH.

### B.1 Implementation

ADH requires six input files. A file with extension `.sup` contains the paths and names of the other input files for the boundary conditions, domain, initial conditions, and data. These are called the boundary condition, mesh, hotstart, and met files with extensions `.bc`, `.3dm`, `.hot`, and `.met`, respectively. The boundary condition file also contains information about all other parameters used in the model. The data stored in the met file is used in calculating the boundary condition. To run ADH from the command line, the user must type:

```
\pathname\adh basename
```

where the `.sup` file is `basename.sup`.

#### B.1.1 Boundary Condition File

The boundary condition file contains most of the information a user may want to change in ADH. The output specifications, initial and final times, maximum allowed time step size, and



all values for parameters are contained in this file. Control cards are character strings which specify values in the boundary condition file [15].

For example, the line

TC JUL 220

tells ADH the start Julian day of the simulation is 220. The first two characters in the line, TC, indicate the control card is time control. The letters JUL specify Julian day and the integer 220 is the value of the day. Another example of a specification in the boundary condition file is:

MP TKS 1 1.451

which indicates the saturated thermal conductivity is 1.451. Here, the control card MP stands for material properties and the letters TKS represent saturated thermal conductivity. The first number, 1, indicates that this is the value for material 1 when working in a domain with multiple materials. The second number is the value of the material parameter. Comments in the boundary condition file begin with an exclamation point. Thus, an example of an entire line may be:

MP TKS 1 1.451    ! Saturated Thermal Conductivity (W) / (m K)

The ADH manual as well as [35] document the control cards and values in the boundary condition file [15].

### **B.1.2   Mesh File**

The mesh file contains the locations of the nodes, edges, and faces of the tetrahedral mesh. It must have extension .3dm for a 3-D domain.

### **B.1.3   Hotstart Files**

The initial conditions for the model are stored in the hotstart files. A constant initial condition across the domain is an easy way to begin the simulation. Another option is to specify the



initial condition as the solution at the final time of a previous AdH run. The initial conditions for the heat and groundwater portions of the model are stored in the heat and groundwater hotstart files, respectively.

#### **B.1.4 Met File**

The met file contains data from a meteorological station. The column headers include: date, hour, minute, pressure, temperature, humidity, wind speed, wind direction, visibility, precipitation, shortwave, and longwave radiation. ADH extracts information from the met file to formulate the boundary conditions in the model.

#### **B.1.5 Sup File**

For an example of a sup file, the contents of the file `basename.sup` are below:

```
GEO domain_mesh.3dm
BC boundary_conditions.bc
HT HOT initial_heat.hot
GW HOT initial_groundwater.hot
MET met_data.met
```

## Appendix C

# Diagram of Simulations

### C.1 PEST + ADH

The details for running PEST with any model are extensively described in [34]. For the dry simulations, we use the model command line:

```
(PATH_TO_ADH/adh base_name > Screen_output.txt) >& Error_ouput.err; /usr/bin/python go_through_list.py  
> python_out_tmp.dat; sleep 5;
```

to call ADH and extract the output at the appropriate nodes and times. For the unsaturated simulations, we call ADH and extract the temperature and saturation output using a bash script:

```
#!/bin/bash  
  
(PATH_TO_ADH/adh base_name > Screen_output.txt) >& Error_output.err  
  
/usr/bin/python go_through_list_tmp.py > python_out_tmp.dat  
  
/usr/bin/python go_through_list_sat.py > python_out_sat.dat  
  
cat python_out_tmp.dat python_out_sat.dat > merged_out.dat  
  
sleep 5
```

## C.2 DRAM + ADH

The DRAM code is written in Matlab. It can be downloaded from

<http://helios.fmi.fi/~lainema/mcmc>.

To run simulations from the command line, we use a bash script containing the following two lines:

```
#!/bin/bash  
  
matlab -nodisplay -r "run_DRAM_Desert; quit;" > Screen_output.txt
```

We use, for example, a Matlab script `run_DRAM_Desert`, which contains the instructions and input to DRAM. Additionally, functions `model_prior.m` and `model_ssq.m` are needed to evaluate the prior distribution and model, respectively. We discuss these three files in sections C.2.1 -C.2.3.

### C.2.1 `run_DRAM_Desert`

At the beginning of this script, we include the path to the DRAM package. We then specify inputs to DRAM: initial parameter values, parameter names, number of data points, upper and lower bounds for parameters, function names for `model_prior.m` and `model_ssq.m`, update of error variance, initial proposal covariance matrix, and number of DRAM simulations. We also include details for running in parallel, initialize random number streams, and save our results. We call DRAM with the line:

```
[results,chain,s2chain,sschain] = mcmcrun(model,data,PARAMS,options);
```

### C.2.2 `model_prior.m`

We use noninformative uniform priors. Thus, our prior distributions are just the product of the ranges of the parameters.

### C.2.3 `model_ssq.m`

The function `model_ssq.m` has the following header:

```
function [ssq]=model_ssq_highlands(theta,CARDLIST)
```

It takes as inputs the parameter values, `theta`, and list of parameter names, `CARDLIST`. Its output is the weighted residual sum-of-squares. See Figure C.1 for a diagram of this function.

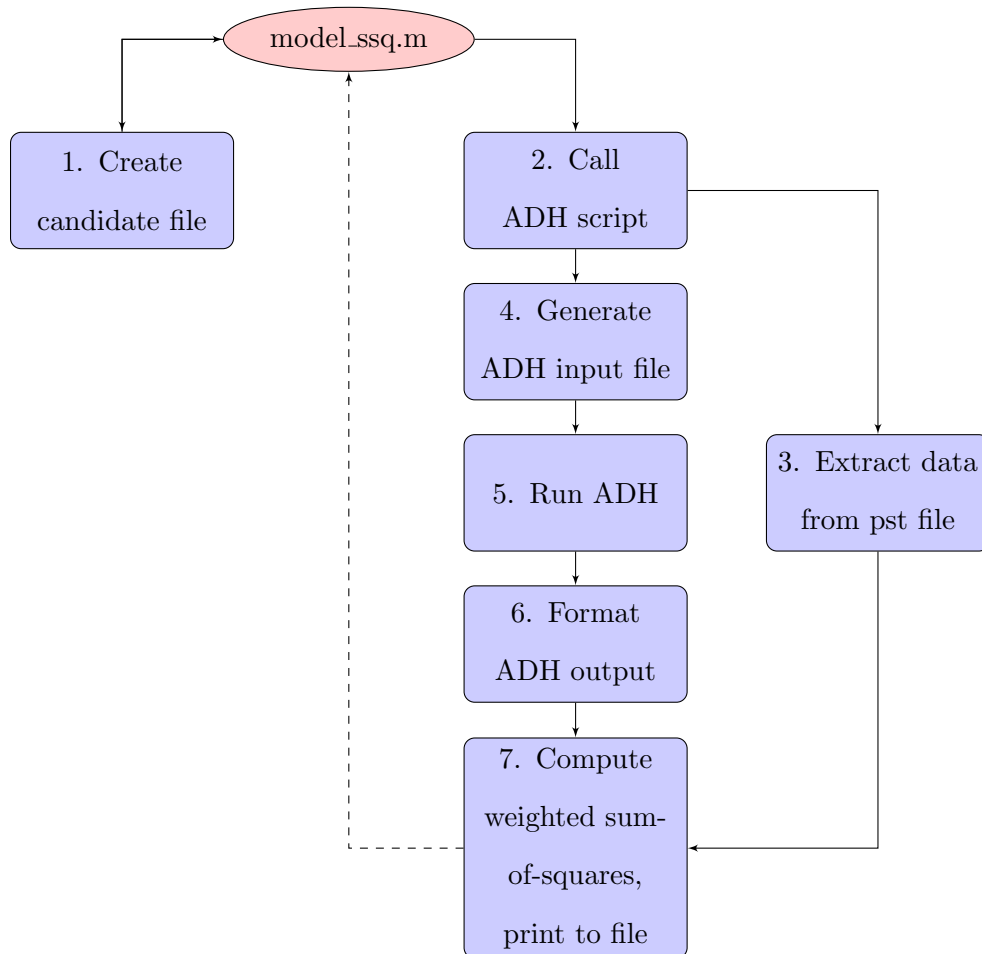


Figure C.1: Diagram of `model_ssqr.m`.

