Towards improved monitoring of changing permafrost by estimating soil characteristics from ground temperature time-series

by

Nicholas Brown

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Science
in
Geography,
Data Science Specialization

Carleton University
Ottawa, Ontario

ABSTRACT

Knowledge of subsurface liquid water content is important in permafrost but continuous measurements are rarely collected at monitoring sites. Two parameter estimation methods are used to estimate soil thermal properties and freezing characteristic curves (SFCC) from temperature time series in order to calculate changes in ground liquid water content. Tests with synthetic data show that even with the addition of noise, estimated SFCCs are visually similar to their true shape. Overall, saturation water content and freezing temperature are easiest to estimate, whereas heat capacity and the van Genuchten n parameter, which controls the curvature of the SFCC, are more difficult. Different calibration periods may result in high variability for estimates of low sensitivity parameters. Weighting model error by ground energy content underestimates saturation water content, but provides good estimates of freezing point temperature. Applying these techniques at monitoring sites encounters challenges when model structure is not well chosen.

ACKNOWLEDGEMENTS

I would like to acknowledge the generous support, financial and otherwise, that has made this research possible. First of all, I want to extend a huge thanks to Stephan Gruber whose seemingly limitless enthusiasm and encouragement has made the past two years an incredible experience. Thank you for all the opportunities to learn, do fieldwork, teach, and explore academic curiosities (even if they were a bit of a derailment from writing!)

Funding for this project has been provided by W. Garfield Weston Foundation, the Ontario Ministry of Training, Colleges and Universities, the David and Rachel Epstein Foundation Fund and the family, friends, and colleagues of Dr. Thomas Betz through an endowed award in his memory. Support for fieldwork and safety training has come from ArcticNet as well as the Northern Scientific Training Program (NSTP) and the resources for the computer simulations have been provided by the Southern Ontario Smart Computing Innovation Platform (SOSCIP).

I would also like to extend my gratitude to Chris Burn, whose wealth of permafrost knowledge and helpful feedback have improved the readability of this thesis tremendously. And also to Julio Valdés who provided valuable insights and ideas that helped kick off the project. Of course, the project would not have been possible without the data provided by Sharon Smith and the Geological Survey of Canada as well as by Ketil Isaksen, or without the Carleton Research Computing team who kept all the simulations running.

Finally, to Rosanna who I can't thank enough for putting up with my writing-induced existential crises, reading endless drafts of my work and, most importantly, for helping me keep things in perspective.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1: INTRODUCTION

1.1 Context

This thesis investigates the use of two techniques to estimate soil thermophysical properties and freezing characteristic curves at permafrost monitoring sites from *in-situ* ground temperature measurements. It builds on and extends previous research by using techniques that require less detailed site information. The estimates allow calculation of changes to subsurface energy, ice, and water content.

1.1.1 Permafrost monitoring

Permafrost is estimated to underlie between 12.8 and 17.8% of the terrestrial northern hemisphere (Zhang et al., 2000) and between 9 and 14% of the exposed land surface north of 60°S (Gruber, 2012). As a consequence of increased air temperatures, permafrost is warming and degrading (Vaughan et al. 2013), resulting in ground subsidence (Nelson et al. 2001; Streletskiy et al. 2017), ground slumping (Lantz & Kokelj 2008; Kokelj et al. 2009), and damage to infrastructure (Fortier et al. 2011).

Higher ground temperatures also lead to a greater amount of unfrozen water in permafrost. The proportion of unfrozen water affects the geotechnical strength of the soil (Williams & Smith 1989; Hivon & Sego 1995; Lovell 1957) and the rate of biological activity (Clein & Schimel 1995). Despite their importance, continuous records of soil water content are not readily available in permafrost environments (Marmy et al. 2016). Although databases of soil water content monitoring data do exist as part of other global monitoring efforts (Pellet et al. 2016; Dorigo et al. 2011), few sites are in permafrost-affected areas, and most measurements are restricted to the active layer.

Ground temperatures, however, are commonly tracked and are key indicators used to monitor changes to permafrost (Smith & Brown 2009). These data are widely used because they are relatively inexpensive to collect over prolonged periods of time. Ground temperature data are now available for hundreds of monitoring sites from national and international databases (Christiansen et al. 2010; Biskaborn et al. 2015; Vonder Mühll et al. 2008).

Unfortunately, the use of temperature measurements alone is not sufficient to quantify the state of permafrost without additional information about soil properties. Warming rates in permafrost are significantly lower for sites with ground temperatures near 0°C compared with colder sites (Smith et al. 2010; Romanovsky & Osterkamp 2000). This is interpreted, in part, as a consequence of the latent heat absorbed during melting of interstitial ice (Smith et al. 2005; Romanovsky & Osterkamp 2000; Smith et al. 2010). As ice-rich permafrost nears 0°C, measured warming rates are reduced due to the additional energy required to melt ground ice (Figure 1). Although the reduced warming rate may be observed in the temperature record and used to infer the melting of ice, the exact rate of ice loss cannot be quantified without knowledge of site-specific soil properties at the site. This means that it is difficult to meaningfully compare rates of change between permafrost sites. However, if soil characteristics are known, changes to both ground energy and liquid water content can be calculated from the temperature record.

1.1.2 Unfrozen water

The soil freezing characteristic curve (SFCC, Figure 2) relates the liquid water content in a soil to its temperature (Kurylyk & Watanabe 2013). In a porous medium such as soil,

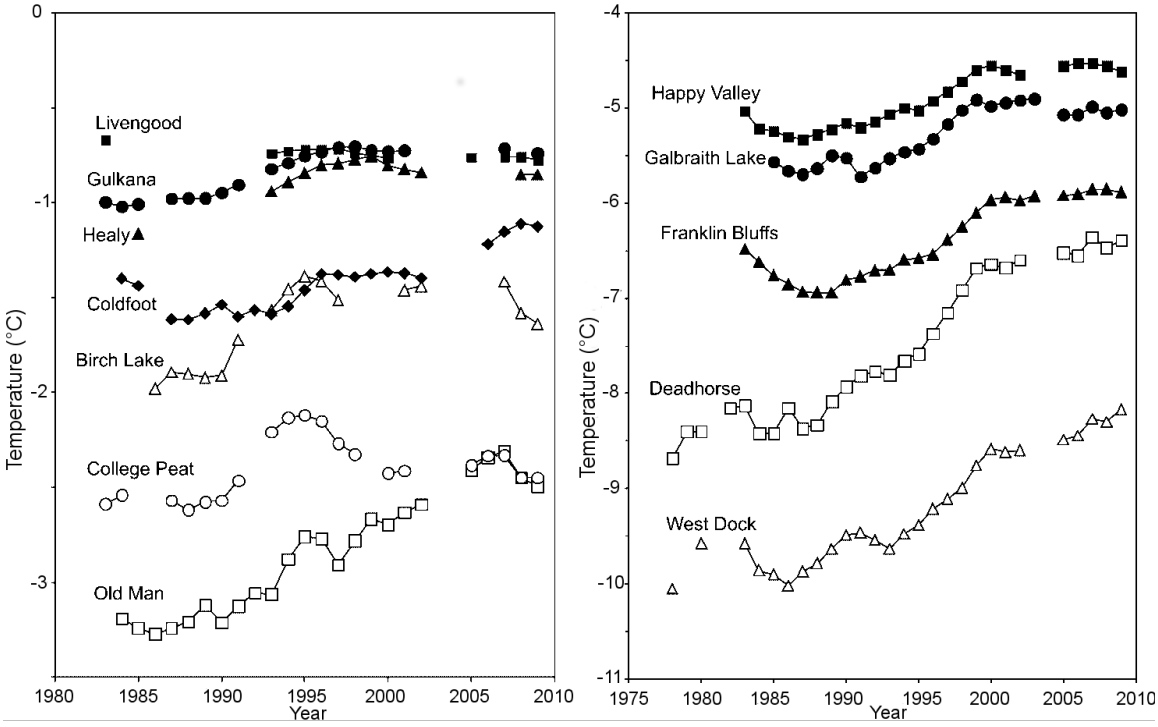

Figure 1*: The ground temperature record from permafrost monitoring sites in Alaska. Warming rates are greater at colder sites than at warmer ones where ice melt may take place. Note the scale difference between the two figures. Figure from Smith et al. (2010), reproduced with permission of John Wiley and Sons (license number 4270240022577).*

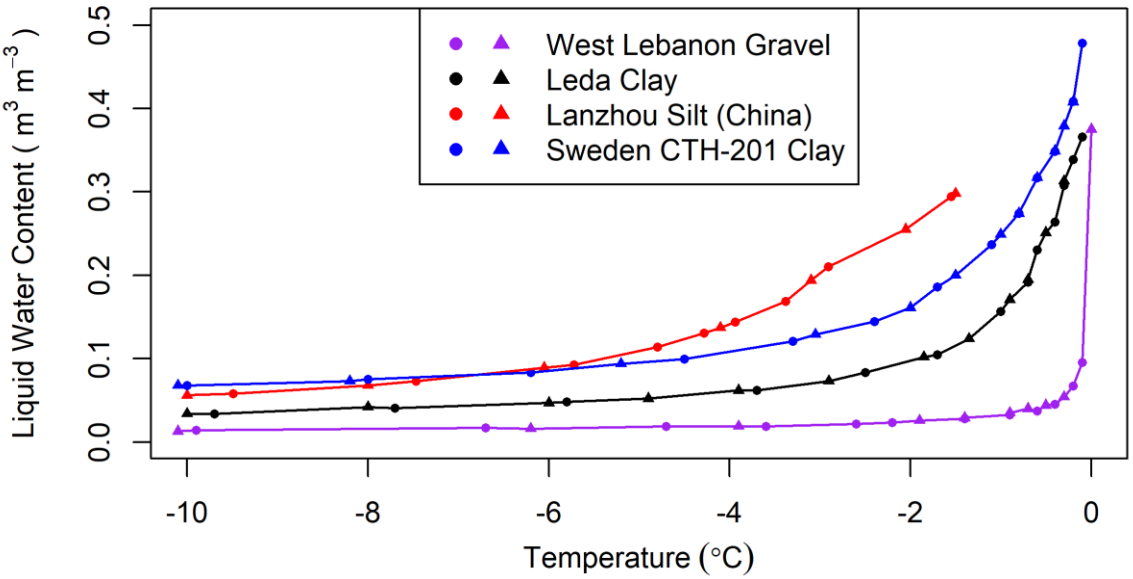

***Figure 2**: Measurements of volumetric soil water content at sub-zero temperatures using NMR (circles) and TDR (triangles). Lines representing approximations of an empirical SFCC have been added as a visual aid. Modified from Smith & Tice (1988).*

water does not melt or freeze at a single temperature. Instead, freezing takes place over a range of temperatures and initiates at a freezing temperature below the equilibrium freezing temperature of the same water without the influence of soil particles and structure (Williams & Smith 1989; Williams 1964; Lovell 1957; Tice et al. 1978).

SFCCs are often determined empirically by measuring soil liquid water content at known temperatures using a variety of methods including time-domain reflectometry (TDR), nuclear magnetic resonance (NMR) and differential scanning calorimetry among others (Tice et al. 1978; Kozlowski 2003a; Smith & Tice 1988). Examples of sub-zero soil water measurements for various soil types are shown in Figure 2. Measuring unfrozen water content becomes difficult near the melting point (Spaans & Baker 1996), and often empirical curves are based on only a few measurements (Kozlowski 2003a). Most measurements also require that an intact sample be collected, which can be difficult. *In-situ* measurements are possible with TDR (Boike & Roth 1997) and borehole NMR (Kass et al. 2017) but, in the case of TDR, the sampling location must be accessible which generally limits applicability to the active layer and shallow permafrost. For borehole NMR, the instruments are not designed to remain in place for prolonged periods of time, making these measurements only snapshots of water content at a single temperature. Therefore, reconstruction of SFCCs using this technique would be challenging.

In the absence of measurements, SFCCs can also be predicted using other available data. For instance, Anderson & Tice (1972) demonstrated that soil specific surface area can be used to predict the SFCC. Alternatively, the theoretical similarities between soil freezing and soil drying (Miller 1966) make it possible to approximate the SFCC using the soil

water characteristic curve which describes the water content of a soil as it dries (Spaans & Baker 1996; Koopmans & Miller 1966). This simplification is advantageous because of the extensive experimental basis for predicting soil water characteristics from easily obtained measurements such as soil texture (Vereecken et al. 2010; Schaap et al. 2001). However, these approximations are limited in the extent to which they can represent the effect of solutes (Nixon 1986; Banin & Anderson 1974) and pore structure (Morishige & Kawano 1999; Kozlowski 2003b).

1.1.3 Parameter estimation from physically-based models

As an alternative to the methods described above, soil properties such as the SFCC can also be estimated by calibrating a physically-based model to fit observed data and then using the calibrated model parameters as estimates of the corresponding soil properties (Bateni et al. 2012; Nicolsky et al. 2007). Mathematically, the calibration process involves finding the parameters that minimize the value of an objective function: a function that relates the parameter set to an associated measurement of model error (Nicolsky et al. 2007).

There are a number of ways to find parameters. Romanovsky & Osterkamp (2000) manually adjusted model parameters to obtain a good fit to ground temperature data. However, this manual approach is not systematic and is cumbersome when the number of parameters becomes larger. Nicolsky et al. (2007b) developed a method to estimate soil properties in the active layer, and this was later extended for use in permafrost (Nicolsky et al. 2009). However, this method requires a series of heuristic steps as well as some prior level of knowledge about site conditions. Furthermore, the method of Nicolsky et al.

(2007b, 2009) uses a local search algorithm to find the minimum value of the objective function. Local search algorithms require that a suitable initial approximation of parameter values be provided in order to converge on the best estimate (Nicolsky et al. 2007). They thus only find a locally-best estimate, i.e. in the neighbourhood of the initial estimate.

In contrast, there are many techniques to estimate model parameters that do not require a high degree of site knowledge but instead require only an acceptable range of values for each parameter. For example, global search algorithms are not sensitive to the choice of the initial approximation and have also been used to estimate parameters for numerical models (Bateni et al. 2012; Duan et al. 1992; Duan et al. 1994). They are designed to converge on the minimum value within a large search space.

Alternately, Monte-Carlo-based techniques randomly sample possible parameter combinations instead of relying on an algorithm to find an optimal value. One example of this is the generalized likelihood uncertainty estimation (GLUE) technique (Beven & Binley 1992; Beven 2012). The GLUE technique rejects the idea that one optimal parameter set can be found because of errors inherent in all models and observations. Instead, GLUE assigns a relative measure of likelihood to each set of parameter estimates based on how well it fits the observed data. Parameter sets that result in unrealistic model behaviour are discarded, and the remainder are used to calculate quasi-probabilistic estimates of model output. This technique is commonly applied to hydrological models (Lamb et al. 1998; Shen et al. 2012; Beven & Binley 1992) but has also been used for models of permafrost (Marmy et al. 2016) and seasonally frozen soil (Hansson & Lundin 2006). Even though GLUE can be used to estimate parameters in the absence of detailed

site information, the degree of prior knowledge must be explicitly stated by specifying the probability distributions from which trial parameter sets are generated (Beven & Binley 1992). Furthermore, different initial probability estimates are known to impact the results of the analysis (Beven & Freer 2001; Beven 2012).

Although parameter fitting techniques have been used to calibrate soil properties for ground thermal models in permafrost environments, the results have chiefly been applied to the prediction of future permafrost states (e.g. Harp et al., 2016; Marmy et al., 2016). Little work has been done that uses the resulting parameters to estimate observed changes to subsurface heat or liquid water content at the monitoring sites. Furthermore, many studies either deliberately avoid treating estimated parameters as representative of their physical counterparts (Marmy et al. 2016) or make the implicit assumption that, because the parameters are able to recreate observations, they are good surrogates for measurements (Bateni et al. 2012). Given the value of soil parameters for permafrost monitoring, a more thorough investigation is needed to determine whether calibrated estimates are accurate.

1.1.4 The objective function

Parameter estimation techniques are dependent on the objective function, which quantifies model error for a given set of parameters. The choice of an appropriate objective function is crucial, as it has been shown to affect the result of parameter estimation when calibrating models (Krause & Boyle 2005; Moussa & Chahinian 2009). For permafrost modelling, the root mean squared error (RMSE) between simulated and modelled ground temperatures is often used

(Atchley et al., 2015; Nicolsky et al., 2007b). However, Marmy et al. (2016) report that using the coefficient of determination (R^2) as an objective function improves calibration results near the freezing point. They also suggest that the total energy content of the ground could be incorporated into the objective function in future research. With the exception of Marmy et al. (2016), there has been little comparison of different objective functions for permafrost models.

1.2 Aim and objectives

The aim of this work is to investigate how well parameter estimation techniques can support the derivation of changed subsurface water and energy content from time series data of ground temperatures monitored in boreholes. Two methods of estimating soil properties that do not require detailed prior site information as an initial estimate are used as exemplars. This is accomplished by addressing three objectives:

The first objective is to estimate soil parameters based on temperature time series, a ground-temperature model, and two differing calibration methods. This will illustrate whether the model is sufficiently sensitive to the parameters to allow for estimation. It may also be the case that the optimization procedure yields parameters that do not reproduce the observations sufficiently well.

The second objective is to evaluate the stability and accuracy of estimated parameters with respect to observational noise and changes to calibration data. This is an extension of the first objective, but also considers whether the calibrated parameters are representative of site conditions instead of only considering whether or not they produce a good fit to observed data. If estimated parameters are overly sensitive to the effect of noise, or to other

small model perturbations, then it is less likely that they are representative of site conditions.

The third and final objective is to investigate the effect of an objective function that weights model errors based on differences in total heat content instead of differences in temperature. Research in hydrological modelling describes how different objective functions can affect the values of estimated parameters. Because temperature differences have different physical meaning at different absolute temperatures, translating temperature error into energy error may improve the estimated parameters, possibly at the cost of degraded temperature estimates.

1.3 Methodology

Two calibration methods, generalized likelihood uncertainty estimation (GLUE, Beven & Binley 1992) and differential evolution (DE) (Storn & Price 1997), are used to estimate soil parameters using synthetic temperature time series. The GEOtop model (Dall'Amico et al. 2011) is used to produce estimated ground temperature data for trial parameter sets by simulating a one-dimensional soil profile. To test the robustness of the calibration methods, two perturbed datasets are created by adding noise alternately to either the temperature values or the depths at which the temperatures are given and repeating the experiments. The resulting error in parameter estimates is compared to an ideal case with perfect data. Synthetic data are used here instead of measured data because they allow estimates to be compared to a known, true value. Synthetic data also make it possible to isolate the influence of observational error.

In the GLUE parameter fitting experiment, an objective function that takes into consideration the energy content of the soil profile is compared with an objective function that only considers observed temperatures. Here the expectation is that by taking into consideration the total ground energy content, the performance of the calibration algorithm should be improved with respect to predicting ground liquid water or energy contents, possibly at the expense of the calibration of temperature.

Lastly, the two methods are tested using borehole data. Although soil liquid water content data are not available for these sites to validate predictions, the calibrated parameters, the modeled temperature, and the ground liquid water and energy content time series are evaluated.

1.4 Thesis structure

Chapter 2 provides an introduction to the subject material and a more in-depth description of the motivation for the research. Chapter 3 introduces the ground temperature data that were used and outlines the details of how the numerical modelling was done. The experiments that were conducted and the rationale behind each one are also introduced. The results of these experiments are presented in Chapter 4 and interpretations in light of the research questions are presented in Chapter 5. The assumptions and uncertainties of the study are discussed in Chapter 6, along with the results of the experiments in the context of the published literature. Finally, in Chapter 7 the important results are summarized in a short conclusion.

CHAPTER 2: BACKGROUND

2.1 Introduction to permafrost

Permafrost is defined as ground material that remains at or below 0°C for two or more years (Harris et al. 1998). It is overlain by a layer of seasonally frozen ground known as the active layer. In some cases, as permafrost thaws, it is possible to develop a persistent layer between the active layer and permafrost that remains for multiple years above 0°C known as a residual thaw layer (French 2007 p. 112). The occurrence of permafrost is governed by processes operating at many spatial scales. When considering the occurrence of permafrost at a global scale, latitude, elevation and continental-scale physiography are the most important parameters as these affect the mean annual air temperature. Permafrost is more common at higher latitudes, greater elevations, and areas where physiographic features restrict precipitation or produce cooler air temperatures (Brown et al. 1998; Heginbottom & Dubreuil 1993).

2.1.1 Occurrence of permafrost

Under equilibrium conditions, lower ground surface temperatures will produce permafrost that is thicker, colder and more spatially abundant. At smaller spatial scales, the thickness and timing of the snowpack provides a strong control on permafrost thickness by insulating the ground during the cold season (Mackay & MacKay 1974; Goodrich 1982). Anything that affects the distribution of snow will also affect the occurrence and characteristics of permafrost on a local scale. This can include trapping of windblown snow by vegetation (Kokelj et al. 2009; Smith 1975) or anthropogenic obstacles (Hinkel & Hurd 2006; Mackay 1993). The presence of peat on the ground surface results in lower

subsurface temperatures because it is as a good conductor of heat in winter due to a high ice content but insulates the ground in the summer when it is dry (Williams & Smith 1989). Consequently, at the margins of permafrost (the discontinuous permafrost zone), its occurrence is commonly restricted to peat-rich terrain (French 2007).

Mapping the spatial extent of permafrost presents challenges. It has traditionally taken an approach that partitions the earth surface into zones characterized by the fraction of ground expected to be underlain by permafrost (e.g. Brown et al. 1998). More recently, an acknowledgement of the difficulty in drawing sharp distinctions between zones has led to the development of continuous permafrost likelihood maps, which describe the qualitative (Gruber 2012) and/or quantitative (Bonnaventure et al. 2012) likelihood of encountering permafrost within a grid cell. Further refinement of this technique provides an interpretation key which identifies whether sub-grid landscape features will tend to be warmer or cooler, thus reducing or increasing permafrost likelihood (e.g. Boeckli et al., 2012). Mapping permafrost is complicated by climate change. Some permafrost maps were created using equilibrium models that assume a steady-state temperature regime (Henry & Smith 2001). However, in the presence of a transient climate, the changes in ground temperature lag behind changes in air temperature and so steady-state models may overpredict rates of change in ground temperature. This effect is especially pronounced when the ground temperature nears 0°C (Riseborough 2007).

2.1.2 Relevance of permafrost change

Improved understanding of permafrost change is relevant for a number of reasons, particularly as it relates to climate change and implications for infrastructure and

development. An estimated 1300 ± 200 Pg of carbon—more than is currently in the atmosphere—is thought to be stored in permafrost regions (Hugelius et al. 2014). Of this, approximately 800 Pg are thought to be perennially frozen. As permafrost thaws, this carbon decomposes, thereby producing methane and carbon dioxide, which further contribute to climate change. The timing and magnitude of carbon release from permafrost into the atmosphere remains an area of active research (Schuur et al. 2015).

In cold regions, permafrost is an important consideration during the construction and maintenance of infrastructure. Changes to the ground surface associated with construction and other disturbances can warm the underlying permafrost and cause it to deform or subside (Fortier et al. 2011). Various techniques have been established to mitigate the effects of construction on the underlying permafrost (Bommer et al. 2010; Doré et al. 2016). The importance of understanding changes to permafrost has led to the development of a number of monitoring efforts.

2.2 Permafrost monitoring and databases

Permafrost has been identified by the Global Climate Observing System (GCOS) secretariat as one of 13 terrestrial essential climate variables (ECV) (Smith & Brown 2009; Sessa & Dolman 2008). The ECVs are a set of global indicators that are feasible to measure on a large scale and that contribute to the characterization of Earth's climate as it changes. Ground temperature and active layer thickness are the two key permafrost measurements that are collected as a part of this program (Smith & Brown 2009). However, both are limited in the extent to which they can resolve certain types of change. Measurements of active layer thickness (ALT) are commonly used to provide first-order information about

changes to permafrost condition (Burn 1998; Mackay & Burn 2002). The Circumpolar Active Layer Monitoring (CALM) program was developed to observe the response of the active layer to climate change (Brown et al. 2000). There are a number of ways these measurements are taken, each with their own challenges: physical measurements of ALT using a metal probe are quick and inexpensive but only measure the instantaneous thickness of unfrozen material, known as the thaw depth. The thaw depth typically approximates the true active layer thickness in the early to late fall, but the exact timing of maximum thaw depth is site-specific and temporally variable (Brown et al. 2000). Active layer or thaw-depth probing is also limited by spatial heterogeneity and by requiring sites with amenable soil textures. Physically probing the active layer is also limited to relatively thin active layers. Mackay (1973) developed a frost tube that comprises a flexible plastic tube filled with ice and placed inside a sealed borehole casing. A bead placed on top of the ice records the maximum annual thaw depth as it descends with the melting ice and refreezes in the winter. Both mechanical probing and thaw tubes require annual site visits to ensure data continuity. Geochemical sampling as a compliment to, or replacement for, active layer probing has also been proposed as a qualitative way to detect fine-scale changes in ALT (Keller et al. 2010).

As noted above, permafrost is defined by temperature. This means that ground temperature measurements are most commonly used to monitor changes to permafrost and consequently, permafrost scientists have an ever-increasing amount of ground temperature data. The Thermal State of Permafrost (TSP) project arose out of the International Polar Year (IPY) and provided snapshots of permafrost temperature profiles in North America

(Smith et al. 2010), the Nordic area (Christiansen et al. 2010), Russia (Romanovsky et al. 2010) and China (Zhao et al. 2010). The Permafrost and Climate in Europe (PACE) project made available borehole temperature data from mountainous environments (Harris et al. 2001) for the purpose of monitoring climatic change. The Swiss permafrost monitoring network PERMOS combines snow, air, surface and subsurface temperature measurements with automated geophysical monitoring and measurements of rock glacier movement (Vonder Mühll et al. 2008). More recently, the Global Terrestrial Network for Permafrost (GTN-P) has developed a web portal from which temperature data associated with TSP and CALM can be obtained (Biskaborn et al. 2015). However, this service is still in its early stages; metadata completeness is 63% and 50% for active layer sites and boreholes respectively.

These databases share a limitation: although ground subsidence is a major concern in permafrost areas, detecting surface displacement is not possible with measurements of temperature or active layer thickness. Instead, repeat surveys can be conducted by collecting measurements of elevation directly, or using remotely-sensed data (Short et al. 2011; Jones et al. 2013). Airborne measurements have the capacity to detect change over a larger area but validation of results can be difficult in remote environments (Short et al. 2011).

2.3 Ice in permafrost

One of the important characteristics of permafrost is that it permits the preservation and growth of subsurface ice. As the ground freezes, the existing moisture within the soil pores is converted to pore ice throughout the permafrost (French & Shur 2010). Sellmann &

Hopkins (1983) clarified the term 'ice-bonded' permafrost to mean permafrost that contains a sufficient quantity of ice to increase the geotechnical strength of the ground. This is in contrast to 'ice-bearing' permafrost which contains ice, but in quantities that do not appreciably affect the strength. This language was originally developed to describe subsea permafrost (Sellmann & Hopkins 1983) but now is also used to describe terrestrial permafrost (Romanovskii et al. 2005; Dallimore & Collett 1995).

Ice in permafrost can accumulate in quantities in excess of the porosity of the unfrozen soil (this is known as excess ice). Excess ice can form as large wedge-shaped bodies from seasonal surface water infiltration into thermal contraction cracks (Mackay & Burn 2002), as lenses formed by moisture migration across a temperature gradient (Cheng & Wang 1983; Wolfe et al. 2014), and as layers through the co-deposition of sediment with ice and snow (Humlum et al. 2007). This is by no means an exhaustive list; for a more in-depth treatment of the subject, the reader is referred to Gilbert et al. (2016) and French & Shur (2010). Although excess ice is not dealt with explicitly in this thesis, it is mentioned here because its exclusion represents a significant limitation in the applicability of the methods described in Chapter 3.

2.4 The importance of latent heat

Latent heat of fusion (L_f) is the energy required to change the phase of some finite volume of material from solid to liquid or vice-versa (Williams & Smith 1989). In permafrost, latent heat influences the response of ground temperature to energy input. This is sometimes expressed as the temperature-dependent apparent heat capacity of a soil (French

2007), and is one of the reasons temperature data are unable to fully describe rates of change in permafrost.

The behaviour of ground temperature in the presence of phase change can in some cases permit the interpretation of more information from temperature changes than would otherwise be possible. Latent heat is responsible for the zero-curtain effect in permafrost regions: as the ground in the active layer freezes back during the fall, the rate of temperature change is dramatically reduced and typically remains near the freezing point for a prolonged period of time until the water has frozen completely (Williams & Smith 1989; Outcalt et al. 1990). By interpreting this phenomenon in terms of latent heat, it is possible to make qualitative inferences about the soil composition and hydrology simply by examining temperature behavior.

The incorporation of latent heat in permafrost modelling is an important consideration. Determination of thermal properties from temperature measurements in tundra soils by McGaw et al. (1978) found that apparent thermal diffusivities decreased as permafrost warmed. They hypothesized that this was a consequence of the latent heat absorbed as ice melted into water. Later, results by Osterkamp & Romanovsky (1997) confirmed this hypothesis by using modelling-derived SFCCs to calculate the apparent thermal diffusivities of freezing soil. They found that the calculated thermal diffusivities matched well to the earlier measurements.

The effects of latent heat reduce rates of permafrost thaw and are important modelling considerations. Nixon (1986) demonstrated that rates of permafrost thaw are significantly lower when subsurface ice melts according to the SFCC than when it melts at

a single temperature. When latent heat is taken into consideration, model results predict decreased warming rates in permafrost close to 0 °C (Etzelmüller et al. 2011). Lunardini (1996) demonstrated that permafrost would persist longer than had been predicted (Nelson & Anisimov 1993; Gavrilova 1993) by using a physically-based model that took into consideration the dynamics of permafrost thaw including the latent heat.

Latent heat is problematic because it precludes meaningful, direct comparison between permafrost change at different temperatures. Warming rates are much lower at warm permafrost sites where ground temperature is near 0 °C and where ice is melting, than at cold sites (Smith et al. 2010). This is explained by the latent heat of ice which increases the apparent heat capacity of the soil resulting in lower rates of temperature change per unit of energy absorbed into the ground.

2.5 Unfrozen water in frozen soil

2.5.1 Causes of unfrozen water in soil

There are several phenomena that permit the existence of liquid water below 0°C. Dissolved solutes will lower the freezing point of water and although this effect is present in all soils, it is especially noticeable in subsea permafrost or other saline materials. For example, depressed freezing point values between -0.8 and -3.0 °C are reported for seabed soils in Alaska (Nixon 1986), whereas values of between 0.0 and -0.40 °C are reported for terrestrial soils beneath a drained lake (Mackay 1997).

Interactions between water and soil at the surfaces of soil particles also reduce the freezing point of water. Variations in the charge distribution proximal to mineral surfaces disrupt the structure of water and inhibit the formation of ice (Mitchell & Soga 2005).

Surface effects are especially pronounced for clay particles, which have a strongly negatively-charged surface. Anderson & Tice (1972) show that the specific surface area of a soil, which is a measure of the total surface area of particles, can be used to predict the unfrozen water content. Furthermore, as ice nucleates in the centre of soil pores, dissolved solutes are concentrated in the remaining water, further reducing the freezing point (Banin & Anderson 1974).

The relationship between the quantity of frozen soil water and the temperature depends on the size distribution of soil pores (Kozlowski 2003b). Water that is closer and more strongly adsorbed to mineral surfaces requires progressively lower temperatures to freeze (Williams & Smith 1989). Therefore, water in larger pores begins to freeze before water in smaller pores because the centre of the large pores is further away from mineral surfaces. This can also be explained by the geometry of the pores. Smaller pores with greater curvature lower the chemical potential of soil water and cause a freezing point depression proportional to the curvature of the pore known as the Gibbs-Thomson effect (Rempel et al. 2004; Watanabe & Mizoguchi 2002).

Finally, increased pressure lowers the freezing point of water and permits liquid water below 0°C. The pressure-dependence of the freezing point is typically not relevant for permafrost studies except at the base of very deep permafrost where it can manifest itself as a change in the geothermal gradient due to the difference in thermal conductivity between water and ice (e.g., Lachenbruch et al., 1982).

2.5.2 Implications of unfrozen water in soil

The presence of unfrozen water in frozen soil is important to understanding both the ecology and engineering properties of the soil. For instance, Clein & Schimel (1995) observed an increase in soil respiration with increased saturation for soils at temperatures below 0 °C. They suggested that the discrepancy may be due to the increased liquid water content present as thin films surrounding soil particles.

The strength of frozen soil is dependent on its temperature (Lovell 1957); colder soils have a higher compressive strength and resistance to creep. Williams & Smith (1989) give three mechanisms to explain this: the decrease of liquid water content at lower temperatures; the increased strength of ice-bonds at lower temperatures; and the reduction in pressure-refreezing of water at lower temperatures for a fixed stress increment. The degree to which unfrozen water content controls soil strength is dependent on the soil texture (Williams & Smith 1989). For fine-grained soils such as silts and clays, it is the dominant effect. In a series of experiments on saline soils, Hivon & Sego (1995) demonstrated that the strength of a soil decreases in response to increases in both soil temperature and salinity which, in turn, increase the liquid water content.

2.6 Soil freezing characteristic curves

Soil freezing characteristic curves (SFCC) describe the relationship between the temperature of a porous media and the volumetric fraction of liquid water contained within it (Figure 2).

2.6.1 Measuring soil water content

Direct measurements of soil water content, often collected by comparing the difference in weight of a soil sample before and after oven-drying, are used as standards against which other methods are compared (Bittelli 2011). However, this method is not practical in the field, destroys the soil structure of the sample, and cannot discriminate between frozen and unfrozen water.

There are a number of methods that have been developed to measure the unfrozen water content directly. Topp et al. (1980) empirically demonstrated a relationship between the liquid water content of a soil and its dielectric constant. The dielectric permittivity (or dielectric constant) of a medium describes its response to an applied electric field. Water has a very high dielectric constant ($\varepsilon \sim 80$) in comparison with ice ($\varepsilon \sim 3$), air ($\varepsilon \sim 1$), or other soil materials ($\varepsilon \sim 3$-40) (Davis & Annan 1989). Consequently, the dielectric response of a soil will depend strongly on the proportion of liquid water therein. The dielectric constant of a soil can be measured using time-domain reflectometry (TDR). The TDR method has been used to calculate the unfrozen water content in soils below 0°C (Patterson & Smith 1981) and reasonably accurate TDR measurements of liquid water content can be readily taken in the field (e.g. Boike and Roth, 1997) .

Another common method of indirectly measuring liquid water is nuclear magnetic resonance (NMR) (Tice et al. 1978; Tice et al. 1981). Here, the sample is placed within a strong magnetic field and exposed to bursts of a secondary magnetic field which perturbs the magnetic moments of the hydrogen nucleii within the sample. The response of the sample nuclei induces measurable, rapidly-decaying field variations called spin echoes.

The decay rates of the spin echoes depend strongly on whether the nucleus is contained in a liquid or solid phase, and therefore the measured response is proportional to the liquid water content in the sample (Kleinberg 2006). Smaller NMR analysers are available which permit measurements in the field with a sample (Kleinberg 2006) or *in-situ* within a borehole (Kass et al. 2017).

A third method is differential scanning calorimetry, which measures the difference in energy needed to heat two vessels, one empty and one containing a sample (Kozlowski 2003a). The amount of energy needed to heat the sample is proportional to the instantaneous quantity of phase change. This method is advantageous because it permits a very high measurement density over a range of temperatures; however, it requires a deconvolution of two functions that is not always possible using analytical methods and so numerical approximations are sometimes required (Kozlowski 2003a).

Even when soil moisture data are available, fitted equations for the SFCCs are not able to represent measured data with perfect accuracy (e.g. Lovell, 1957; Tice et al., 1981). Moreover, the use of any single function to describe the SFCC for a given soil is complicated by the fact that observations demonstrate a hysteretic behavior: multiple liquid water contents are possible for a given temperature depending on whether the soil is warming or cooling (Spaans & Baker 1996). Repeated cycles of freezing and thawing, which may homogenize soil over time by sorting particles by size (Corte 1963), can also affect SFCCs over time by changing the pore geometry and distribution of pore sizes.

2.6.2 Parameterization of SFCCs

Several different mathematical expressions have been developed to describe SFCCs (Kurylyk & Watanabe 2013). A simple power relation is often used (Lovell 1957; Black & Tice 1989) to relate the unfrozen water moisture content *w* to temperature using two empirically-derived coefficients:

$$\theta_w = \mathrm{AT}^{\mathrm{B}} \quad \mathrm{T}<\mathrm{T}^*$$

Where θ_w is the percentage of total water that is unfrozen, T is the temperature in °C, T* is the freezing point of the soil water (°C) and A and B are empirical parameters used to fit the curve to data. A similar version of this equation was used by Nicolsky et al. (2007) which includes an additional term (*T**) for freezing point depression:

$$\theta_w = \begin{cases} 1 & T \geq T^* \\ |T^*|^B |T|^{-B} & T < T^* \end{cases}$$

In some cases, the simple power function is insufficient because of the strong nonlinearities near the freezing point, and a piecewise function comprising two exponential curves is used instead. Exponential functions are also commonly used (Kozlowski 2007).