#### 1. Create candidate file

Use the following code from `model_ssqr.m`:

```

cd('./Desert_files')
fileID=fopen('candidate.txt','w');
formatSpec='%-5s %10.6e \r\n';

```

```

for i=1:length(CARDLIST)
    fprintf(fileID,formatSpec,CARDLIST{i},theta(i));
end
fclose(fileID);

```

## 2. Call ADH script

Make a system call in `model_ssq.m` from Matlab to the script `get_ssq.csh`. Additional arguments are the candidate parameter file, `candidate.txt`, the base name for the ADH simulation `Desert.col` (in this example), and the data file `data_Desert.pst` (in this example).

```
system('csh get_ssq.csh candidate.txt Desert.col data_Desert.pst')
```

The file `get_ssq.csh` defines the system arguments:

```

set input_file = $1
set base_name = $2
set pest_file = $3

```

## 3. Extract data from pst file

Use a python script to extract data and weights from the data file. We use the PEST data file as it contains the necessary information.

```
python pst_data_grab.py $pest_file > data_weights.txt
```

## 4. Generate ADH input file

Use a python script to change the parameter values in the ADH input file, the boundary condition file.

```
python input_grab.py $input_file
```

## 5. Run ADH

Call ADH using the command line format:

```
(path_to_adh/adh $base_name > adh_output.out) >& adh_errors.warnings.err
```

## 6. Format ADH output

Use a python script to extract output at the appropriate nodes and times:

```
python go_through_list.py $base_name'_tmp.dat' > adh_output_formatted.txt
```

## 7. Compute weighted sum-of-squares and print to file

First, paste the values from ADH, the data, and the weights in one file:

```
paste -d ' ' data_weights.txt adh_output_formatted.txt > merged_file.txt
```

Then, use a python script to calculate the weighted sum-of-squares residual. Print scalar value to file `Final_sum.txt`:

```
python pst_calculation.py merged_file.txt > Final_sum.txt
```

At the end of the ADH script, `get_ssq.csh`, we remove the files generated in computing the sum-of-squares. The Matlab function `model_ssq.m` uses system calls and Matlab functions to extract the sum-of-squares:

```
system('cp Final_sum.txt ..')
ssq=textread('Final_sum.txt','%f');
cd('../')
```

## C.2.4 Prediction Intervals

To compute prediction intervals, we run ADH with numerous parameter values sampled with DRAM. We generally choose to obtain a sample of 1000 from the posterior values generated by DRAM. The process is similar to the one described above.

First, we call the Matlab script with a bash script at the command line:

```
!/bin/bash matlab -nodisplay -r "get_pred_out; quit;" > Screen_output.txt
```

In the Matlab script, `get_pred_out.m`, we add the path to the DRAM package. We also load the results from a previous DRAM run and specify parameter names. If we ran DRAM in parallel, we take extra steps to ensure we include the results from all chains. Then, we call a variation of the DRAM function for generating prediction intervals with the line:

```
out = mcmcpred_anna_adh(results, chain, s2chain, [], 'get_adh', 1000, PARAMS);
```

Finally, in this script we save the output, `out`.

We edited the DRAM function `mcmcpred.m` in order to accommodate our weighting of the data by depth. We include the following vector of weights

```
weights=[0.08 0.1 0.125 0.21 0.3125 0.5 0.7140]';

w_here=1./weights;

tell ADH which parameters to change by including extra input arguments
```

```
y = feval(modelfun,datai,th,varargin{:});

and include weighting in generating the prediction intervals

y + randn(size(y))*diag(w_here)*diag(sqrt(s2chain(isample(iisample),:)));
```

The function `mcmcpred.anna_adh` calls `get_adh.m` to evaluate ADH. The process is very similar to that in section C.2.3 except that we do not compute the weighted sum-of-squares residual and must be careful that the ADH output is in matrix form. We get the ADH output in matrix form using the Matlab `reshape` function:

```
adh_out_orig= dlmread( 'adh_output.txt','%f');
adh_out=reshape(adh_out_orig,[],7);
```

## C.2.5 Parallel Simulations

In the Matlab script `run_DRAM_Desert`, we make the necessary changes to run our simulations in parallel. First, we start a Matlab parallel pool and determine the number of workers:

```
poolobj = parpool;
if isempty(poolobj)
    poolsize = 0;
else
    poolsize = poolobj.NumWorkers
end
```

Then, we initialize a different stream of random numbers on each worker

```
spmd
s = RandStream.create('mlfg6331_64','NumStreams',numlabs,'StreamIndices',labindex);
RandStream.setGlobalStream(s);
end
```

We call DRAM with an SPMD block and create a temporary directory for each worker:

```
spmd
    currDir = pwd;
    addpath(currDir);
    addpath(strcat(currDir,'/Desert_files'));
    tmpDir = tempname;
    mkdir(tmpDir);
    copyfile('*',tmpDir)
    cd(tmpDir);
    pwd
    %%% run dram
    p0 = mvnrnd(theta, pestcov)
    PARAMS= { {CARDLIST{1},p0(1),pr_low(1),pr_up(1)} };
    model.sigma2 = gamrnd( 0.1, 0.1 );
    [results,chain,s2chain,sschain] = mcmcrun(model,data,PARAMS,options);
    cd(currDir);
    rmdir(tmpDir,'s');
    rmpath(currDir);
end
```

Finally, we save our results from all workers and close the Matlab parallel pool.

```
dsave file_name results chain s2chain sschain options PARAMS data model
delete(poolobj)
```



## Appendix D

# Additional Diagnostics

### D.1 Constant Saturation

#### D.1.1 Least Squares (PEST) with Numerical Model (ADH)

Below are the plots of the ADH with optimal parameter values found using PEST for the constant saturation full datasets.