A more highly parameterized description of the SFCC is used by Dall'Amico et al. (2011)

$$\theta_w = \theta_r + \frac{(\theta_s - \theta_r)}{(1 + [-a\,\psi(T)]^n)^{\frac{1}{n}-1}}$$

Where $\psi(T)$ is physically-derived equation expressing liquid water matric potential as a function of temperature (Figure 3a). Here, *n* is a measure of the pore size distribution and α is related to the inverse of the air entry suction (van Genuchten 1980; Schaap et al. 2001). θ_s is the volumetric water content of the soil when it is saturated and θ_r, the residual water content, is the volumetric fraction of water that remains unfrozen even at extremely low

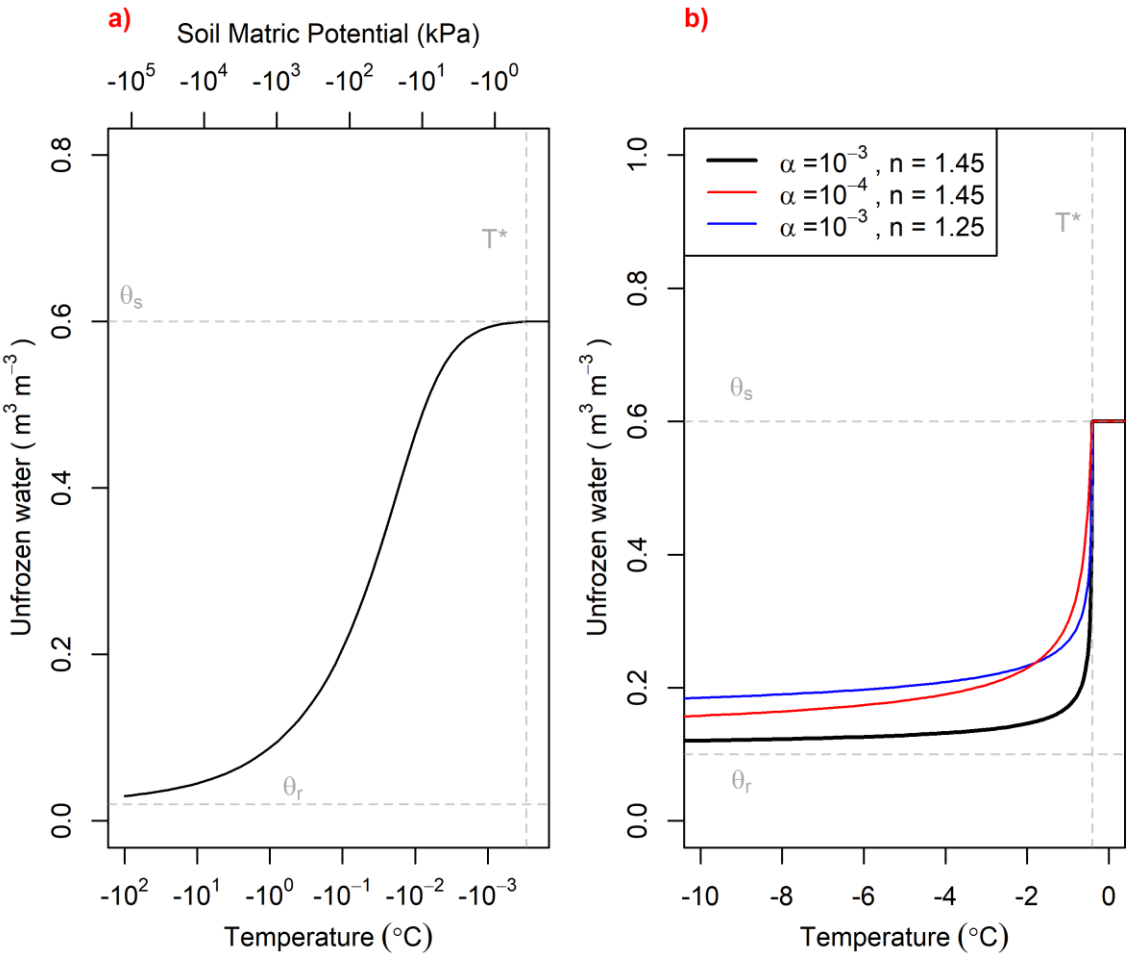

***Figure 3**: a) General model of van Genuchten water retention curve with no freezing point depression. The van Genuchten model relates soil matric potential to water content as a soil dries. The assumption that freezing and drying behaviours are similar allows the van Genuchten curve to be used to model unfrozen water content provided that a function relating matric potential to temperature can be derived. b) Idealized soil freezing characteristic curves (SFCC) for saturated soils showing the relationship between SFCC shape and the depressed freezing temperature (T*), saturation water content (θ_s), residual water content (θ_r) and van Genuchten shape parameters α and n. Note the scale differences between the two figures.*

temperatures. The form of this equation is based on the work of van Genuchten (1980) and was originally used to relate moisture contents to matric potentials during soil wetting and drying. This is known as the soil water characteristic curve (SWCC) or soil water retention curve. As will be discussed below, there are a number of similarities between soil freezing and soil drying which permit the SWCC to be used in place of the SFCC in many circumstances. In all of the described models, there is a combination of empirical parameters that are fit to observed data and physically-meaningful parameters that can be measured directly, such as saturation water content or freezing-point depression.

The theoretical similarities between soil drying and soil freezing were investigated by Koopmans & Miller (1966). Subsequent work demonstrated that soil freezing and drying can be described using the same equation with the same parameters provided that the soil has the same history (uniformly warming/wetting or cooling/drying) and bulk density (Black & Tice 1989).

More recent studies show good similarity between the two curves provided that temperature measurements are sufficiently accurate and are taken at equilibrium conditions (although this can take days to weeks depending on the technique) (Liu et al. 2012). Similarly, the SFCC has been used in place of the SWCC due to the difficulties in measuring soil moisture at low soil matric potential typical of very dry soil (Flerchinger et al. 2006). There are a few criticisms of this approach including the lack of consideration of solutes in the SWCC (Azmatch et al. 2012) which can significantly alter the shape of the SFCC across a range of subzero temperatures (Wu et al. 2015), particularly as solutes are expelled from ice forming at the centre of the pores (Banin & Anderson 1974).

2.6.3 Estimating SFCCs and SWCCs

In the absence of direct measurements of soil water content at known temperatures or matric potentials, a number of techniques have been used to estimate SFCC or SWCC indirectly based on a variety of other measurements which are more readily or more commonly available. One such approach relies on pedotransfer functions (Patil & Singh 2016).

Pedotransfer functions relate easily measured soil properties such as textural information to other properties that are more difficult to obtain (Patil & Singh 2016). Soil water characteristic curves (and, by extension, SFCCs) can be estimated from soil textural information using pedotransfer functions (Schaap et al. 2001; Vereecken et al. 2010; Wösten et al. 2001). These functions empirically relate data such as particle-size distribution and bulk density to van Genuchten parameters. The van Genuchten model was originally developed to describe the SWCC for unfrozen soil (van Genuchten 1980), but because of the similarity between the SFCC and SWCC (Koopmans & Miller 1966), it is possible to use the parameters to describe the SFCC (e.g. Dall'Amico et al., 2011). A database of 2134 soil samples with accompanying soil water retention data has been incorporated into a software package called ROSETTA (Schaap et al. 2001). Depending on the level of prior knowledge of the soil sample, ROSETTA can take a variety of different input data ranging from simple soil textural classification (e.g. 'sand' or 'silty loam') to increasingly sophisticated data such as quantitative soil textural measurements or bulk density. The model also produces an estimation of the standard deviation for each of the parameter estimates. However, both the prediction of the van Genuchten parameters and

the total moisture content is limited by the data used to calibrate the model. Evaluating soil samples that are not in the database inevitably results in higher RMSE and lower R^2 (Stumpp et al. 2009; Schaap & Leij 1998). Despite these limitations, the ROSETTA model remains a benchmark against which new estimates of soil water retention parameters are measured (e.g. Ghanbarian-Alavijeh et al., 2010; Stumpp et al., 2009; Yang and You, 2013). Other methods of estimating SFCC or SWCC include models based on Atterberg liquid limits (Tice et al. 1976) or a suite of other geotechnical data (Aubertin et al. 2003). These methods rely to some extent on the assumption that particle size can be used as a proxy measurement for pore structure and size distribution when, in fact, these are independent of the particle size and can change through soil consolidation or compaction. Furthermore, the soil samples used to develop SWCC estimation techniques have not been collected from permafrost environments (e.g. Schaap et al. 2001).

Instead of using soil textural data, SFCC parameters may be estimated indirectly using ground temperature data. Nicolsky et al. (2007) developed a technique to estimate active layer thermal and hydrologic parameters (including those necessary to reconstruct the SFCC) by modelling ground temperatures with trial parameter sets and calculating the RMSE from observational data. A search algorithm is then used to minimize the RMSE by iteratively changing the parameters. The search algorithm requires an initial approximation which is determined by a combination of a heuristic technique and extensive prior information about the study site which constrains the ranges for the parameters and provides an initial approximation to the search algorithm.

2.7 Conductive heat transfer in one dimension

Simulation of permafrost requires an understanding of the physics of heat conduction. One of the foundational equations for heat flow is the Fourier equation (Williams & Smith 1989) which, in one dimension, is represented by:

$$Q = -\lambda \frac{dT}{dx}$$

This equation describes how, in response to a temperature gradient dT/dx, heat will conduct through a solid at a rate Q proportional to the ability of that solid to conduct heat, λ. This ability of a material to conduct heat is known as the thermal conductivity [W m^{-1} K^{-1}]. Fourier's law can be used to derive the general heat flow equation (Carslaw & Jaeger 1959; Williams & Smith 1989), which is defined in one dimension as:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial x^2}$$

In plain language, this states that the rate of temperature change in a material is directly proportional to the curvature of the temperature gradient at that location. The constant of proportionality, κ is called the thermal diffusivity [m^2 s^{-1}] (Williams & Smith 1989; Carslaw & Jaeger 1959). The thermal diffusivity is related to the thermal conductivity by the volumetric heat capacity c [J m^{-3} °C^{-1}]:

$$\kappa = \frac{\lambda}{c}$$

Subject to certain simplifying assumptions, there are a number of exact solutions to the heat equation. One in particular—the Stefan solution—is commonly used to model permafrost phenomena such as active layer changes following surface disturbance (Burn 1998; Mackay 1995), permafrost degradation associated with thaw slumps (Kokelj et al.

2009), and the growth of pingos (Mackay 1990). However, in scenarios that involve changing surface conditions and climates, or more complex ground conditions, the assumptions of exact solutions limit their applicability.

Instead, it is possible to solve the heat equation numerically. Numerical methods require *discretization*—splitting up the physical space of the model into pieces of finite dimension, and splitting up time into 'steps' of finite duration (Mathews & Fink 2004; Riseborough et al. 2008). At each time-step, an approximation is made as to the trajectory of the original equation, and this trajectory is projected forward to the following time step (Mathews & Fink 2004). A number of numerical models exist that have been used to simulate one dimensional heat transfer in permafrost such as GIPL2 (Jafarov et al. 2012), ATS (Painter 2011; Atchley et al. 2015) and GEOtop (Dall'Amico et al. 2011; Endrizzi et al. 2014). Numerical models are powerful because they allow for complexities such as variable soil composition, parameterization of SFCCs and arbitrary driving temperatures at the model boundaries. However, the accuracy of modelling results depends on how well these complexities are described.

2.8 The simulator GEOtop

This thesis uses the GEOtop model to simulate the 1d soil profile. The model GEOtop was originally developed as a distributed hydrological model (Rigon et al. 2006), but has been adapted for use in permafrost environments (Dall'Amico et al. 2011; Endrizzi et al. 2014). GEOtop has been used for other glacial and periglacial research including the simulation of cold-firn temperature (Buri 2013), the effects of water flow on thaw depth (Endrizzi & Gruber 2012) and simulation of active layer temperature dynamics (Pan et al. 2016).

Analyses of model uncertainty and parameter sensitivity have been conducted for permafrost environments (Gubler et al. 2013). GEOtop uses the van Genuchten parameterization to describe water retention (van Genuchten 1980). This parameterization has been extended to describe the SFCC by invoking the similarity between soil freezing and soil drying (Miller 1966; Spaans & Baker 1996). This is advantageous because of the abundance of experimental work done to approximate van Genuchten parameters for soil water retention (e.g. Schaap et al., 2001; Vereecken et al., 2010), and the possibility of using the parameterization for unsaturated conditions (Dall'Amico et al. 2011). The four relevant parameters for the van Genuchten parameterization are: θ_s, the saturation water content, θ_r, the residual water content and two shape parameters α and n.

2.9 Estimating parameters for physical models

Modellers are often faced with the task of estimating model parameters because acquiring information about subsurface properties is difficult and costly. Permafrost models often require parameters describing the freezing characteristic curve as well as thermal properties λ and c discussed earlier (Nicolsky et al. 2007; Marmy et al. 2016; Atchley et al. 2015). One solution is to select model parameters based on how well they are able to reproduce observations. For example, Romanovsky & Osterkamp (2000) manually chose parameters for the SFCC by matching model output with observations.

There are many methods that generate and evaluate trial parameter combinations instead of picking them by trial-and-error; one option is to use an optimization algorithm. Local optimization algorithms (Mathews & Fink 2004) require an initial approximation of parameter values from which they move to parameter sets with a lower objective function.

Many of these algorithms require that the objective function be differentiable, while others such as the Nelder-Mead algorithm have no such requirement (Mathews & Fink 2004). Local search algorithms may become stuck in local minima of the objective function and fail to reach the global minimum if the initial approximation is inappropriate. Nicolsky et al. (2007) use a Nelder-Mead algorithm to estimate soil parameters with a permafrost model. In order to converge on an appropriate result, the initial approximation is selected by a sequence of heuristic steps. First, a range of likely values is established for each parameter from detailed field data. Second, an initial approximation within this range is selected. And finally, the data is split up into seasonal periods and a series of successive local optimizations are performed manually for each parameter in the season where the model error is most sensitive to variations in that parameter (Nicolsky et al. 2007). The resulting parameter set is then selected as an initial approximation for the subsequent optimization. This method unfortunately requires a great deal of prior measurements and familiarity with the site in question.

Global optimization methods overcome some of the deficiencies of local optimizations by exploring the entire search space instead of being restricted to a single basin of attraction (e.g. Duan et al., 1992). Global optimization yields a single parameter set which theoretically provides the best model fit in the search space. Global optimization methods have been used to calibrate simplified rainfall-runoff models (Duan et al. 1992), hydrologic models of full-watersheds (Duan et al. 1994) and soil thermophysical properties (Bateni et al. 2012).

Parameter estimation techniques for permafrost have their limitations: Harp et al. (2016) were unable to determine appropriate values for the van Genuchten α and n because the model was not sufficiently sensitive to these parameters. Methods that require an initial approximation (Nicolsky et al. 2007) can only be used for sites with sufficient prior information. Although certain highly-documented study sites may have this level of information, it is difficult to scale up the procedure, or to use at sites where only the ground temperatures are available. This thesis aims to address these limitations by using two calibration techniques that are not sensitive to the choice of initial estimate and therefore require less prior information about the monitoring site. These methods— differential evolution (DE) and GLUE—are discussed below.

2.9.1 Differential Evolution

Differential evolution (Storn & Price 1997) is used in this thesis as an exemplar of a global search algorithm. Differential evolution (DE) begins with a relatively small set of trial parameter sets and iteratively tests and modifies them. Modification is done in such a way as to cause the set of parameters to converge on an optimal value that represents the parameter set that causes the model to fit the data best (Storn & Price 1997). For each iteration, a single parameter set (referred to as the target parameter set) $\mathcal{C}_i$ from the population is tested against a modified parameter set $\mathcal{C}_i^*$. If the modified parameter set performs better than the original, the original is replaced. The following procedure (Storn & Price 1997) is used to modify and test any given parameter set $\mathcal{C}_i$: first, the scaled vector difference between two randomly chosen parameter sets is added to a third parameter set to create a mutated parameter set: $\mathcal{C}_4 = k * (\mathcal{C}_1 - \mathcal{C}_2) + \mathcal{C}_3$ (Figure 4). The third

parameter set ($\mathcal{C}_3$) is typically the best candidate in the population, but can also be randomly selected. Next, a trial candidate $\mathcal{C}_i^*$ is created using parameters chosen randomly from either $\mathcal{C}_i$ or $\mathcal{C}_4$ (Figure 5). At least one parameter is always chosen from $\mathcal{C}_4$.
If the objective function of the trial candidate $J(\mathcal{C}_i^*)$ is less than that of the original $J(\mathcal{C}_i)$, then the original is replaced. Each parameter set in the population is iteratively mutated and evaluated in this way, and the population convergence is evaluated at the end of each pass. If all members of the population are sufficiently close, the optimization terminates, otherwise the procedure is repeated.

2.9.1 GLUE and equifinality

Instead of using an algorithm to estimate the 'best' parameter set, the GLUE technique acknowledges that for a set of observations, there may be a number of parameter sets which fit the data equally well (Beven & Binley 1992). The GLUE method also recognizes that the simulation error may be due not only to incorrectly specified parameters but also to model structural error, or the inherent inability of a model to fully capture the behavior of the system under consideration. In the GLUE method, the model error is used as an estimate of model likelihood. This is not a statistically rigorous likelihood but rather a measure of relative likelihood between simulations (Beven & Binley 1992). GLUE is widely used to calibrate hydrological models (Smith et al. 2008; Shen et al. 2012) and has also been used to estimate parameters for permafrost models (Marmy et al. 2016). Implementation of the GLUE technique requires that a series of subjective decisions be made by the researcher (Beven 2012): first, prior parameter distributions must be selected to represent the *a*

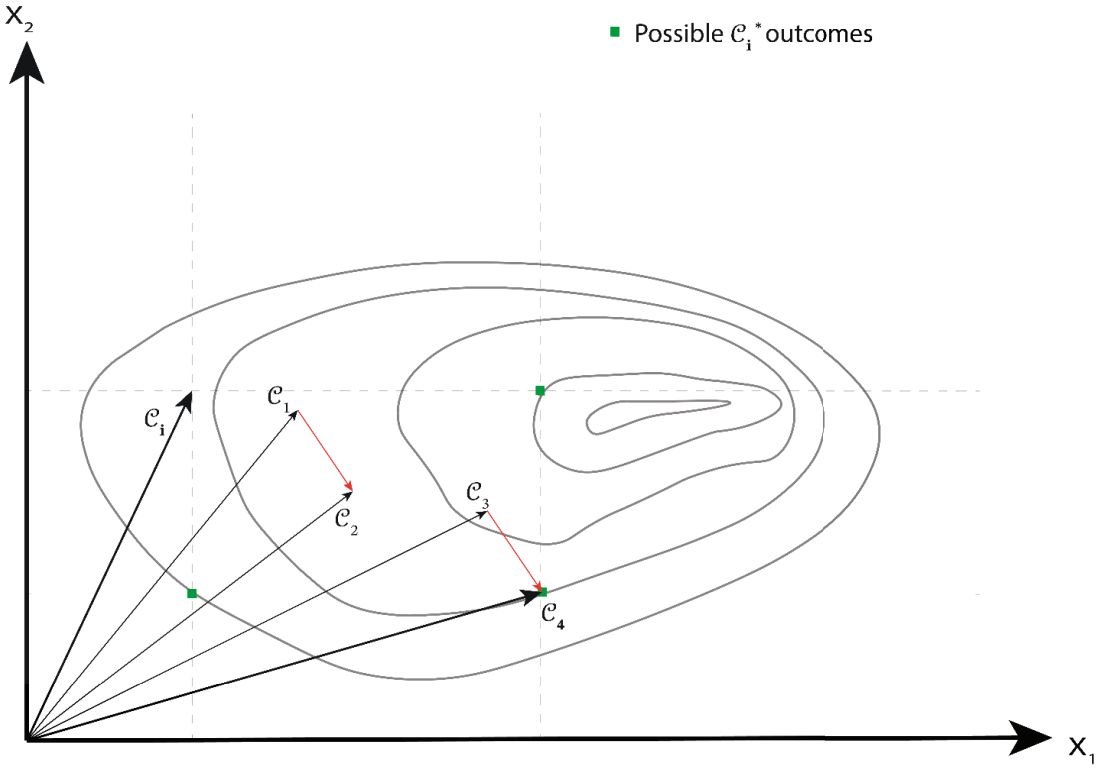

***Figure 4**: Creation of mutated parameter set $\mathcal{C}_4$ from other population members. A trial vector is created by taking components from either the target vector $\mathcal{C}_i$ or the mutated vector. At least one component must be taken from the mutated vector. Possible outcomes for $\mathcal{C}_i^*$ are shown as green squares. Contours represent isolines of a hypothetical two-dimensional objective function. Figure modified from Storn & Price (1997).*

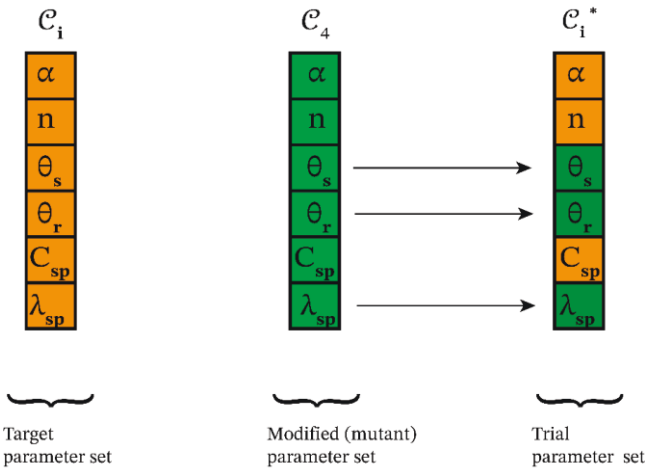

***Figure 5**: Creation of trial parameter set from components of target parameter set and mutated parameter set. Note that while a two-dimensional objective function is used in Figure 4 for the purpose of explanation, a six-dimensional objective function is used here. Figure modified from Storn & Price (1997).*

priori state of knowledge about the parameter likelihood; second, a sampling strategy must be chosen to create trial parameter combinations by sampling from the prior parameter distribution; and third, a likelihood function must be chosen based on the model error compared to observations. In determining the likelihood measure, a criterion must be defined to separate simulations into behavioural and non-behavioural. Non-behavioural simulations are those that do not reproduce observations sufficiently well (Beven 2012). This criterion can be defined either as a threshold likelihood value below which model runs are discarded (e.g. Shen et al., 2012) or as a percentage of the best simulations that are retained (Lamb et al. 1998; Sreelash et al. 2012). The likelihood function also must increase monotonically with increasing goodness-of-fit (Beven 2012).

During the GLUE procedure, test parameter sets are generated by sampling the prior distribution of each parameter (Figure 6a & b). Next, the model error is calculated using the objective function for each parameter set (Figure 6c). Any parameter sets associated with non-behavioural model results are discarded and the likelihood values are calculated (Figure 6d) for the remaining parameter sets. Finally, uncertainty ranges can be determined.

The probabilistic nature of GLUE allows for the calculation of uncertainty ranges. Uncertainty ranges are computed separately for each depth and time. To do this, first the simulated value of the variable of interest (temperature, for instance) is calculated at some depth and time for all behavioural parameter sets.

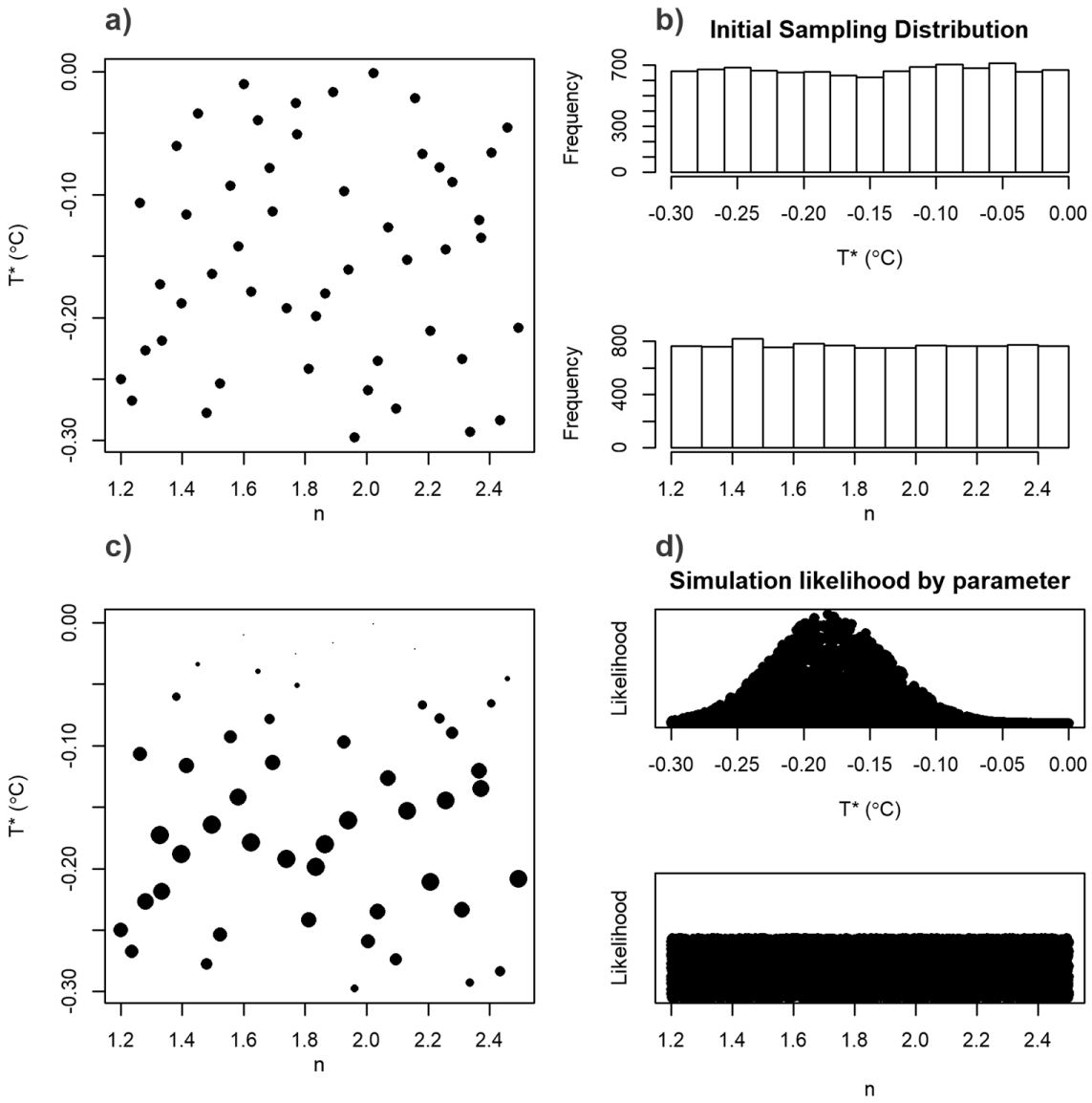

Figure 6*: Idealized description of GLUE procedure for two unknown parameters: T* and n. a) First, trial parameter sets are generated by randomly sampling parameter values from their prior distributions. b) In this case, prior distributions are uniform, but this choice must be made by the researcher ahead of time. c) The model error for each parameter set is determined. Larger circles correspond to lower model error and higher likelihood. d) Finally, the likelihood of each parameter set is normalized and can be compared to parameter values. Here, model likelihood does not change for different values of n suggesting that the model is not sensitive to the n parameter.*

Second, the temperature values and their associated likelihoods are ordered by the temperature value (Figure 7a) and then the cumulative likelihoods are calculated. This produces a distribution of temperature values at that particular depth and time (Figure 7b). Finally, the temperatures associated with the 5 and 95% cumulative likelihoods are used as the upper and lower bounds for that time step. Quantile bounds of 5 and 95% most commonly used to plot the range of possible outcomes in GLUE (e.g. Beven & Binley, 1992; Hansson & Lundin, 2006; Uhlenbrook & Sieber, 2005). This process is iterated over every time step and repeated for all depths at which uncertainty bounds are desired. Note that the parameter sets associated with the upper and lower uncertainty limit are not necessarily the same between subsequent time steps.

2.9.1 Choice of objective function

An objective function relates a parameter set to a quantitative measure of model error. Choosing an objective function is of critical importance. The choice of an objective function to compare observed to modeled data is nontrivial. Common objective functions include: RMSE, coefficient of determination, Nash-Sutcliffe efficiency, R^2, mean error and Kling-Gupta efficiency (Romanowicz et al. 2013; Gupta et al. 2009; Marmy et al. 2016). For instance, in hydrological models objective functions that consider the squares of errors tend to weight periods of peak flow more heavily than periods of low flow (Moussa & Chahinian 2009). More recent model calibration uses multi-objective calibration where two or more objective functions are used (Moussa & Chahinian 2009). Others highlight the necessity of selecting a specific objective function to satisfy site- or study-specific criteria

a)

Parameter set	Value	Likelihood	Cumulative Likelihood
$\mathcal{C}_{253}$	20.1	.004	0.004
$\mathcal{C}_{42}$	22.4	.061	0.065
$\mathcal{C}_{2003}$	23.6	.001	0.066
…	…	…	…
$\mathcal{C}_{968}$	58.9	.033	0.967
$\mathcal{C}_{522}$	60	0.007	1

b)

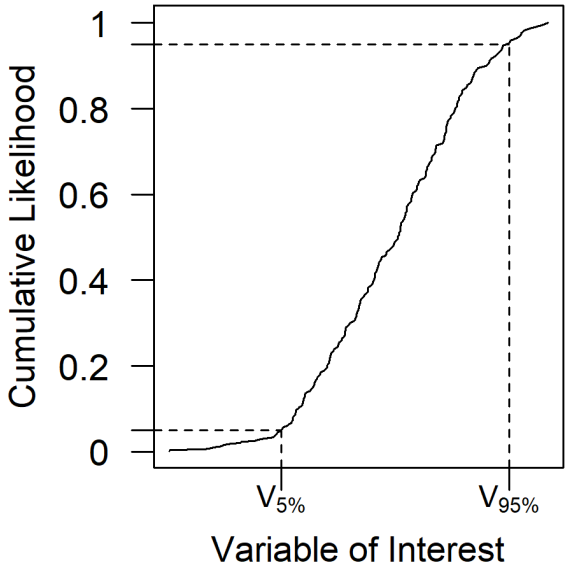

***Figure 7**: Determination of GLUE 5 – 95% quantile uncertainty bounds for a single depth and timestep. a) First, the likelihoods are ordered by the value of the variable of interest and cumulative likelihoods are calculated. b) Second, the values associated with the 5 and 95% cumulative likelihood ($V_{5\%}$ and $V_{95\%}$ in the figure) are used as the bounds for the variable of interest. In this thesis, the variable of interest is either temperature, total energy content, or liquid water content.*

(Garcia et al. 2017). Estimated parameter values are ultimately dependent on the choice of objective function.

In permafrost calibration studies, the RMSE error between modeled and measured temperatures over all sensor depths is commonly used as an objective function (Nicolsky et al. 2009; Nicolsky et al. 2007; Atchley et al. 2015; Harp et al. 2016). However, Marmy et al. (2016) use a combination of mean error, and R^2 to balance the calibration of temperature and latent heat effects. They also suggest incorporating some measure of total energy content into the calibration step in the future. Objective functions for permafrost models differ from hydrological models insofar as they will almost always have multiple observations corresponding to different depths, whereas hydrological models often calibrate to a single hydrograph.

It is important to make the distinction between calibrating a model to reproduce observations and project future conditions (e.g. Marmy et al., 2016), or fitting a model with the expectation that the parameter estimates accurately reflect local soil characteristics and can be used in lieu of measurements (e.g. Bateni et al., 2012b). In the first case, model parameters are seen to be somewhat divorced from the physical properties they represent, or the scale at which the parameter is represented in the model may not correspond to the scale at which meaningful measurements can take place (Fiddes & Gruber 2014; Gubler et al. 2011). Successful reproduction of an observed variable in the model could therefore be 'right for the wrong reason' in the sense that the parameters used do not depict reality but do give a good model fit. In the latter case, there is an acknowledgement that

although estimated parameters may not be exactly verifiable by measurement (because of scaling issues for instance), they are somehow indicative of site condition.

2.10 Summary of problem

There is an interest in monitoring permafrost as it changes. Of particular relevance is the change in liquid water content because of its consequences for ecology, the integrity of engineered structures or infrastructure, and the change in total heat content because of the implications for global heat partitioning between the atmosphere, oceans, and continental crust. For a given study site, we may be limited in the amount of information we are able to collect; however, for many monitoring boreholes, temperature time series data are available. Unfortunately, temperature is of reduced value for inferring changes to liquid water content near 0°C.

Depending on the extent of information about a site, different techniques are available to estimate the soil parameters necessary to calculate changes in liquid water or heat. There are uncertainties and challenges associated with each of these techniques. Even if measurements of SFCCs are taken, there are errors associated with the measurement of water content and in fitting a functional relationship to the data.

If no direct measurements of SFCC are available but soil texture information is available, it is possible to use pedotransfer functions to estimate water retention parameters and from those, infer SFCCs. Alternately, if site information is detailed enough, a local, gradient descent algorithm coupled with a heuristic approach can be used (Nicolsky et al. 2007). On the other hand, by using a global search algorithm, less detailed site information is necessary because the result is not sensitive to the initial approximation. Furthermore,

based on hydrological literature, there is a dependence of parameter estimates on the choice of objective function. This dependence has not been investigated for fitting permafrost models. Finally, many studies either deliberately avoid treating estimated parameters as representative of their physical counterparts (Marmy et al. 2016) or make the implicit assumption that, because the parameters are able to recreate observations, they are good surrogates for measurements (Bateni et al. 2012). Given the value of soil parameters for permafrost monitoring, a more thorough investigation is needed to determine whether calibrated estimates of the SFCC are accurate and, therefore able to inform accurate derivation of value-added information such as water content.

CHAPTER 3: METHODS AND DATA

3.1 Introduction

This thesis used GLUE and DE to estimate soil parameters by comparing GEOtop model output to observed or synthetic temperature time series. The parameters estimated are listed in Table 1 with a short description of each one. The residual water content (θ_r) was not estimated and was set to a fixed value because early experimentation suggested that estimating θ_r and θ_s simultaneously was significantly more difficult. This is discussed further in Chapter 5.

In the following chapter, I introduce the temperature time series used as observational data and describe the set-up of the permafrost model GEOtop. Then, I provide details of how the parameters were estimated. Finally, I describe the individual experiments that were performed and how they relate to the objectives of the thesis.

3.2 Data

Both synthetic and measured temperature time series were used in this thesis. Synthetic data were used to determine how well parameters can be estimated using GLUE and DE in a best-case scenario. There are several advantages to using synthetic data for this: first, since the true parameter values are known ahead of time, the accuracy of estimates can be evaluated; second, synthetic data make it possible to isolate the effect of errors such as sensor bias or inaccurate field measurements. Parameter estimation was repeated using measured data in order to identify some of the challenges associated with applying these methods to field sites.

***Table 1**: Parameters that were estimated in this thesis.*

Parameter	Description
θ_s [m^3 m^{-3}]	Saturation water content: what fraction of the soil volume is occupied by water when the soil is fully saturated
α [mm^{-1}]	Water retention parameter described by van Genuchten (1980): affects the curvature of the SFCC.
n	Water retention parameter described by van Genuchten (1980): affects the curvature of the SFCC.
T* [°C]	Freezing point temperature: the temperature at which water in soil pores begins to freeze.
c_{sp} [J m^{-3} K^{-1}]	Heat capacity of the soil particles: how much the temperature of soil particles changes in response to the addition of energy.
λ_{sp} [W m^{-1} K^{-1}]	Thermal conductivity of the soil particles: how well the soil particles transmit heat in response to an applied temperature gradient.

3.2.1 Synthetic data

A synthetic ground temperature dataset was generated by using 3 years of near-surface ground temperature data (Figure 8a) to run the GEOtop model with specified ground properties (Table 2). The soil properties for α, n, θ_s and θ_r were based on a ice-bonded silt soil (Dall'Amico et al. 2011). Upper and lower boundary temperatures were specified at 0 m and 15 m to drive the model (Figure 8a). These temperature data were obtained from the PERMOS database for the Murtèl-Corvatsch site at 0.55 and 15.57 m (Vonder Mühll et al. 2008); however, the intention was not to simulate this particular site but rather to obtain a realistic-looking upper boundary for the synthetic data. Temperature output was saved at 0 m, 50 cm, 1 m and every metre thereafter to 15 m.