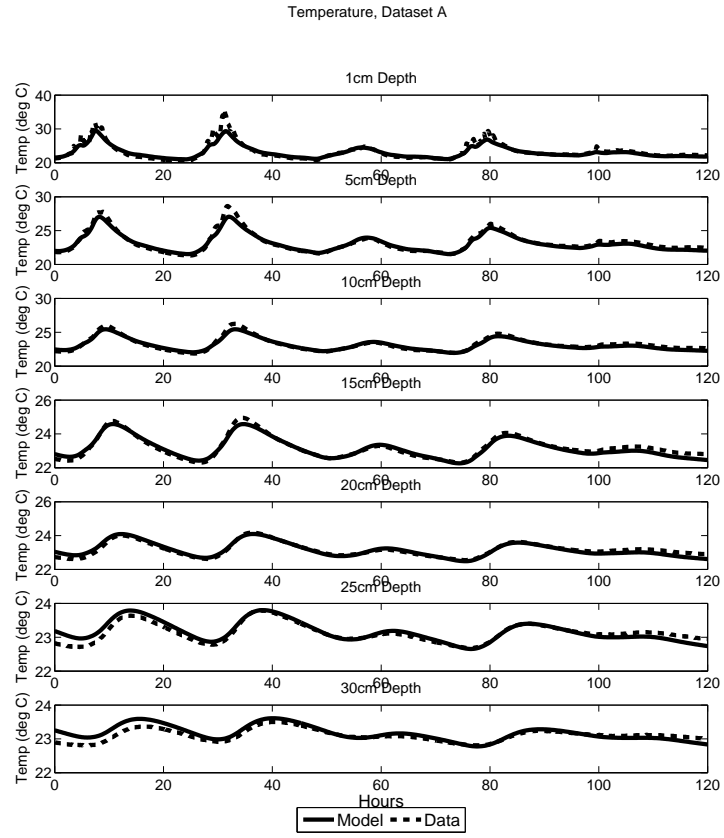


Figure D.1: Dataset A model and data over all depths for the constant saturation case (PEST+ADH).

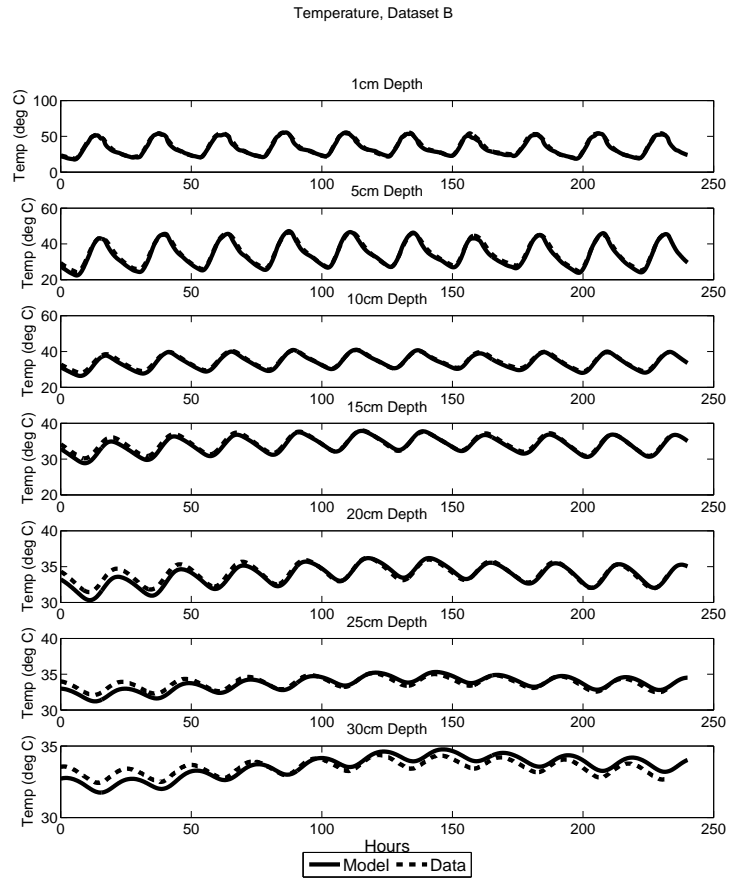


Figure D.2: Dataset B model and data over all depths for the constant saturation case (PEST+ADH).

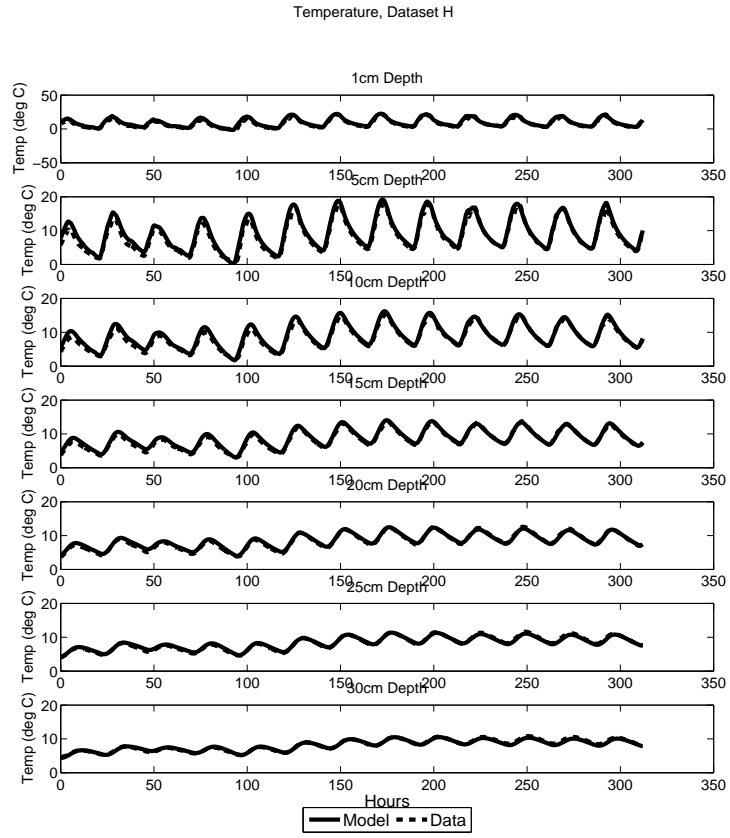


Figure D.3: Dataset H model and data over all depths for the constant saturation case (PEST+ADH).

### D.1.2 DRAM

Below are the additional diagnostics for Datasets B and H using DRAM with the analytic solution to the heat equation as a model.

## Dataset B

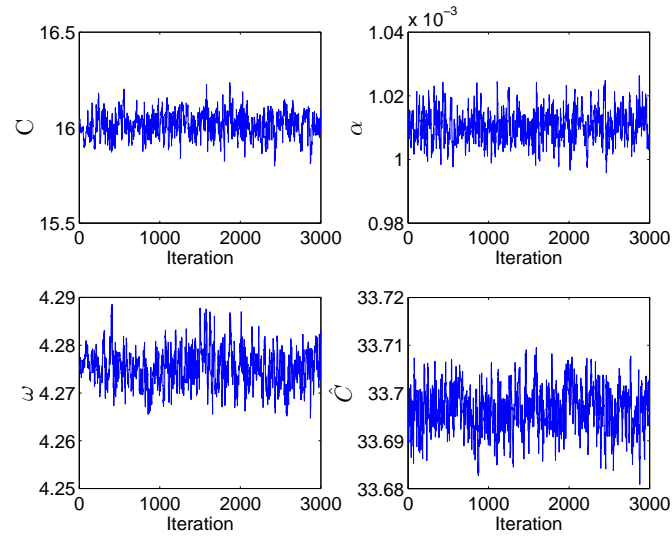


Figure D.4: Chain history from 1 to 3,000 samples, Dataset B.