To examine the evolution of liquid water and energy content over a long period of time, a secondary, 100-year synthetic dataset was created using the same soil properties but different forcing data. At the upper (0 m) boundary, the GEOtop model was driven by an annually-oscillating temperature with an amplitude of 26 °C and a mean temperature of -0.5 °C. The lower (15 m) boundary was set to a fixed temperature of -3.3 °C. After 50 years, a warming signal of 0.08 °C yr^{-1} was added to the surface temperature to simulate climatic warming and a much smaller warming signal of 0.005 °C yr^{-1} was added to the lowest temperature to account for warming from the geothermal flux. The model was then run for another 100 years (Figure 8b)

3.2.2 Perturbed synthetic data

To investigate the effect of imperfect data on results, two sets of synthetic data were created with added error. To simulate the imperfect measurement of sensor depths,

***Table 2**: Soil properties specified for synthetic dataset.*

α [mm^{-1}]	n	θ_s [$m^3\ m^{-3}$]	θ_r [$m^3\ m^{-3}$]	c_{sp} [$J\ m^{-3}\ K^{-1}$]	λ [$W\ m^{-1}\ K^{-1}$]	T* [°C]
1.49×10^{-3}	1.67	0.49	0.05	2.5×10^6	2.5	0

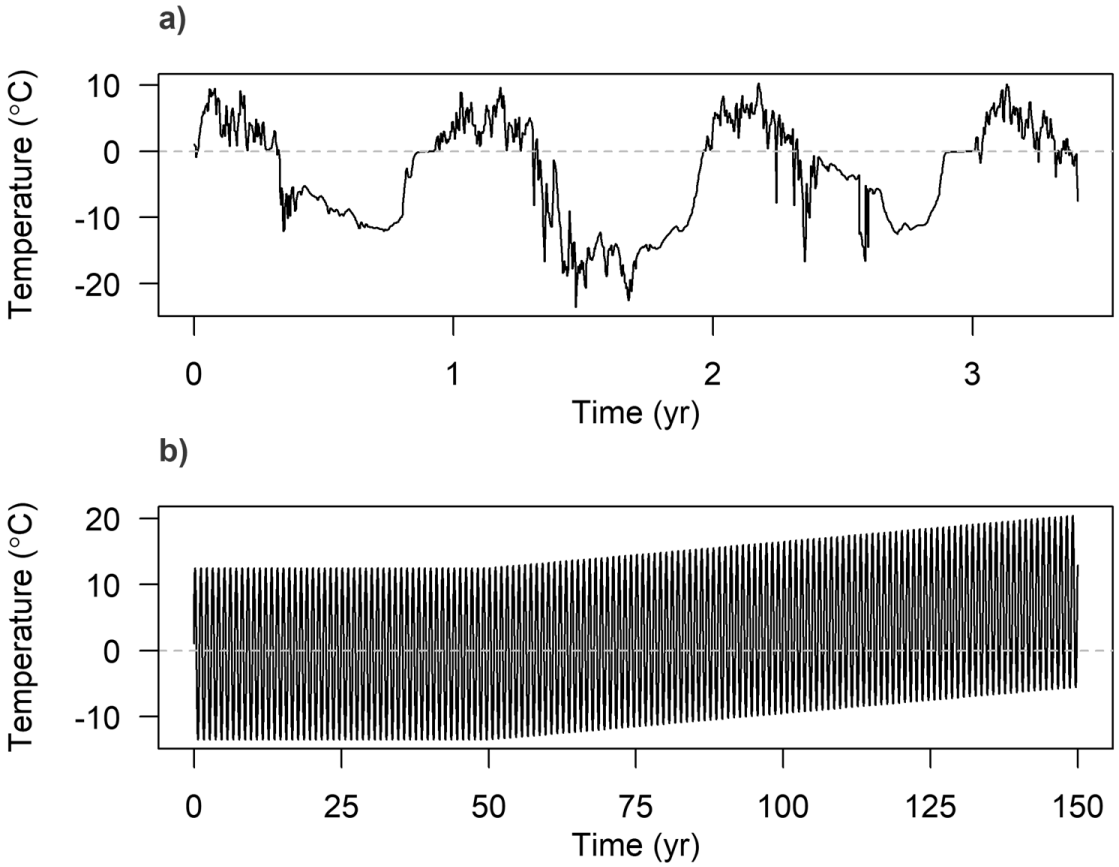

Figure 8: Surface boundary conditions used to generate synthetic ground temperature datasets.

uniformly distributed random error of ±2.5 cm was added to each of the depths in the synthetic dataset to create a perturbed dataset called D-BIAS. A second perturbed dataset, T-BIAS, was created by adding a uniformly distributed random constant error of ±0.05 °C to each sensor to simulate calibration bias. For each type of error (depth or temperature), a total of 20 different perturbed datasets were created, each with a different realization of the random noise. To distinguish the original synthetic data from the biased datasets, the original was identified as PURE.

3.2.3 Measured data

Ground temperature and soil textural data from two sites along the Mackenzie Valley Corridor were obtained from the Geological Survey of Canada (GSC) (Smith et al. 2009). Detailed drill logs are available in the original publication. Two sites—Gibson Lake and Chick Lake—were selected based on the homogeneity of the soil profile, ground temperatures close to 0 °C, and the presence of soil textural information from drill samples. Thermistor depths for each of the two sites are listed in Table 3.

The Chick Lake CL01 borehole samples comprise primarily clay with a high silt content and only trace amounts of sand. Visible ice ranges from 25 to 30% down to 9 m and contains intermittent ice-rich zones with up to 60% visible ice below 9 m. Only a subset of the full data period was used. The first four months of data were removed in order to mitigate the effects of thermal disturbance resulting from the heat generated while drilling the borehole (Lachenbruch et al. 1982). The four-month cut-off was chosen by visual inspection of the data for anomalous warming periods.

***Table 3**: Metadata for permafrost monitoring sites used in this thesis.*

Site	Thermistor Depths (m)	Data Period	Collection Frequency
Chick Lake (CL01)	0.5, 1, 1.5, 2, 4, 7, 12, 20	04/2007 – 05/2011	8-hour
Gibson Lake (GL01)	0.5, 1, 1.5, 2, 4, 7, 12, 20	04/2007 – 09/2011	8-hour
Janssonhaugen (P11)	0.2, 0.4, 0.8, 1.2, 1.6, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 7.0, 10.0, 13.0, 15.0	04/1999 – 02/2017	6-hour

The Gibson Lake GL01 borehole is also a silt-rich clay but with higher sand content than CL01. Visible ice ranges between 0 and 15% for the entire soil profile. As with Chick Lake, the first four months of data were not included in order to mitigate the effects of thermal disturbance.

For both the Gibson Lake and Chick Lake data, not all thermistors were used. The thermistor at 7 m depth was used as a lower boundary for the Gibson Lake, and the thermistor at 12 m was used as a lower boundary for Chick Lake. Below these depths, there was very little temperature variation and there was considerably more noise in the data. The total temperature range below the cut-off depths over the period of data availability was less than 0.15 °C for Chick Lake and less than 0.02 °C for Gibson Lake. Several discrete, temperature discontinuities of up to 0.01 °C and periods of noise with a range of up to 0.01 °C were present below the cut-off depths.

Ground temperature data were also obtained from the Janssonhaugen P11 borehole in Svalbard. The full details of the installation and the site are documented in Isaksen et al. (2000). The borehole was instrumented to a depth of 15 m. Thermistor depths are listed in Table 3. All sensors were used in the analysis. The borehole was drilled into sandstone with measured porosity between 20 and 25% and a bulk specific heat capacity of 800 J kg^{-1} K^{-1}. The mean in-situ bulk thermal diffusivity of the sandstone was calculated to be $1.1x10^{-6}$ m^2 s^{-1} (Isaksen et al. 2000).

3.3 The model: GEOtop

In this thesis, the GEOtop model was used to test how well trial parameter sets were able to reproduce observations. To run the model, soil properties were specified for a 1D soil

column, and the initial temperature profile of the model soil column was taken from the first measured temperature profile in the observations. By doing this, both the observations and modeled data began from the same initial state.

The model was driven by the observed or synthetic temperatures at the upper- and lowermost thermistors in the observations; the temperatures at the top and bottom of the soil column were obtained from the observations at corresponding depths. At each hourly time step, the temperature profile of the soil column was calculated based on the temperature gradients and the properties of the soil.

3.3.1 Soil column discretization

In order to run GEOtop, the soil column was discretized. To illustrate how this was done, consider a column of soil and two measurable properties of interest (Figure 9) where the soil properties vary continuously over the depth of the soil column (Figure 9a). To aid in analysis, it is typical to classify the soil column into conceptual layers (Figure 9b). This is referred to here as the soil abstraction. Each layer of the soil abstraction is homogenous, therefore some of the detail and variability of the soil column is lost in the first step of approximation. Of course, the number of layers and position of boundaries in the soil abstraction are defined subjectively, and a higher-resolution model could be made that captures more detail. However, as more layers are added to the soil abstraction, more parameter sets are required to define the model.

Numerical models require that both time and space be discretized into steps of finite duration or distance. Coarser discretization corresponds to faster model execution but a

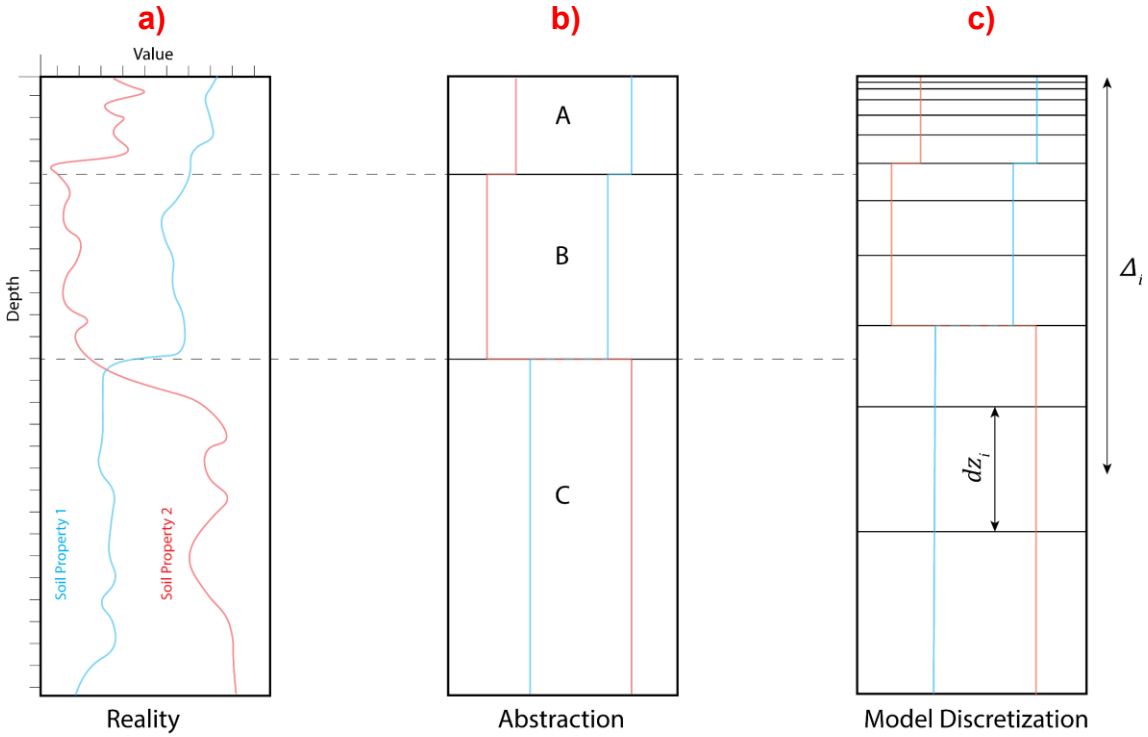

***Figure 9**: Generating a discretized numerical approximation of a soil column. The red and blue lines represent two arbitrary soil properties. The leftmost representation is the closest to reality. The middle and rightmost representation are increasingly abstracted from reality. Boundaries between discretized model cells do not necessary align to boundaries in the simplified conceptualization of the soil. In the model discretization, cell thickness (dz) increases with depth. Δ is the depth to a model node (centre of a discretized cell).*

degradation of model accuracy. A compromise can be made by using finely-discretized cells near locations where change takes place rapidly and using coarser cells at locations where changes are slower. For permafrost models, this corresponds to thin cells near the surface and coarser cells at greater depth based on an assumed dominance of conductive heat transfer behaviour. At the surface, large temperature gradients can develop over relatively short time scales because the rate of air temperature fluctuation is short relative to the timescale of heat transfer in and out of the ground.

To create a model discretization for GEOtop with 200 grid cells, a sequence of monotonically increasing cell thicknesses dz_i was created such that the total thickness of all cells was equal to the thickness of the desired soil profile Z and the cell thickness grew according to a power law. Mathematically, a constant k was found to satisfy the following equation with a small error tolerance $\varepsilon \ll 1$:

$$Z = \left(\sum_{i=1}^{200} dz_i\right) \pm \varepsilon \quad dz_i = i^k dz_1$$

The depth to the centre of any soil discretization cell, also referred to as a node, was given by:

$$\Delta_i = \sum_{j=1}^{i-1} dz_j + \frac{dz_i}{2}$$

For this thesis, a one-layer soil column model was used, which simplified the soil column to be homogenous with respect to soil properties. This corresponds to a single layer in Figure 9b.

3.3.2 Boundary values and initial conditions

Once the soil column has been discretized, the model state at the start of the simulation (initial condition) must be specified. Some information must also be provided about the conditions at the upper and lower limits of the soil column (boundary values). In this thesis, temperatures were specified at the upper and lower boundaries: the shallowest and deepest temperature values of the modeled soil column were taken directly from the corresponding thermistors in the observation dataset (Dirichlet boundary conditions). An initial model state was determined by interpolating the first temperature profile of the observation dataset (Figure 10).

To set the initial condition for the model, the first soil profile in the observations was interpolated to produce a function $\tau(z)$ that allowed temperature values to be approximated at any depth. The initial interpolated temperature (at t = 0) for each of the soil discretization nodes was calculated as:

$$\widehat{T}(\Delta_i, 0) = \tau(\Delta_i)$$

During each time step, GEOtop interpolated between adjacent soil discretization nodes to produce output temperatures at any depth required. The model did not extrapolate past terminal nodes, so any temperature values obtained from output depths below the midpoint of the deepest cell used the temperature value of the nearest terminal cell (Figure 10).

3.4 Parameter estimation

A schematic of the parameter-fitting procedure is shown in Figure 11. In brief, the steps that were taken are as follows: (1) a trial parameter set was generated, (2) the upper- and lowermost observations were used to drive the GEOtop model with the trial parameter set,

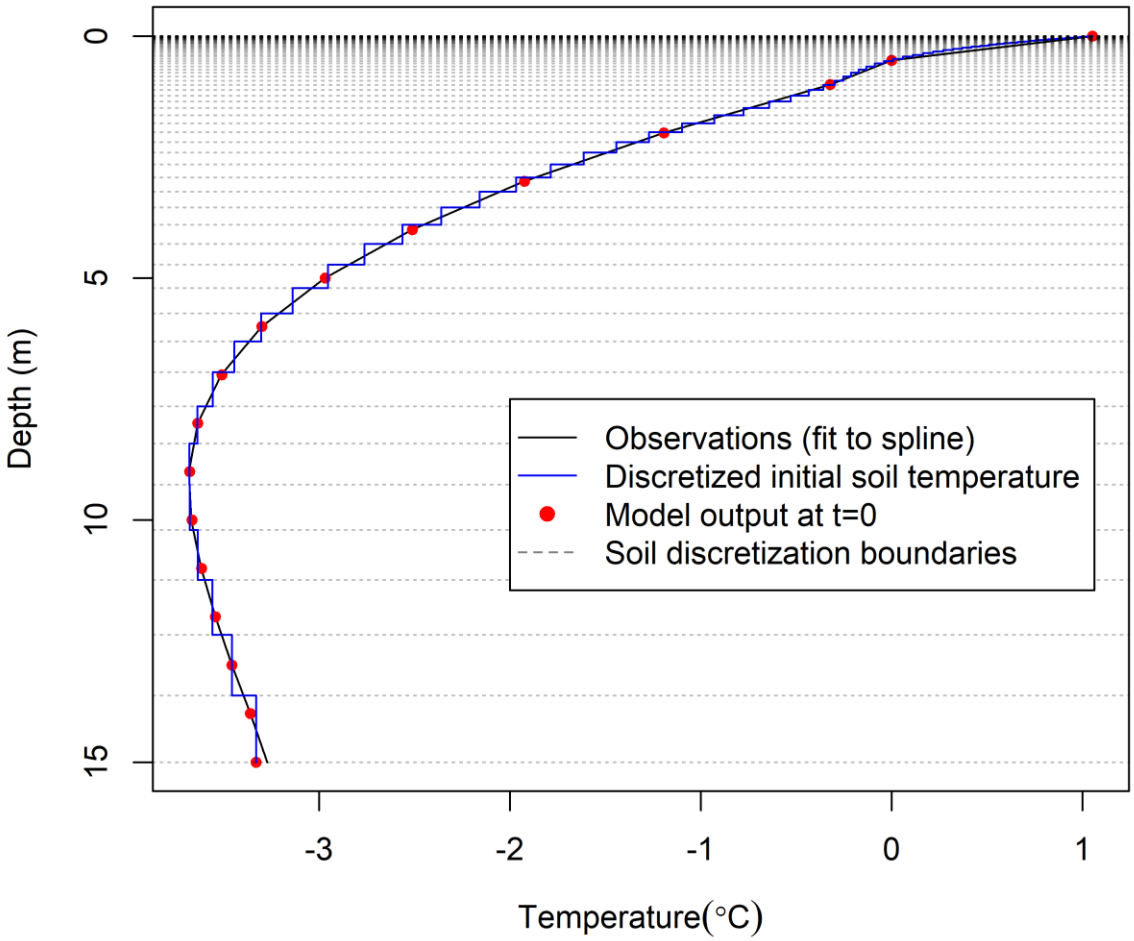

***Figure 10**: The model initial condition is approximated from the first temperature profile in the observations. The deepest temperature output at 15 m is below the midpoint of the lowest discretized cell and therefore takes on the value of the lowest cell.*

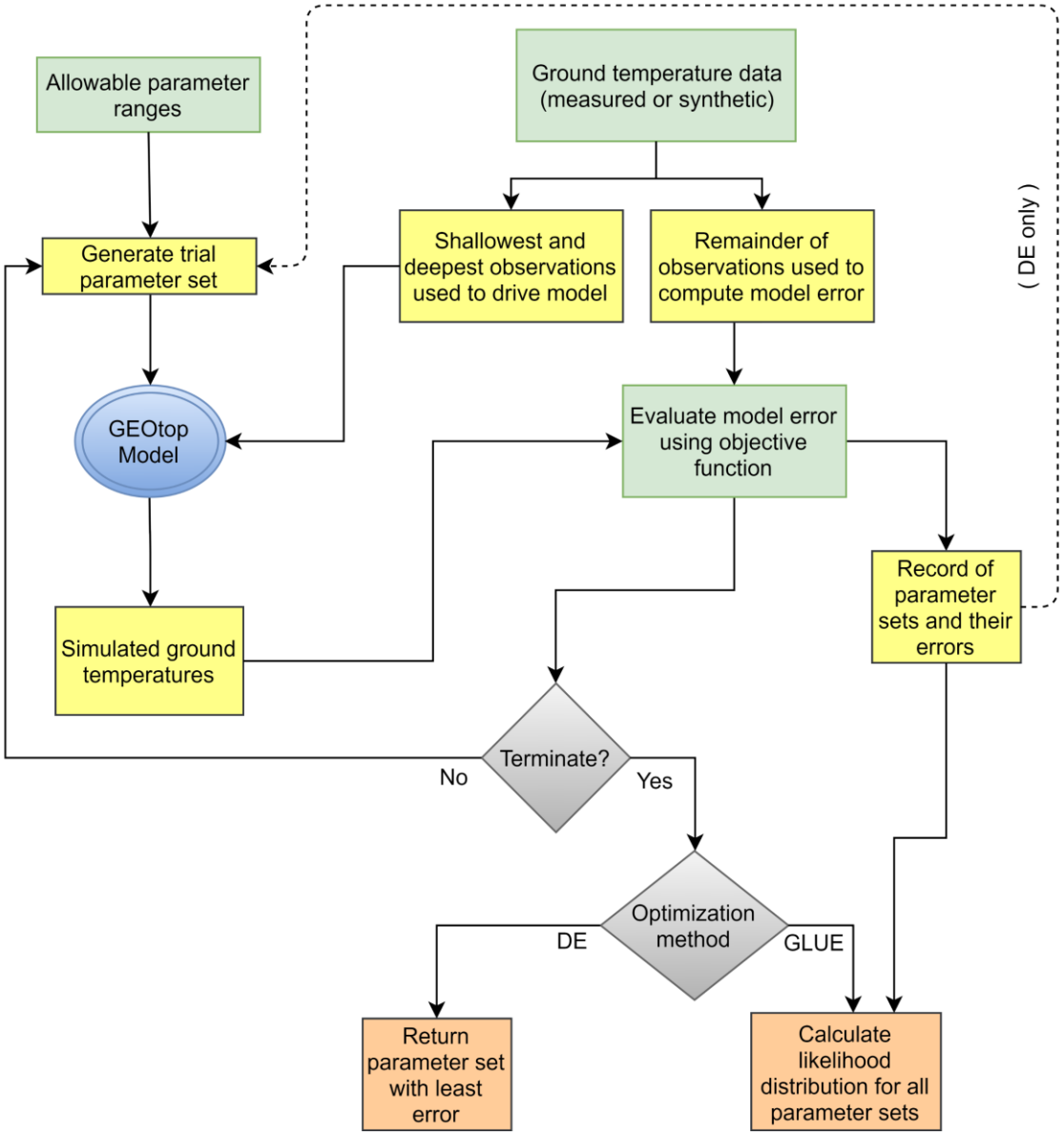

***Figure 11**: Schematic of the parameter-fitting procedure. Green boxes represent the parts of the procedure that are changed between different experiments. Yellow boxes represent intermediate steps, grey diamonds represent decision points, and orange boxes are the output parameter set (DE) or likelihood distributions (GLUE). One of the ways in which the two methods differ in that DE generates parameter sets by modifying previously tested candidates. This is indicated by the dashed line.*

(3) model accuracy was evaluated with the remaining observation depths using one of two objective functions (discussed in more detail below), (4) the process was iterated until either the population of parameter sets became sufficiently similar to one another (DE) or a pre-determined number of parameter sets had been evaluated (GLUE). Below, I outline the allowable ranges of values for each of the parameters and describe the formulation of the two objective functions used to evaluate model performance. The details describing how each of the methods were implemented in this thesis are also presented.

3.4.1 Parameter ranges

When selecting trial parameter sets, the values of individual parameters were limited to ranges that would be expected in permafrost environments (Table 4). For experiments with synthetic data, this set of parameter ranges represented a site for which very little prior information was available—all that was known about the site was that the parameters must be physically realistic. A second set of parameter ranges was also used (Table 5) which simulated a hypothetical scenario in which the parameter ranges were narrowed ahead of time based on some prior site information. For many permafrost sites, only temperature measurements are available while for others, soil texture or measured soil properties may have been recorded. By using two different parameter sets, it was possible to test, using synthetic data, the effect of more or less prior site information on the results.

Parameter ranges for the Chick Lake and Gibson Lake sites were obtained using the software program ROSETTA v1.2 (Schaap et al. 2001). ROSETTA provides estimates of the mean value and standard deviation for θ_s, θ_r and the two van Genuchten parameters α and n. For each measurement depth in the Chick Lake and Gibson Lake data, the percent

***Table 4**: Tolerable parameter values for optimization. Ranges were narrowed for some experiments. Parenthetical references indicate either the literature reference from which the limit was obtained, or the earth material associated with the limiting value.*

Property	**Minimum**	**Maximum**
λ_{sp} [W m^{-1} K^{-1}]	1.0	8.0 (quartz)
c_{sp} [J m^{-3} K^{-1}]	8.5×10^5	4×10^6 (water)
θ_s [m^3 m^{-3}]	0.10 (Dysli 2003)	0.75 (Dysli 2003)
α [mm^{-1}]	1×10^{-5} (silt, (Dall'Amico et al. 2011))	0.1(water is 0.4, (Dall'Amico et al. 2011))
n	1.25 (clay (Dall'Amico et al. 2011))	2.5, (water, (Dall'Amico et al. 2011))
T* [°C]	-1	0 (pure water)

***Table 5**: Parameter ranges for experiments with synthetic data which simulated some degree of prior site knowledge.*

Property	**Range**	**Property**	**Range**
λ_{sp} [W m^{-1} K^{-1}]	[2.40, 2.60]	log(α)	[-4, -2]
c_{sp} [J m^{-3} K^{-1}]	[2.4x10^6, 2.6x10^6]	n	[1.55, 1.75]
θ_s [m^3 m^{-3}]	[0.39, 0.59]	T* [°C]	[-0.3, 0]

weight of the gravel, sand, silt and clay size fraction were determined from the nearest drill sample that was located within the same textural unit (e.g. 'silty clay', 'sand') in the drill log (Smith et al. 2009). Sand and gravel size fractions were added to yield a fraction representing coarse material which was entered into ROSETTA as 'sand'. The value of θ_r for the soil profile was set to the average of the estimated values from ROSETTA. Averages were based on all soil samples collected between the upper and lower thermistors.

3.4.2 Defining the objective function

As discussed above, an objective function J relates a parameter set $\mathcal{C}$ to a quantitative measure of model error. As was discussed in Chapter 2, the choice of objective function can affect the results of parameter estimation.

In this thesis, two different objective functions were used to evaluate their influence on parameter estimation results for the GEOtop permafrost model. Both were based on RMSE error (following Nicolsky et al., 2007) but each with different weighting schemes. In calculating model error, only those model results with a corresponding T(z, t) measurement in the observational dataset are used to calculate error. In other words, if data were missing from the observations for a range of dates, those dates are not considered when calculating model error.

The first objective function J_1 was an unweighted RMSE over all times and depths defined as:

$$J_1(\mathcal{C}) = \sqrt{\frac{1}{N * M}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(T(z_i, t_j) - \hat{T}(z_i, t_j, \mathcal{C})\right)^2}$$

Where N is the total number of thermistors, M is the total number of observations in time and $\hat{T}$ is the modeled temperature. This was referred to as the temperature-based RMSE.

A modified RMSE objective function J_2 was created by transforming the observed and modeled temperatures into a total thermal energy content U and then calculating the RMSE error between the energy contents. This was done to take into consideration the effect of latent heat, H_l, associated with soil freezing or thawing. This is important because an error of ±0.3°C near 0°C may correspond to a large difference in total energy, whereas the same temperature error at -10 °C may represent an order of magnitude less energy. The total thermal energy was calculated as the sum of the sensible and latent heat contents of the soil volume:

$$U = H_s + H_l$$

The sensible heat H_s was calculated from the contribution of sensible heat from each soil component (ice, water, air, soil particles) using the volumetric fraction θ_i and heat capacity c_i for each component. By convention, values are calculated with respect to a reference value associated with a temperature of -273.15 °C:

$$H_s = (T + 273.15) * \left(c_i\theta_i + c_w\theta_w + c_a\theta_a + c_{sp}\theta_{sp}\right)$$

The latent heat was calculated based on the instantaneous volumetric unfrozen water content and the volumetric latent heat of fusion for water, L_f:

$$H_l = L_f\theta_w$$

These values were used to determine the total energy difference between observed and simulated soil profiles at each observation depth. The trial parameter set $\mathcal{C}$ was used to calculate the total energy content for both the modeled and the observed temperatures:

$$J_2(\mathcal{C}) = \sqrt{\frac{1}{N * M}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(U\left(T\left(z_i, t_j\right), \mathcal{C}\right) - U\left(\hat{T}\left(z_i, t_j, \mathcal{C}\right), \mathcal{C}\right)\right)^2}$$

This formulation of the objective function was referred to as the heat-based RMSE.

3.4.3 Generalized likelihood uncertainty estimation (GLUE)

As mentioned in Chapter 2, GLUE uses a statistical distribution of prior parameter estimates and a likelihood function to produce a new distribution of parameter estimates based on how well they fit a set of observations. Put another way, given a set of prior estimates for parameters of interest, GLUE returns the relative likelihood for various combinations of parameters. These parameter combinations are then used to make predictions about the monitoring site.

In this thesis, prior parameter estimates were created for each experiment by creating uniform distributions with maximum and minimum values equal to the parameter range. As discussed above, parameter ranges were based on either a hypothetical site with limited (Table 4) or detailed knowledge (Table 5) for synthetic data or, for measured data, on estimates from soil textural data for the measured data.

Sets of trial parameter sets were generated by Latin Hypercube sampling (LHS). This sampling strategy subdivides the range of each variable into a number of sub-ranges, and ensures that one sample is taken from each sub-range. LHS ensures that the parameter space is sampled in a way to maximize coverage over all variables and reduce clustering of samples (Mckay et al. 1979; Manache & Melching 2004). In each experiment, 10^5 LHS parameter combinations were generated using the SPOTPY python package (Houska et al.

2015). The GLUE procedure also required that a likelihood measure be chosen. In this thesis, the likelihood L was defined as the inverse of model RMSE error, normalized so that the sum of all likelihoods is equal to 1:

$$L(\mathcal{C}_i) = \frac{1}{\sum J(\mathcal{C}_j)^{-1}} J(\mathcal{C}_i)^{-1}$$

Where J was the RMSE objective function based either on temperature or energy content and $\mathcal{C}_i$ was a set of model parameters. For the purposes of separating simulations into behavioural and non-behavioural, the best 1% of parameter sets were retained as behavioural. This is an entirely subjective choice but follows Lamb et al. (1998) who used a cut-off of 10% and only tested 10^4 parameter sets.

3.4.1 Understanding GLUE results

The output of GLUE is a collection of parameter sets, each of which has a measure of likelihood corresponding to the accuracy of modeled ground temperatures using those parameters. These values are typically plotted as cumulative likelihood distributions and referred to as the *posterior* distributions. The trial parameter sets, each used in a model run and then assessed for its accuracy in representing synthetic or observed ground temperatures, are sampled from a distribution that is defined ahead of time—the *prior* distribution.

Cumulative distribution functions are suitable for the combined display of several distributions, but can be less intuitive to interpret than histograms. As an aid to visual interpretation, the equivalencies between the two are presented in Figure 12. Briefly, a cumulative distribution function describes the probability of obtaining a value less than

some arbitrary value, X (Dalgaard 2008), or, in the case of GLUE, the likelihood that the correct parameter value is less than a specified value.

To relate these distributions back to the objective of estimating ground properties, the degree to which each model parameter can be estimated by GLUE can be evaluated by comparing the prior and posterior cumulative likelihood distributions. If the posterior distribution is the same as prior distribution (in this case, if both are uniform) then it suggests that model is not sensitive to that parameter and, therefore, it is not possible to estimate (Figure 12b). If the distribution is clustered to one side (Figure 12 c and d), it suggests that the best-performing parameter is found either at or outside the calibration bounds. If the total range of the parameter distribution is reduced by the GLUE procedure, then it means the estimate of the parameter value has improved over the initial estimate. Finally, distributions that are normally distributed and have a clear modal value (Figure 12a), indicate that the model is sensitive to that parameter, and that the best performing values can be reasonably identified with GLUE. With synthetic data, one can judge the accuracy of the GLUE estimate by comparing the true value to the approximate modal value of the distribution, corresponding to the steepest part of the cumulative distribution curve.

3.4.1 Differential evolution

The DE search algorithm (Storn & Price 1997) was used to test the effectiveness of global search algorithms for parameter estimation. This algorithm was obtained from the open-source SciPy package for the Python programming language (Jones et al. 2001). In all

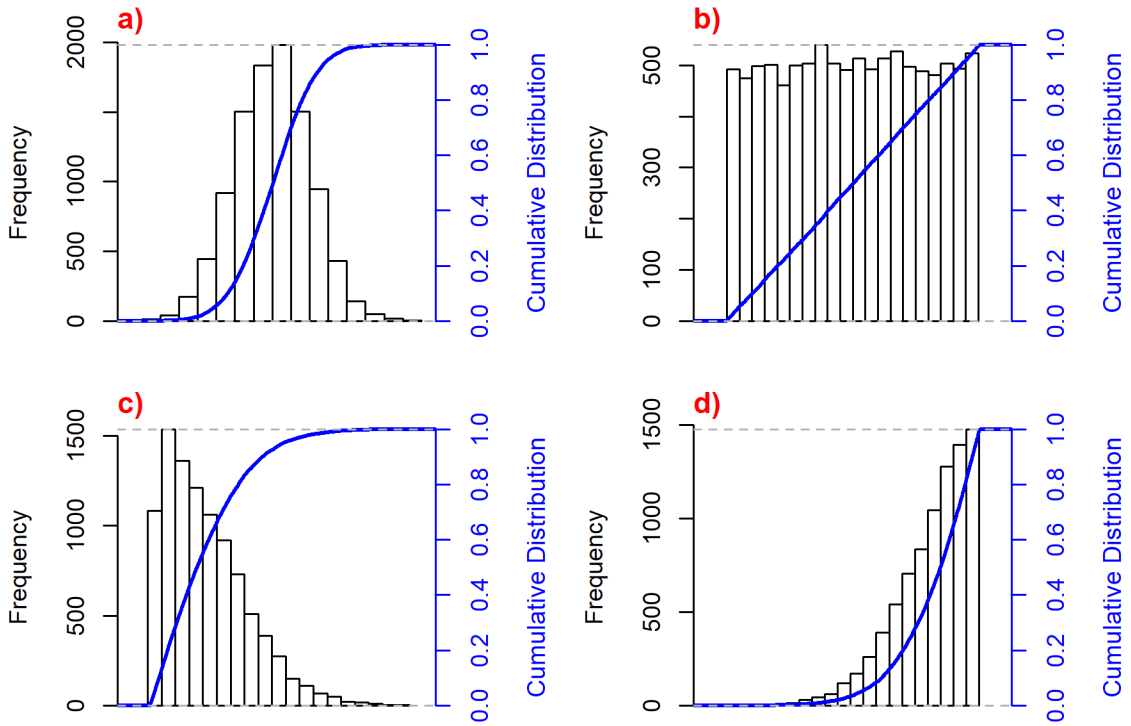

***Figure 12**: Examples of histograms (in black) and their corresponding empirical cumulative distribution curves (in blue). Left-to-right from top: a) A normal distribution, b) a uniform distribution, c) a positively-skewed distribution and d) a negatively-skewed distribution. The modal value of a cumulative distribution can be identified as the value associated with the steepest part of the blue curve.*

experiments using DE, the temperature-based RMSE was used. For experiments with DE, the algorithm was allowed to select any physically meaningful parameter values (Table 4).

3.5 Experiments

3.5.1 Outline

A schematic of all the experiments conducted is presented in Figure 13. In the material that follows, the rationale and details of each experiment is presented in more detail.

3.5.2 Parameter estimation with GLUE using synthetic data

In the first experiment, the capability of the GLUE technique to estimate soil parameters was tested in an idealized scenario. The effect of different formulations of the objective function on the resulting parameter sets was also tested. An unperturbed synthetic dataset was used to determine the best-case scenario to be expected. Two sets of parameter ranges were used: one that simulated a scenario of very limited prior site knowledge (Table 4), and one that simulated a scenario in which the parameter uncertainty had been reduced, for instance by site measurement or pedotransfer functions (Table 5).The GEOtop model was run for each sample set. Both objective functions were evaluated for each model run. Cumulative likelihoods of parameter frequencies for each cut-off level were plotted for each of the RMSE cost criteria.

3.5.3 Sensitivity of GLUE to observational error

To investigate the effect of imperfect data and measurement errors on the GLUE results, the first experiment was repeated using the two perturbed datasets: T-BIAS and D-BIAS.

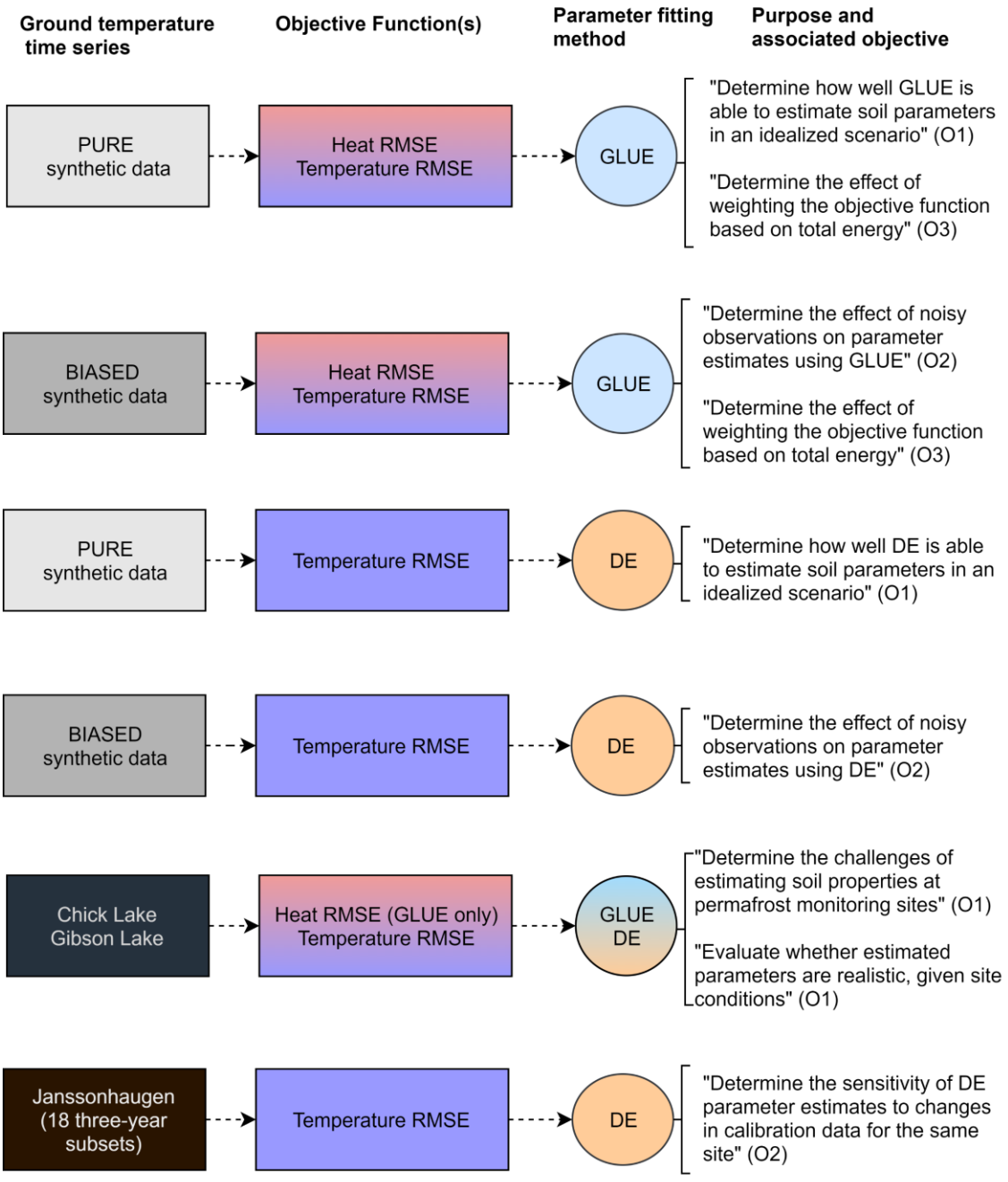

***Figure 13**: Schematic of the experiments conducted. In each experiment, soil parameters are fit using either GLUE or DE. The temperature time series and the objective function are changed in each experiment to address a part of one of the objectives.*

Because of the computing time requirements to run the GLUE analysis, only one realization of each noisy dataset was investigated and only the narrow parameter range (Table 5) was used. In a subsequent experiment using the DE algorithm (described below), all 20 different realizations of noisy data were used. Estimated parameters were used to calculate energy and liquid water content.