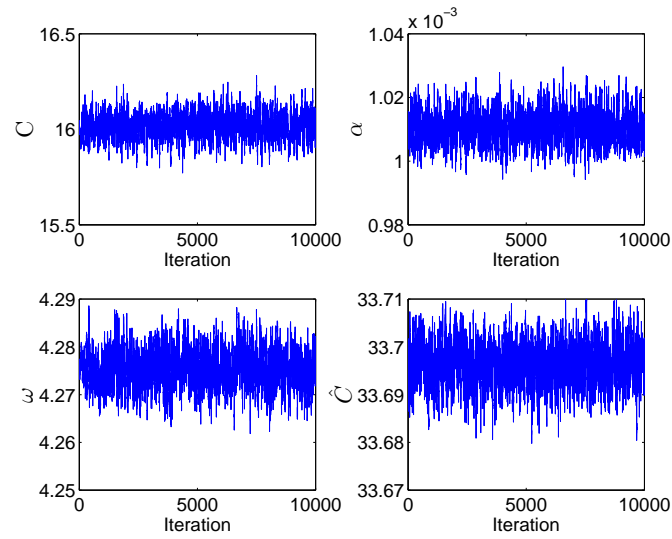


Figure D.5: Chain history from 1 to 10,000 samples, Dataset B.

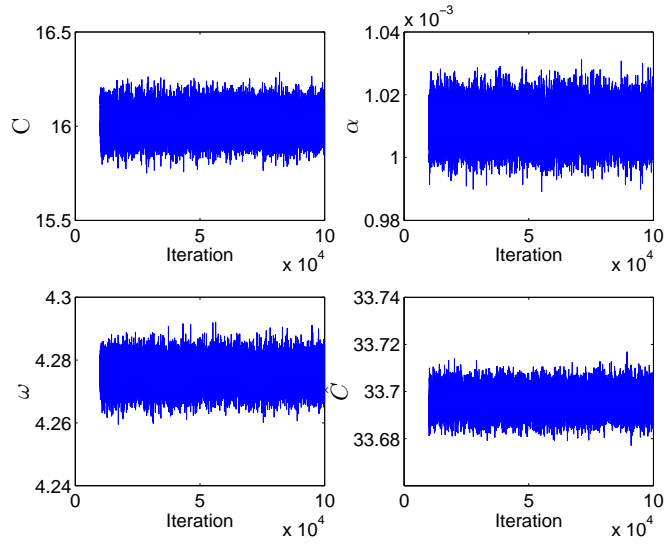


Figure D.6: Chain history from 10,000 to 100,000 samples, Dataset B.

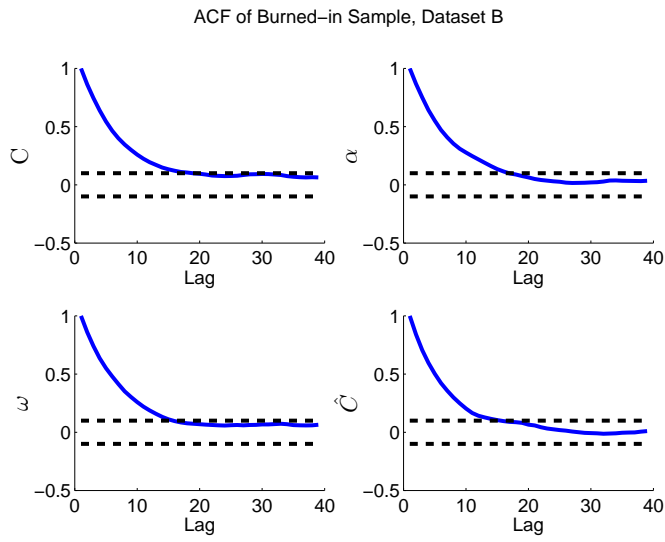


Figure D.7: Autocorrelation for burned-in portion of chain, Dataset B.

Table D.1: Mean, standard deviation and Monte Carlo standard error (MCSE) for  $\alpha$  from Dataset B.

Sample	Mean	St. Dev.	MCSE
1:3000	1.138 E-01	2.693 E-04	1.500 E-05
1:10,000	1.138 E-01	2.832 E-04	8.822 E-06
10,000:100,000	1.138 E-01	2.901 E-04	3.509 E-06
3000:10,000	1.138 E-01	2.890 E-04	1.006 E-05

## Dataset H

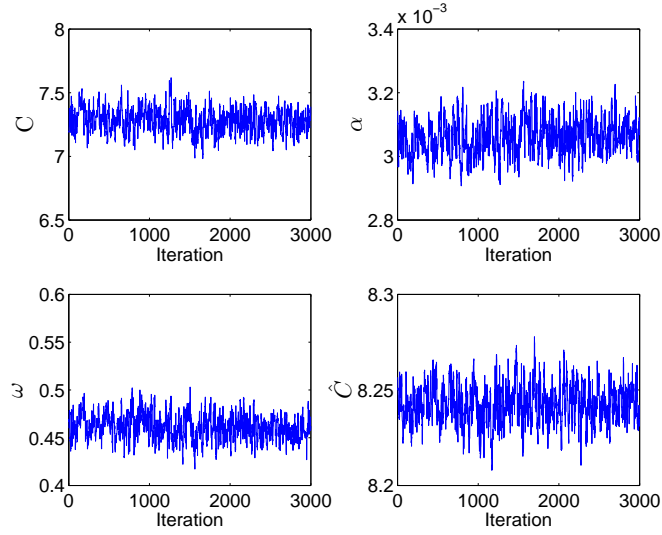


Figure D.8: Chain history from 1 to 3,000 samples, Dataset H.

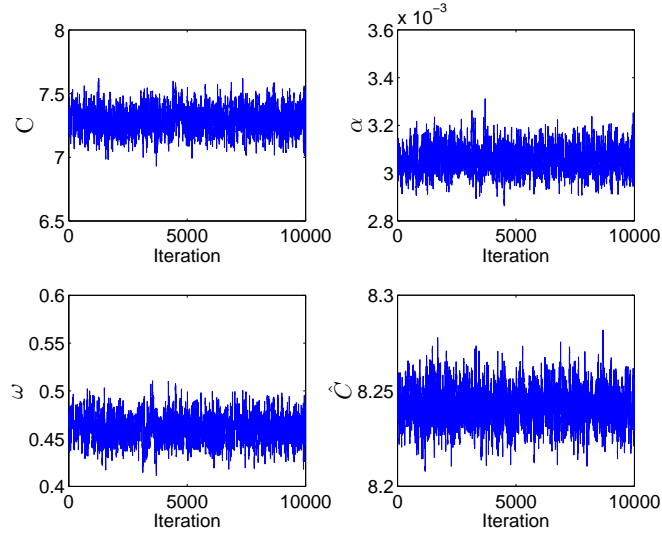


Figure D.9: Chain history from 1 to 10,000 samples, Dataset H.

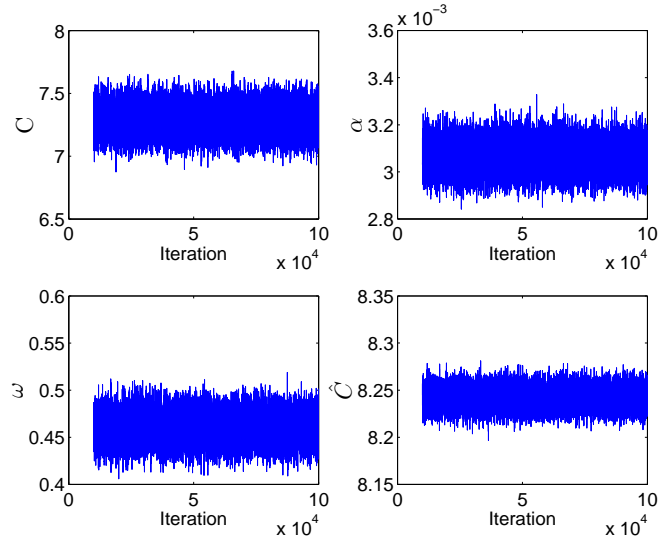


Figure D.10: Chain history from 10,000 to 100,000 samples, Dataset H.



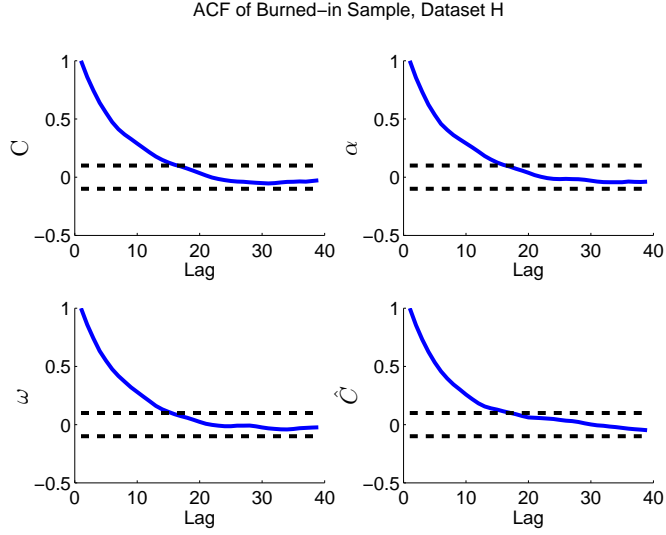


Figure D.11: Autocorrelation for burned-in portion of chain, Dataset H.

Table D.2: Mean, standard deviation and Monte Carlo standard error (MCSE) for  $\alpha$  from Dataset H.

Sample	Mean	St. Dev.	MCSE
1:3000	6.539 E-02	5.915 E-04	4.231 E-05
1:10,000	6.540 E-02	5.999 E-04	2.504 E-05
10,000:100,000	6.540 E-02	5.959 E-04	6.316 E-06
3000:10,000	6.541 E-02	6.033 E-04	2.500 E-05

## D.2 Variable Saturation

The following are supplementary plots for the results in Chapter 6.

### D.2.1 PEST

Below is the plot of the ADH output and data for Dataset H in the variably saturated domain. ADH was run with the optimized parameter values found using PEST.

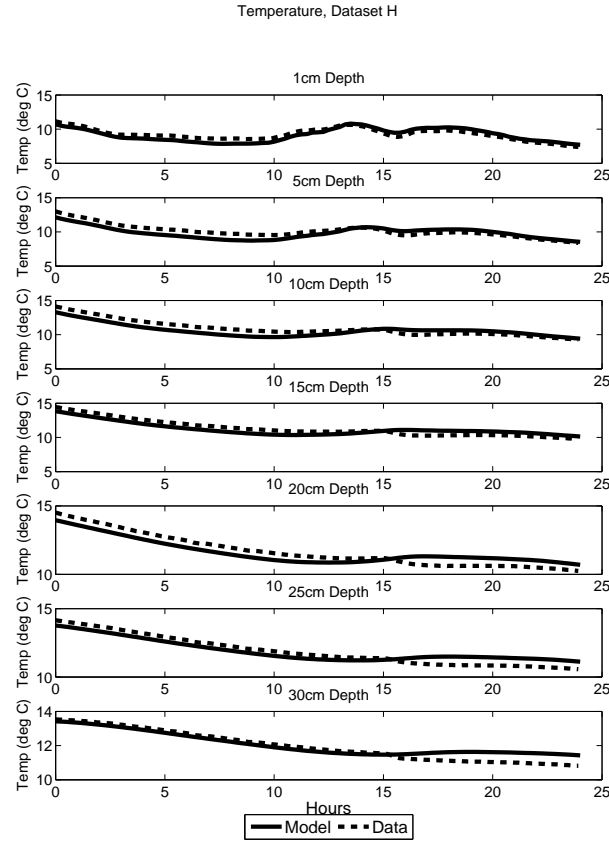


Figure D.12: Dataset H model and temperature data all depths over calibration period, unsaturated case.

### D.2.2 24-hour Constant Saturation Results

Below are plots of ADH evaluated with optimal parameter values found using PEST.

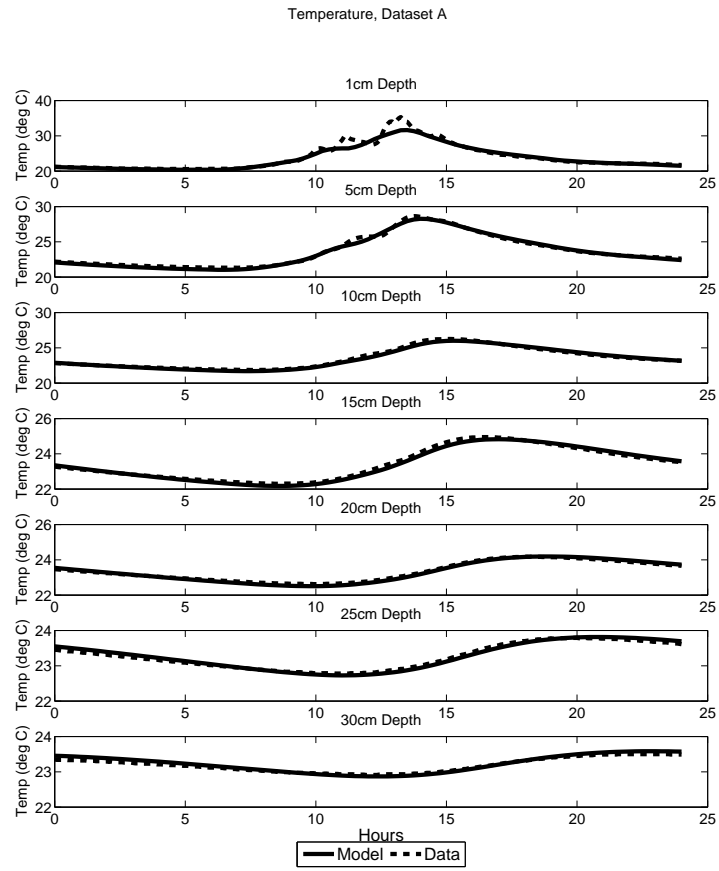


Figure D.13: Dataset A model and temperature data all depths over calibration period, constant 24-hour case.