3.5.4 Parameter estimation with DE

The DE search algorithm was used to fit the 6 soil properties (Table 1) to the unperturbed (PURE) synthetic data as well as the perturbed T-BIAS and D-BIAS datasets. This experiment was conducted to determine whether soil parameters for monitoring sites without detailed soil property information could be estimated using the DE search algorithm (Storn & Price 1997), and what the effect of measurement errors were. Parameter ranges were based on the assumption of limited prior knowledge (Table 4).

Because DE uses a random process to create trial parameter sets, there is an inherent tendency for the final estimates to differ. If this method is to be applied to permafrost sites, it is desirable that repeated estimation produces similar results. For the unperturbed observation set, 20 optimization procedures were run in parallel with the same observation data to ensure that the inherent variability of the technique did not produce highly variable results. For the experiments with noisy synthetic data, all 20 realizations of each biased dataset were tested to investigate the effect of imperfect data on the parameter estimates. The resulting parameter estimations were plotted and compared to the true parameter values.

To visualize the range of the associated estimates for total energy or liquid water content, a process similar to that of GLUE was devised: after calculating the values of energy or liquid water content for all 20 estimated parameter sets, the maximum and minimum value of energy or liquid water was used as an uncertainty envelope. This is similar to GLUE in that the maximum or minimum value at each time step may be a result of a different combination of parameter sets. However, because the DE method does not yield any information on the relative likelihood of each result, the two methods cannot be rigorously compared.

3.5.5 Application to borehole data

To attempt to constrain the liquid water history at permafrost monitoring sites, and to investigate challenges associated with estimating soil parameters using both GLUE and DE at field sites, temperature time series from two sites—Chick Lake and Gibson Lake—were used to estimate soil characteristics. Recall that these sites were chosen because the soil stratigraphy was relatively uniform and could be reasonably modeled as a one-layer soil column.

Following Hansson & Lundin (2006), parameter ranges for the prior distributions of α, n, and θ_s were obtained using the estimates from ROSETTA (Schaap et al. 2001). However, instead of using the mean and standard deviation as reported by ROSETTA, uniform distributions were created by taking the maximum and minimum values associated with 4 standard deviations from the mean based on the values obtained from ROSETTA. Wider parameter ranges were chosen to account for possible model mis-specification,

heterogeneity in the soil layers and other possible errors. θ_r was set to a constant value equal to the mean estimate from ROSETTA for each site.

A plausible range for T* was evaluated by visual inspection of the ground temperature data. The temperature at which the zero-curtain could be identified during cooling was used to inform the lower bound. Zero curtain temperatures appeared to be consistently above -0.2 °C, however, a slightly more conservative range between -0.3 and 0° C was chosen to account for possible thermistor error or inhomogeneity between soil layers. Parameter ranges for k and c_{sp} were based on the assumption of limited prior knowledge (Table 4) because measurements were not available.

3.5.6 Sensitivity of DE to calibration period

If estimated parameters are to be used in place of field measurements, they must be representative of site conditions, and repeated estimates should yield similar results even with slightly different input data. To test if this is the case, the DE technique was applied to the Janssonhaugen borehole data.

Data were available for a longer continuous period at Janssonhaugen (18 years) than at the GSC sites. This offered a better test of the effect of using different calibration periods on the parameter estimates. A total of 14 calibration runs were performed using three-year subsets of the data to investigate whether the estimated parameters were highly sensitive to input data. The hypothesis here is that if different calibration periods result in similar values for one or more parameters, those parameters can be used reliably to calculate energy and liquid water. Alternatively, if the parameters are dissimilar between calibration periods, either site conditions (such as pore structure or water content) change

significantly, or small variations in the data produce large and unstable variation in the best-fit parameters. If this is the case, it would be inappropriate to assume that the parameters represent the site conditions, and their suitability to predict energy and liquid water contents would be questionable. The Janssonhaugen site also has several measured or independently calculated parameter values. Where available, these were compared to the estimated parameters.

CHAPTER 4: RESULTS

4.1 Parameter estimation with GLUE using synthetic data

In what follows, the posterior GLUE distributions are presented for the experiments described in Chapter 3, and the parameter uncertainties are used to calculate uncertainty bounds for ground liquid water and energy content.

4.1.1 Parameter estimation with GLUE: wide parameter bounds

The posterior cumulative likelihoods for GLUE parameter estimation using wide parameter boundaries are shown in Figures 14 through 16.

Posterior likelihoods for c_{sp} are skew-uniform across the allowable parameter range. The temperature objective function places higher likelihoods on values greater than the true value, while the opposite is true of the heat-based objective function, which places slightly higher likelihoods on values lower than the true value. The posterior likelihood distribution for λ_{sp} is positive skew-normal. For the temperature RMSE, the modal value is a close approximation of the true value, whereas for the heat-based RMSE, the modal value underestimates the true value. The posterior likelihood distribution for κ is positive skew-normal with a modal value close to the correct one.

Parameter distributions for α are narrowed slightly for the temperature-based RMSE, with $\log(\alpha)$ values between -5 and -4 being very uncommon. The heat-based RMSE is positive skew-uniform with the modal values being in the -5 to -4 range. Parameter ranges for n are approximately uniform for both objective functions. No reduction in parameter range is present. For the temperature-based RMSE, the θ_s distribution is approximately normal and centred on the true value. For the heat-based RMSE, the

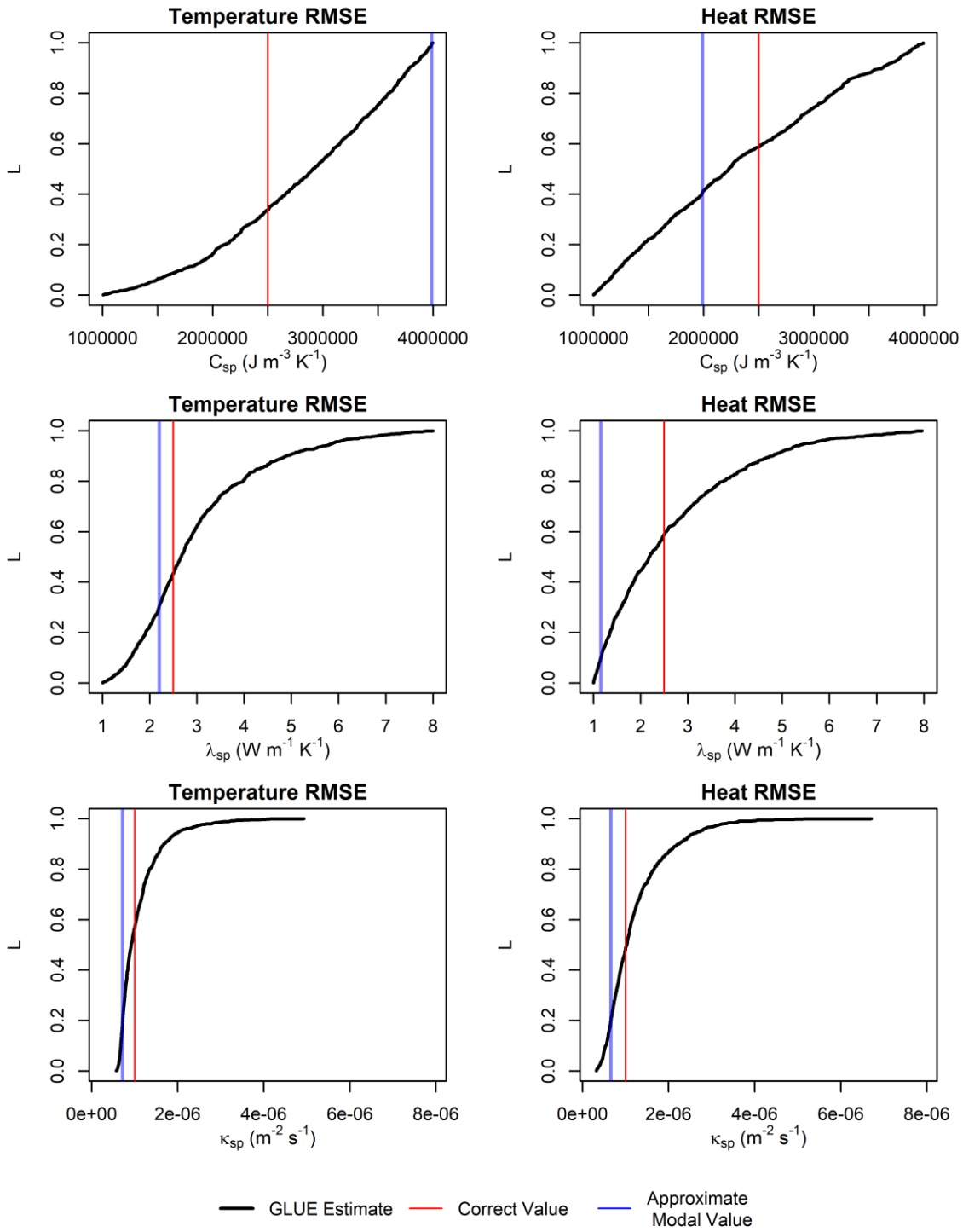

***Figure 14**: GLUE cumulative parameter likelihoods for c, λ and κ based on wide parameter bounds and unperturbed synthetic data. Vertical lines signify the approximate modal value of distribution as determined from the steepest part of the CDF. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

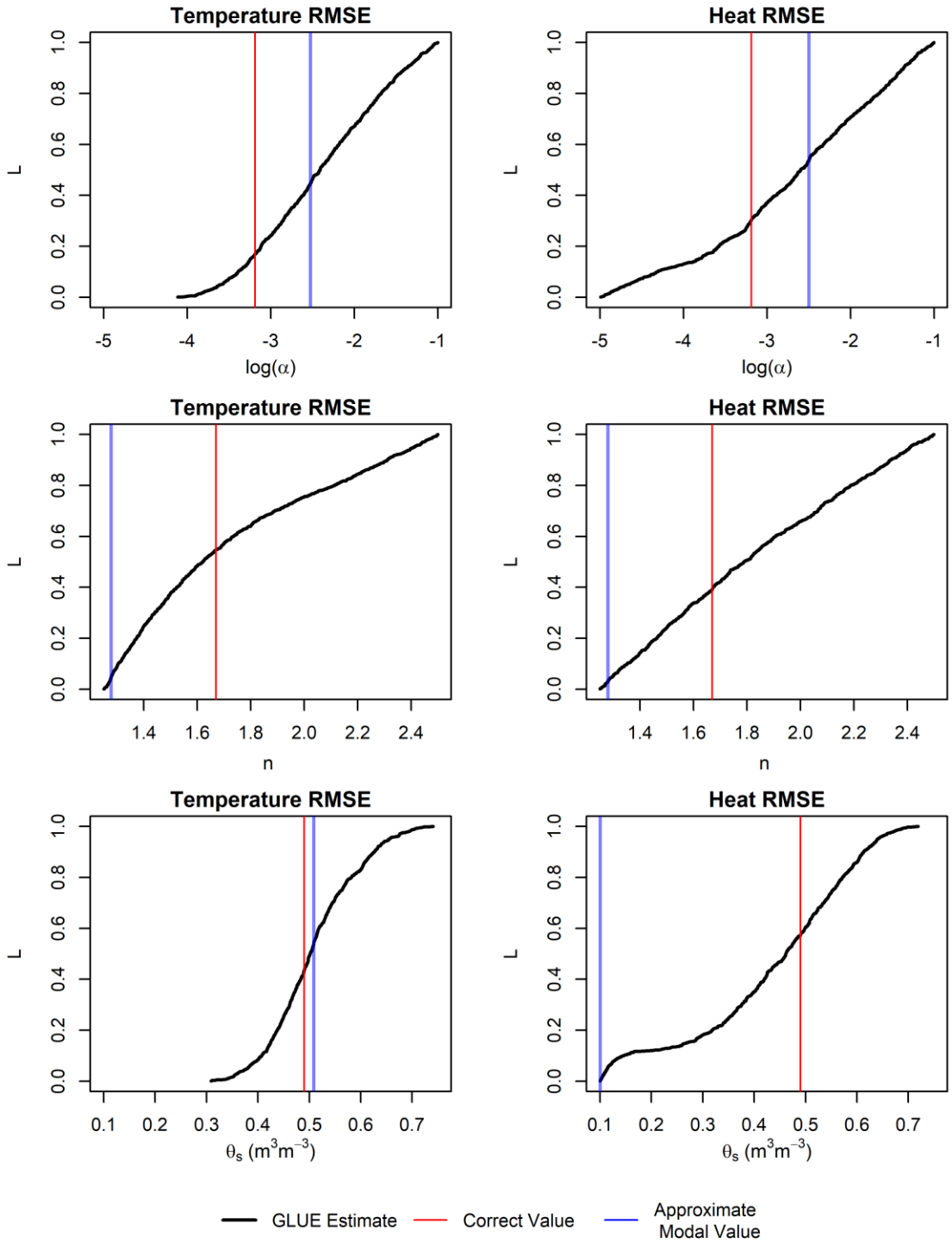

***Figure 15**: GLUE cumulative parameter likelihoods for α, n and θ_s based on wide parameter bounds and unperturbed synthetic data. Vertical lines signify the approximate modal value of distribution as determined from the steepest part of the CDF. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

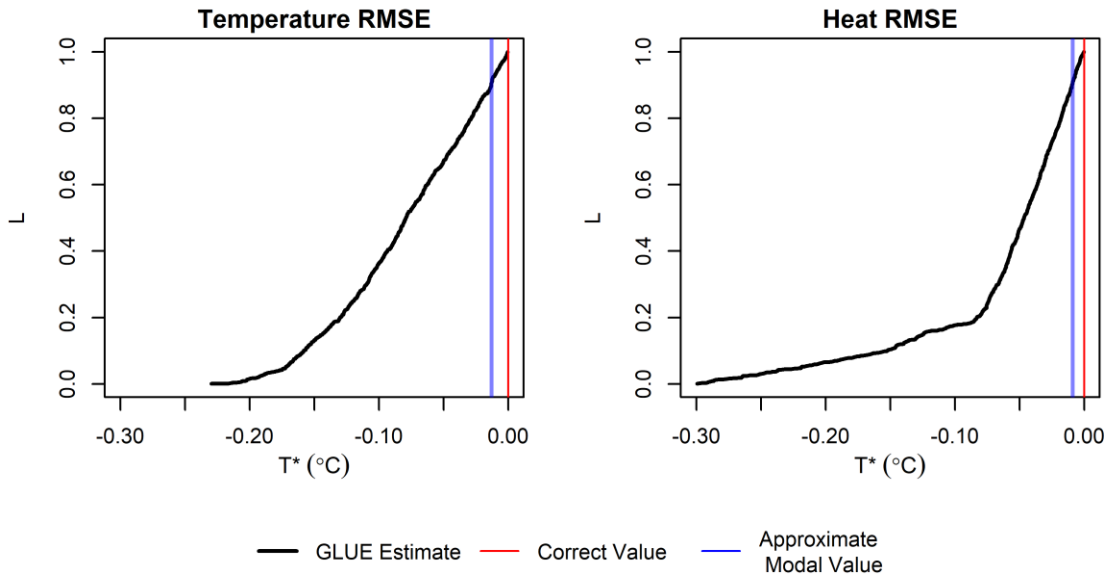

***Figure 16**: GLUE cumulative parameter likelihoods for T* based on wide parameter bounds and unperturbed synthetic data. Vertical lines signify the approximate modal value of distribution as determined from the steepest part of the CDF. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

distribution is bimodal with one peak around the correct value, and the other at the low end of the parameter range.

For the T* parameter, the temperature RMSE yields distributions that are concentrated at the true value of 0 °C and negatively skewed. The heat-based RMSE distributions are generally clustered within 0.08 °C of the correct value and uniformly distributed across the rest of the parameter range. The GLUE-calibrated results at 3 m for wide uncertainty bounds are shown in Figure 17. The temperature-based objective function produces 5 - 95% likelihood bounds that are very close to the correct value. The heat-based objective function on the other hand produces noticeably wider uncertainty bounds. Both uncertainty envelopes contain the correct temperature values.

Based on the limited knowledge of site conditions, the prior uncertainty of liquid water is typically between 0 and 50% and for total energy content this range is between 380 and 1100 MJ m^{-3}, although the upper boundary fluctuates annually. Estimates using GLUE are able to significantly reduce the uncertainty for objective functions. The heat-based objective function predicts lower values for both liquid water and energy content.