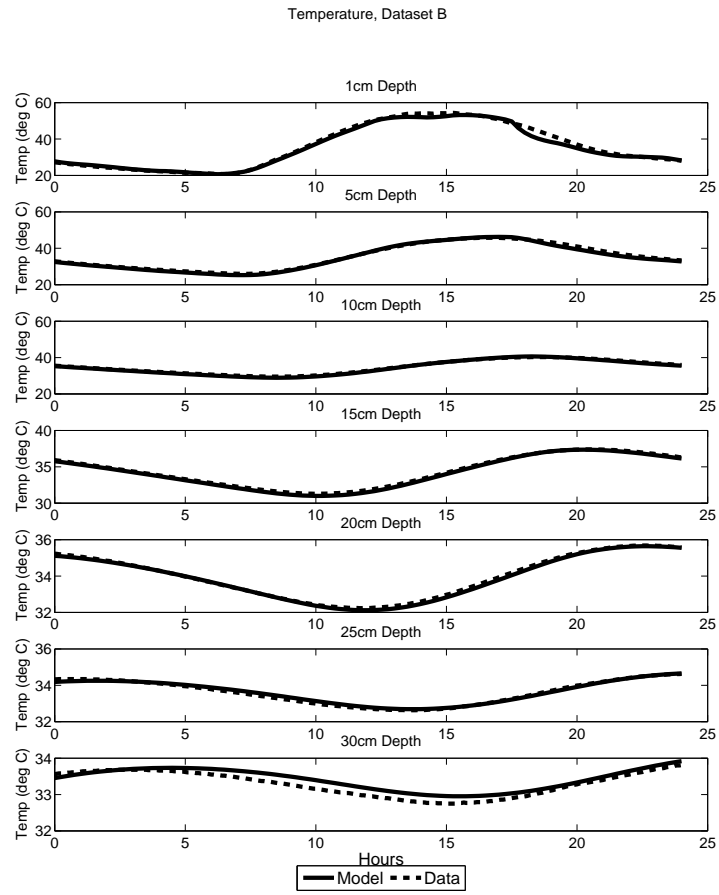


Figure D.14: Dataset B model and temperature data all depths over calibration period, constant 24-hour case.

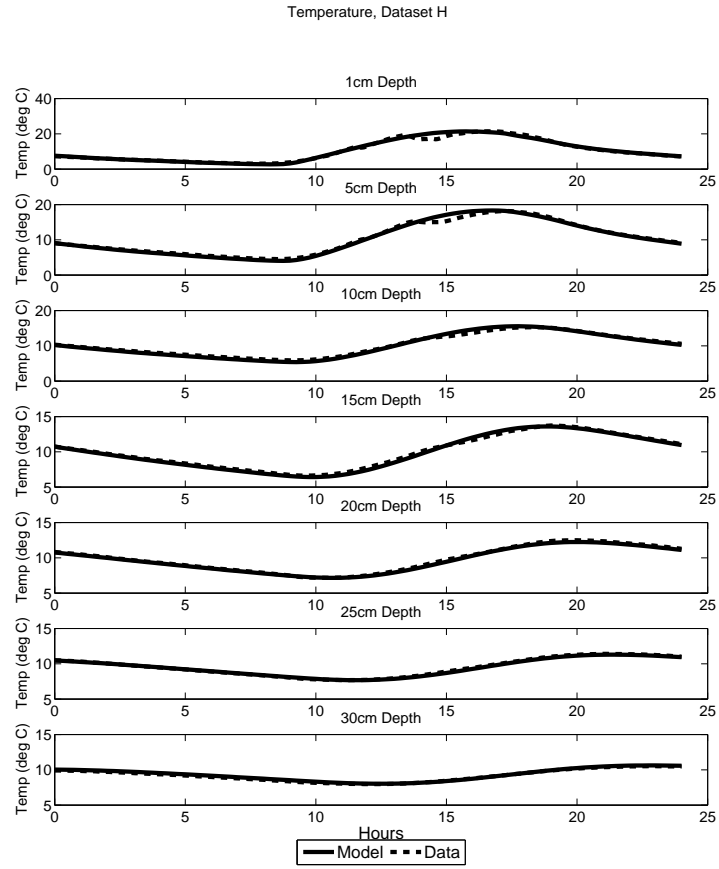


Figure D.15: Dataset H model and temperature data all depths over calibration period, constant 24-hour case.

### D.2.3 DRAM

Below are additional diagnostics for Datasets B and H using DRAM with ADH as a model for the 24-hour constant saturation case.

## Dataset B

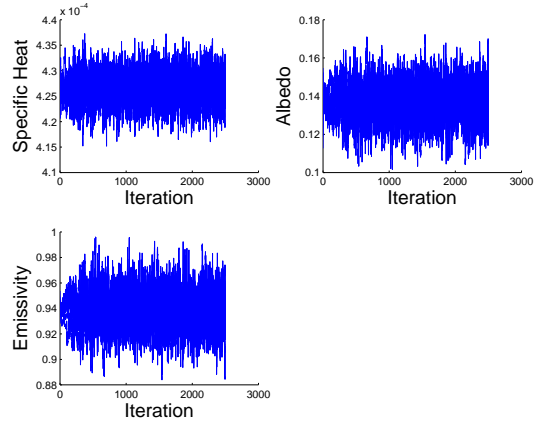


Figure D.16: Dataset B entire simulation for all chains, constant saturation 24-hour case.

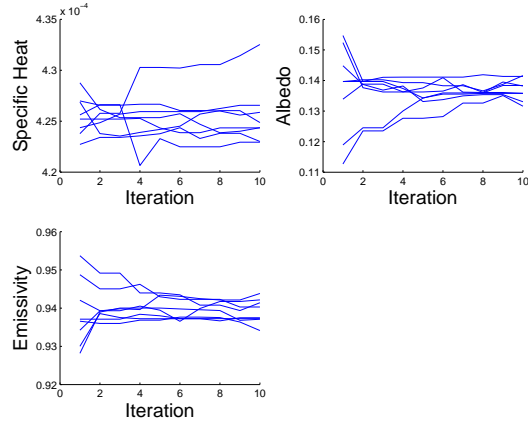


Figure D.17: Dataset B first few iterations across all chains, constant saturation 24-hour case.

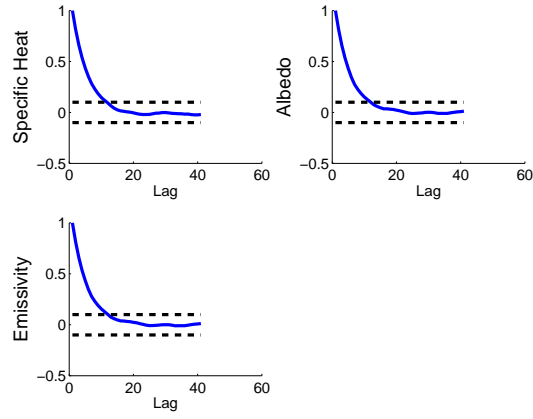


Figure D.18: Dataset B autocorrelation function, constant saturation 24-hour case.

Table D.3: Mean, standard deviation and Monte Carlo standard error (MCSE) for each parameter from Dataset B.

Parameter	Mean	St. Dev.	MCSE
Specific Heat	4.264 E-04	3.028 E-06	8.264 E-08
Emissivity	9.382 E-01	1.497 E-02	4.167 E-04
Albedo	1.371 E-01	9.492 E-03	2.563 E-04

## Dataset H

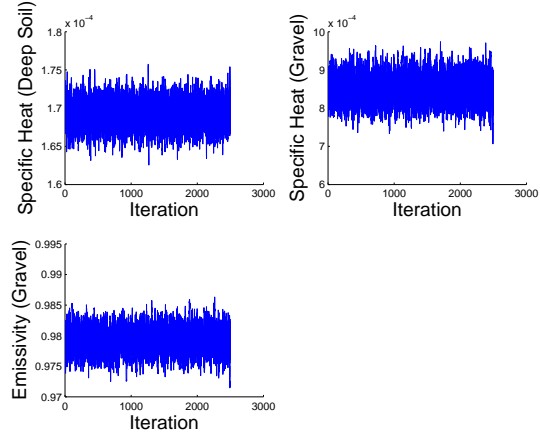


Figure D.19: Dataset H entire simulation for all chains, constant saturation 24-hour case.

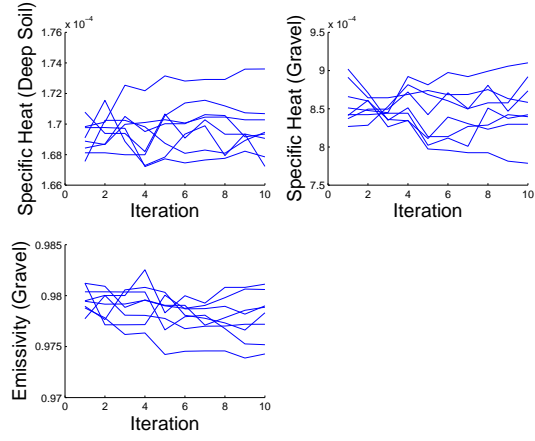


Figure D.20: Dataset H first few iterations across all chains, constant saturation 24-hour case.



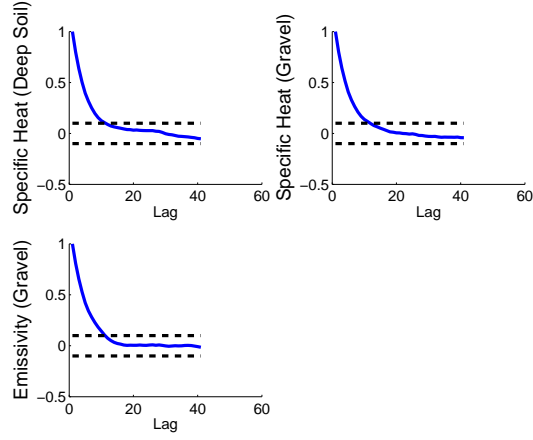


Figure D.21: Dataset H autocorrelation function, constant saturation 24-hour case.

Table D.4: Mean, standard deviation and Monte Carlo standard error (MCSE) for each parameter from Dataset H.

Parameter	Mean	St. Dev.	MCSE
Specific Heat (Deep Soil)	1.692 E-04	1.642 E-06	3.767 E-08
Specific Heat (Gravel)	8.510 E-04	3.330 E-05	8.173 E-07
Emissivity (Gravel)	9.792 E-01	1.877 E-03	2.964 E-05

## Prediction Intervals

Below are plots of the prediction intervals at all depths from using DRAM with ADH in the constant saturation 24-hour case.

Dataset A 95% Prediction Intervals

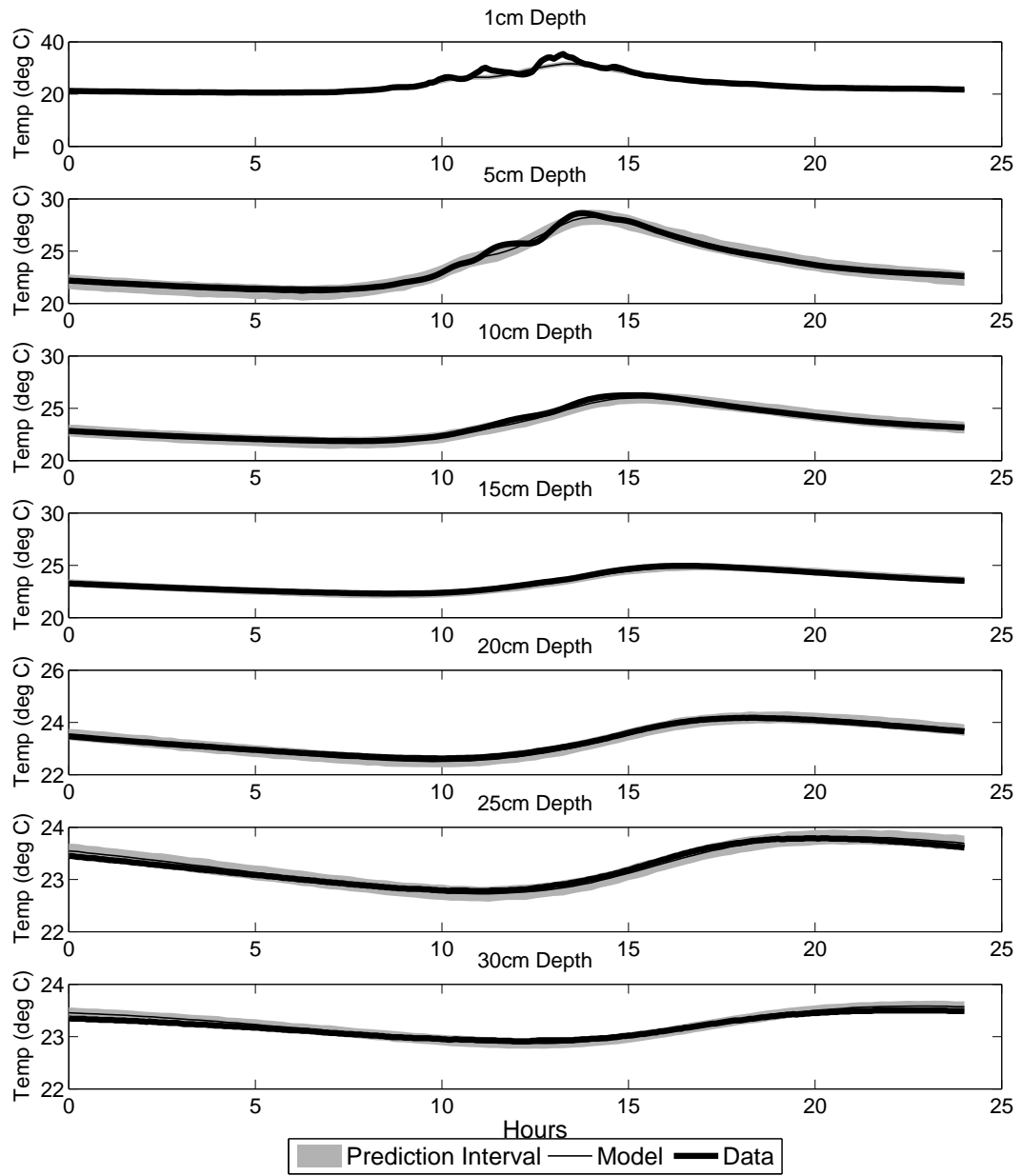


Figure D.22: 95% Prediction intervals for Dataset A, constant saturation 24-hour case with ADH as model.

Dataset B 95% Prediction Intervals

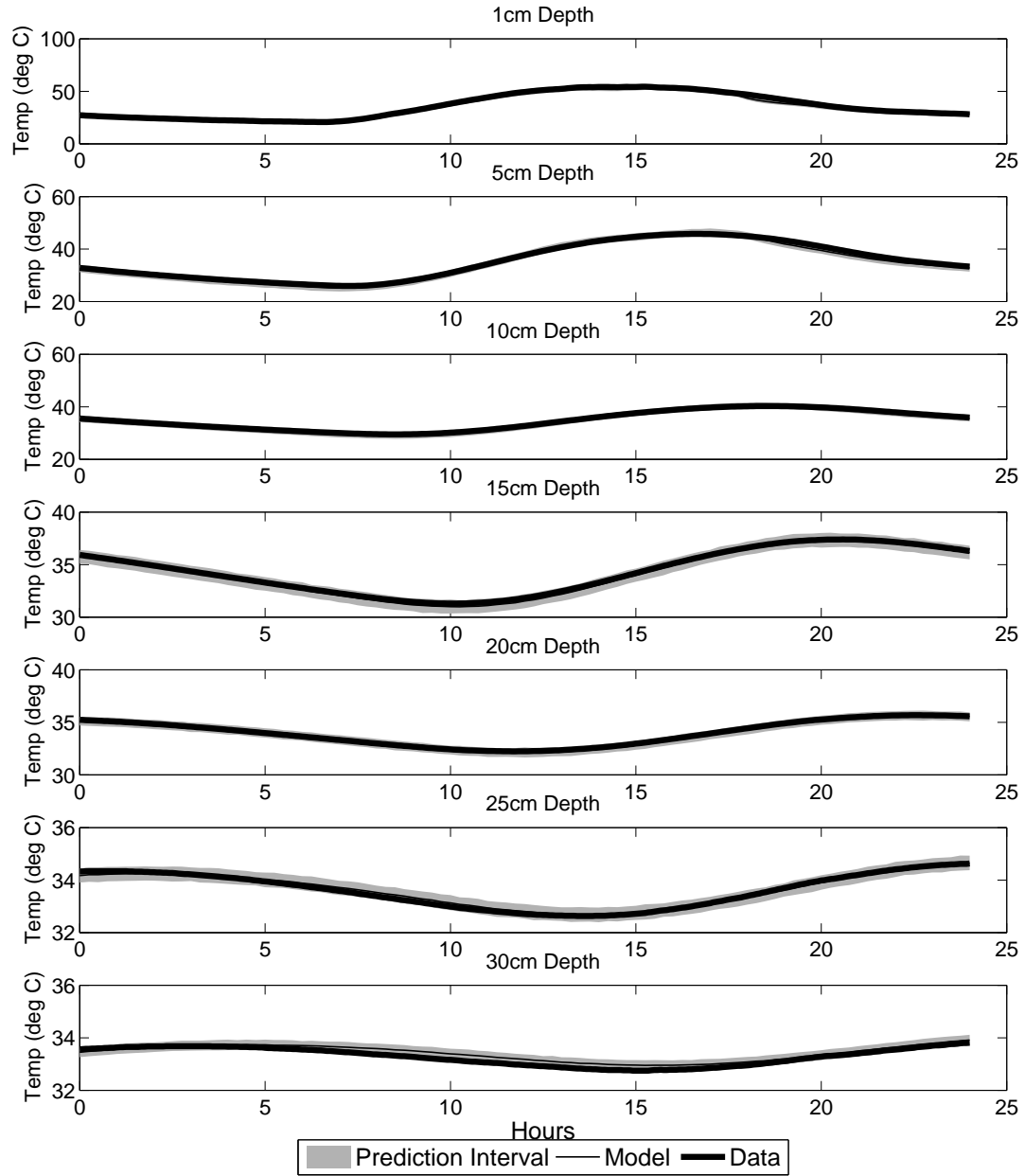


Figure D.23: 95% Prediction intervals for Dataset B, constant saturation 24-hour case with ADH as model.

Dataset H 95% Prediction Intervals

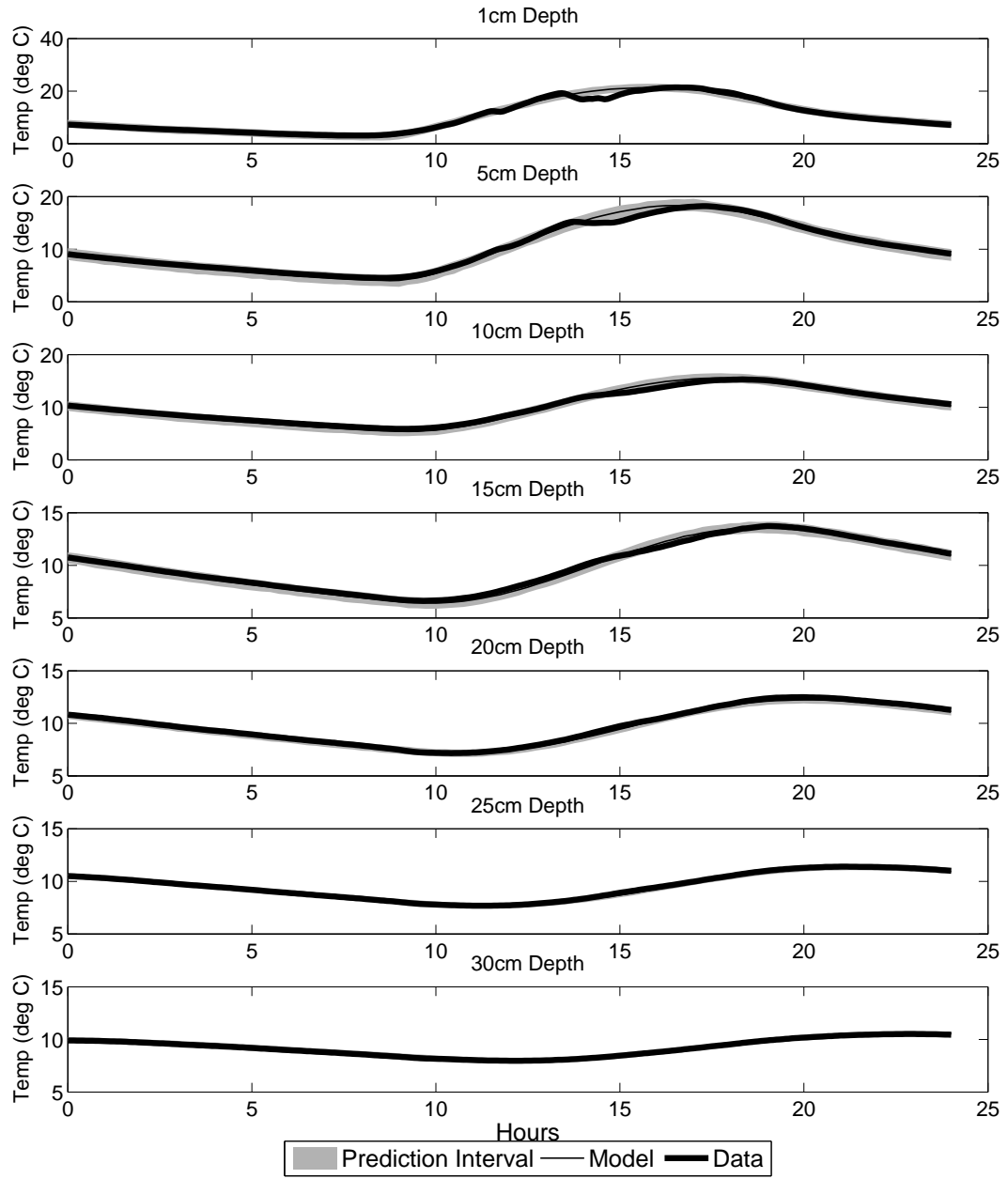


Figure D.24: 95% Prediction intervals for Dataset H, constant saturation 24-hour case with ADH as model.