4.1.1 Parameter estimation with GLUE: narrowed parameter bounds

The posterior cumulative likelihoods for GLUE parameter estimation using narrowed parameter boundaries are shown in Figures 18 through 20. For both the c_{sp} and λ_{sp} parameters, posterior distributions are uniform across the parameter range. The distribution of κ was approximately normal and centred on the correct value of 1×10^{-6} m^2 s^{-1}. When using the temperature RMSE, the posterior distribution of the van Genuchten log(α) parameter is approximately normal with a modal value of -3. The distribution for the heat-

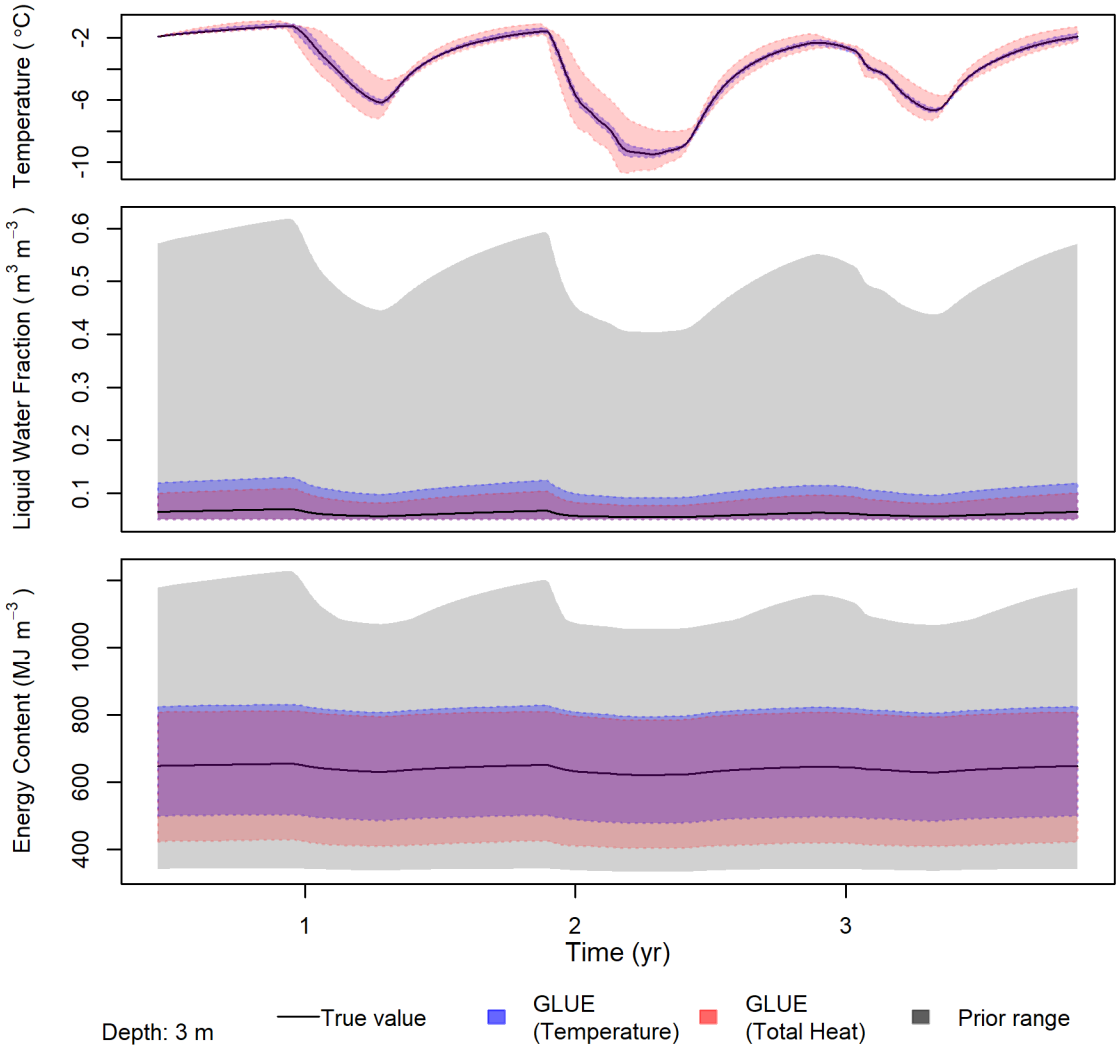

***Figure 17**: Uncertainty envelopes for temperature (middle panel) and liquid water content (bottom panel) for a 3-year synthetic dataset at 1m depth. The 5-95% likelihood uncertainty envelopes are calculated based on the GLUE technique using either the heat-based (red) or the temperature-based (blue) objective functions and a prior distribution that assumes very little information about the site.*

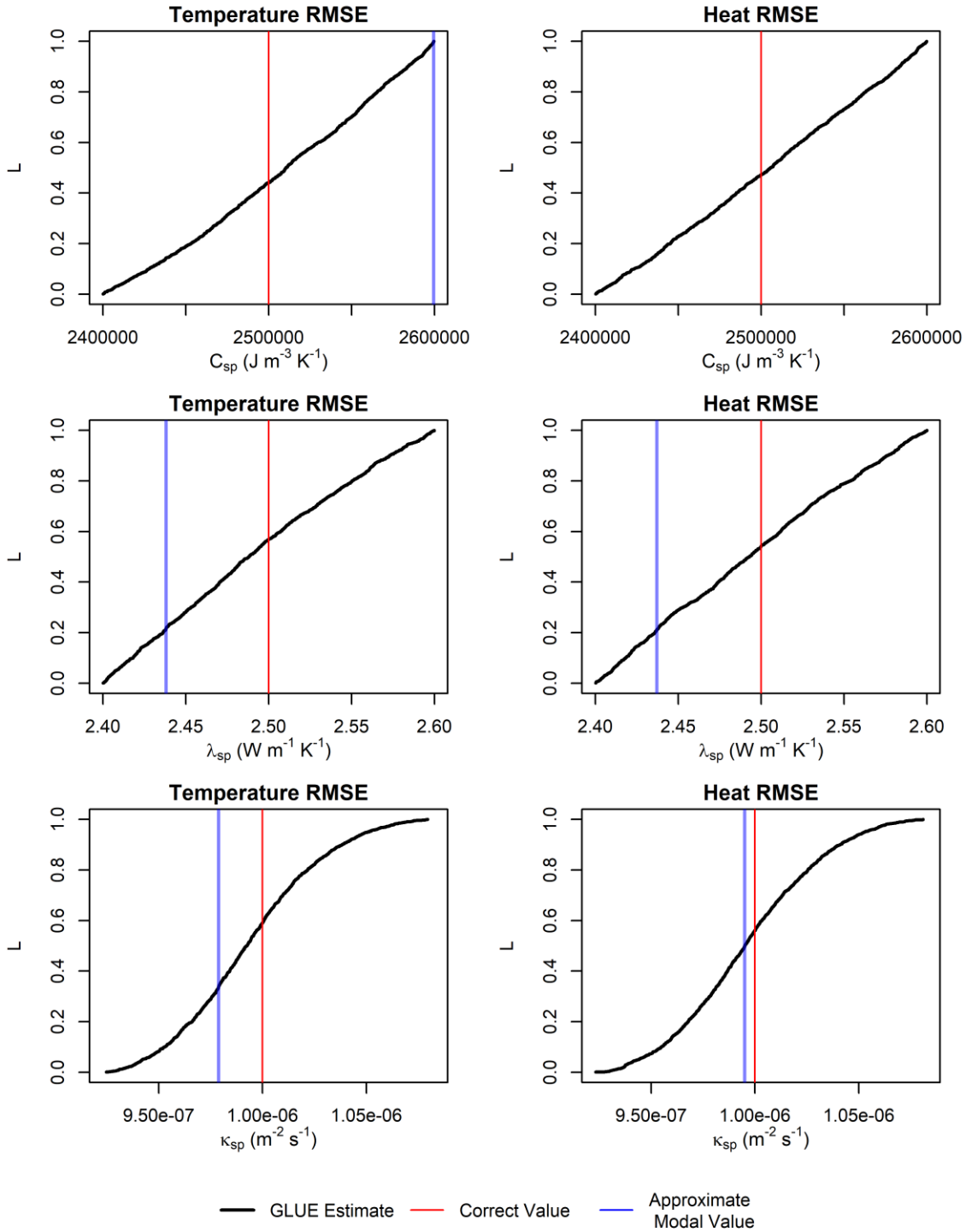

***Figure 18**: GLUE cumulative parameter likelihoods for c_{sp}, λ, and κ based on narrow parameter bounds and unperturbed synthetic data. Vertical lines signify the approximate modal value of distribution as determined from the steepest part of the CDF. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

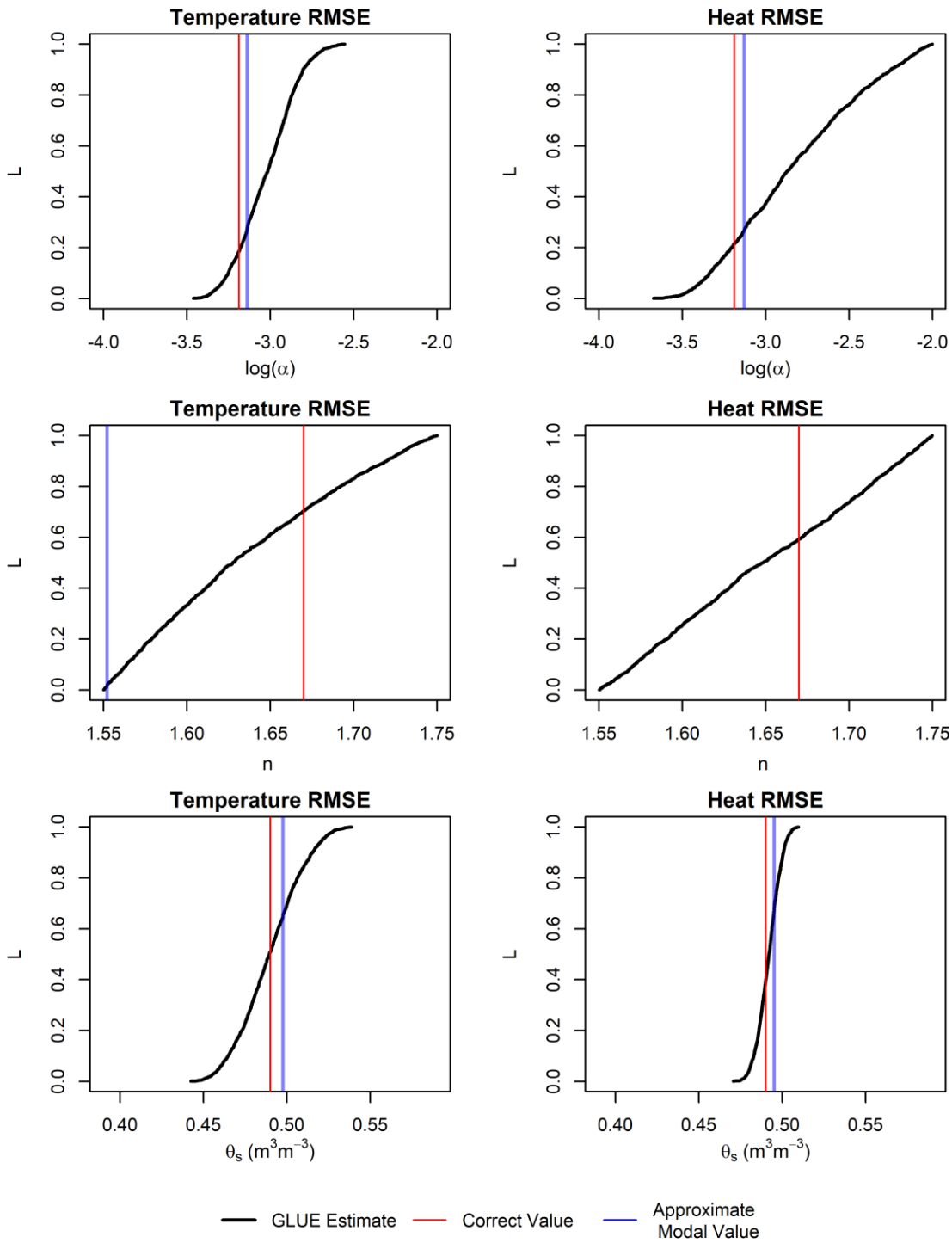

***Figure 19**: GLUE cumulative parameter likelihoods for α, n and θ_s based on narrow parameter bounds and unperturbed synthetic data. Vertical lines signify the approximate modal value of distribution as determined from the steepest part of the CDF. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

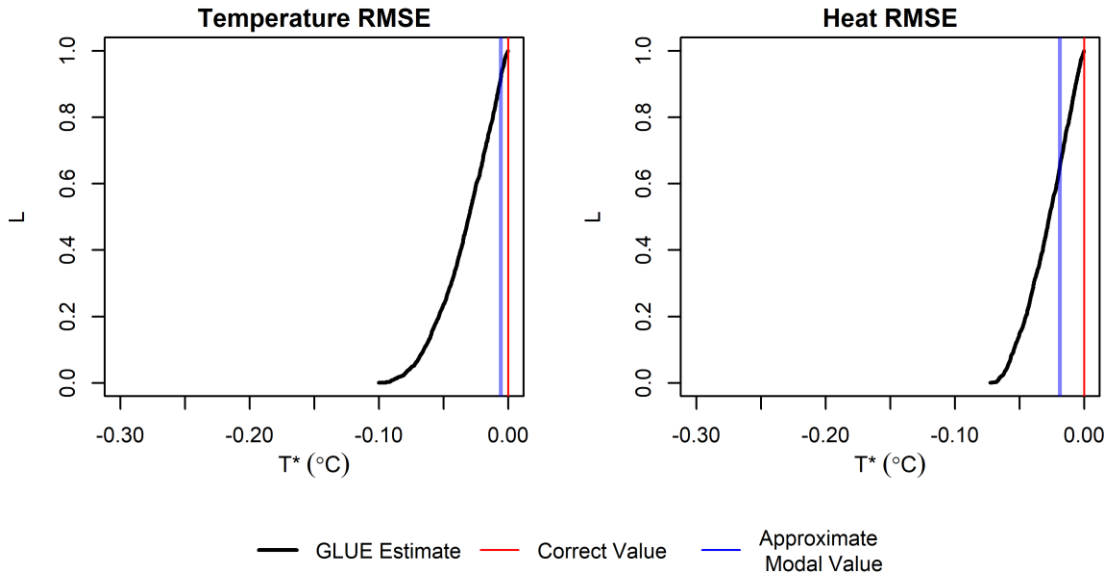

Figure 20*: GLUE cumulative parameter likelihoods for T* based on narrow parameter bounds and unperturbed synthetic data. Vertical lines signify the approximate modal value of distribution as determined from the steepest part of the CDF. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

based RMSE is much more uniform, and the range is only slightly reduced. In contrast, the distribution of the van Genuchten *n* parameter is relatively uniform over the estimation bounds for both formulations of the objective function. Distributions for the θ_s parameter are approximately normal for both formulations of the objective function and centered on or near the true value of 0.49. The total parameter range has been significantly reduced and includes the correct value. The heat-based RMSE is consistently narrower than the temperature-based RMSE.

The T* parameter range is reduced with both formulations of the objective function, although it is slightly more tightly constrained around the true value in the case of the heat-based RMSE.

4.2 Sensitivity of GLUE to observational error

The results of GLUE parameter estimation after adding noise to the synthetic dataset are shown in Figures 21 to 23. Recall that the biased datasets in the GLUE experiments represent only one realization of random noise. This is because of the time requirements for completing the full GLUE analysis. For parameters n, c_{sp}, λ_{sp} and T^*, the noisy data does not appreciably affect the resulting posterior likelihood distribution for either the heat- or temperature-based objective function. In the case of α, the heat-based RMSE is more strongly affected by depth and temperature perturbations. Of the two, the depth perturbation has a greater effect. The θ_s likelihood distribution is most strongly affected by depth perturbation, but only in the case of the heat-based RMSE, where the distribution is shifted to the left. The shape of the distribution is unaffected by either perturbation.

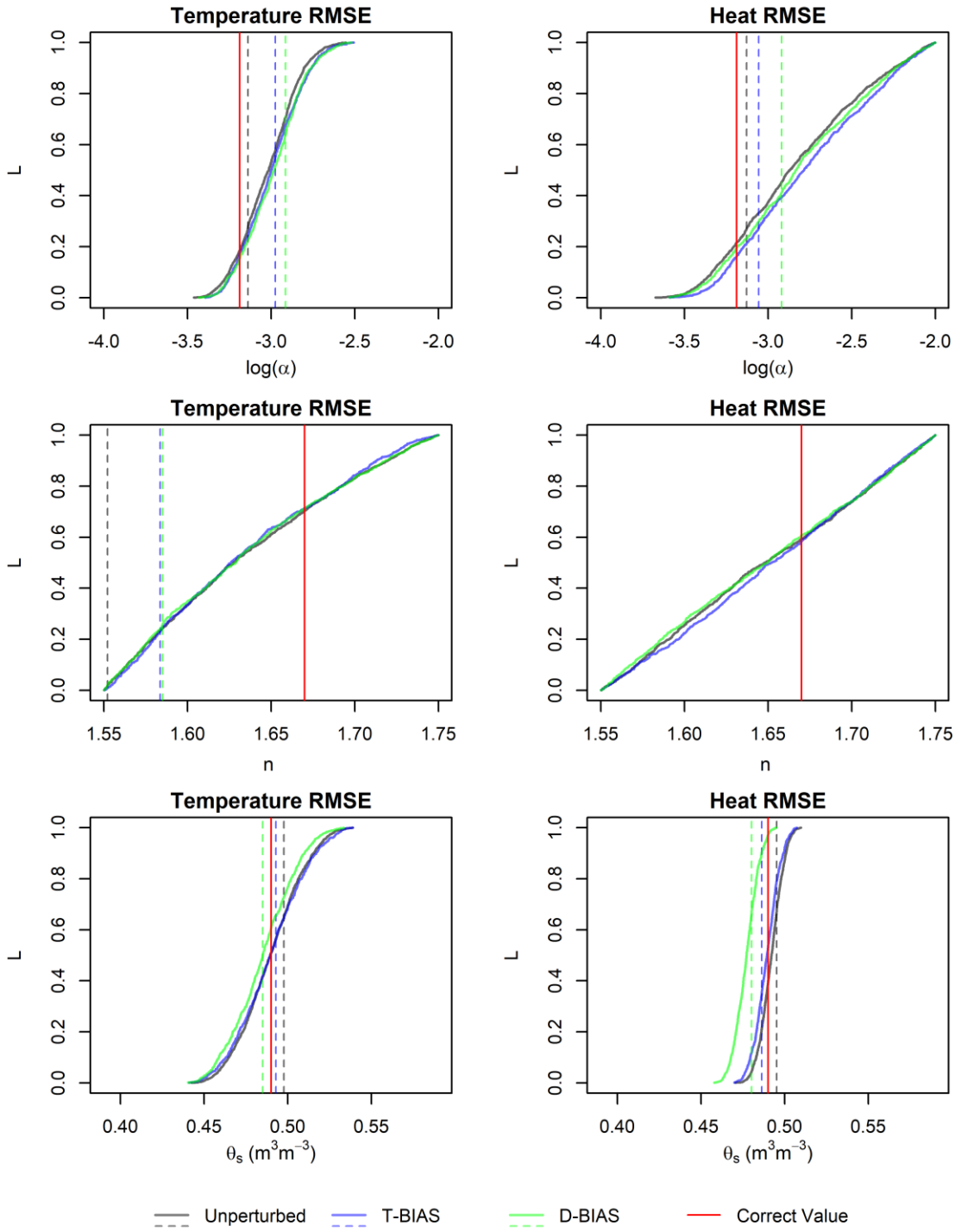

***Figure 21**: GLUE cumulative parameter likelihoods for α, n and θ_s based on narrow parameter bounds and synthetic data with noise added to temperature (T-BIAS) or depth (D-BIAS). Likelihood function retains best 1% of simulations. Dashed line signifies modal value of distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

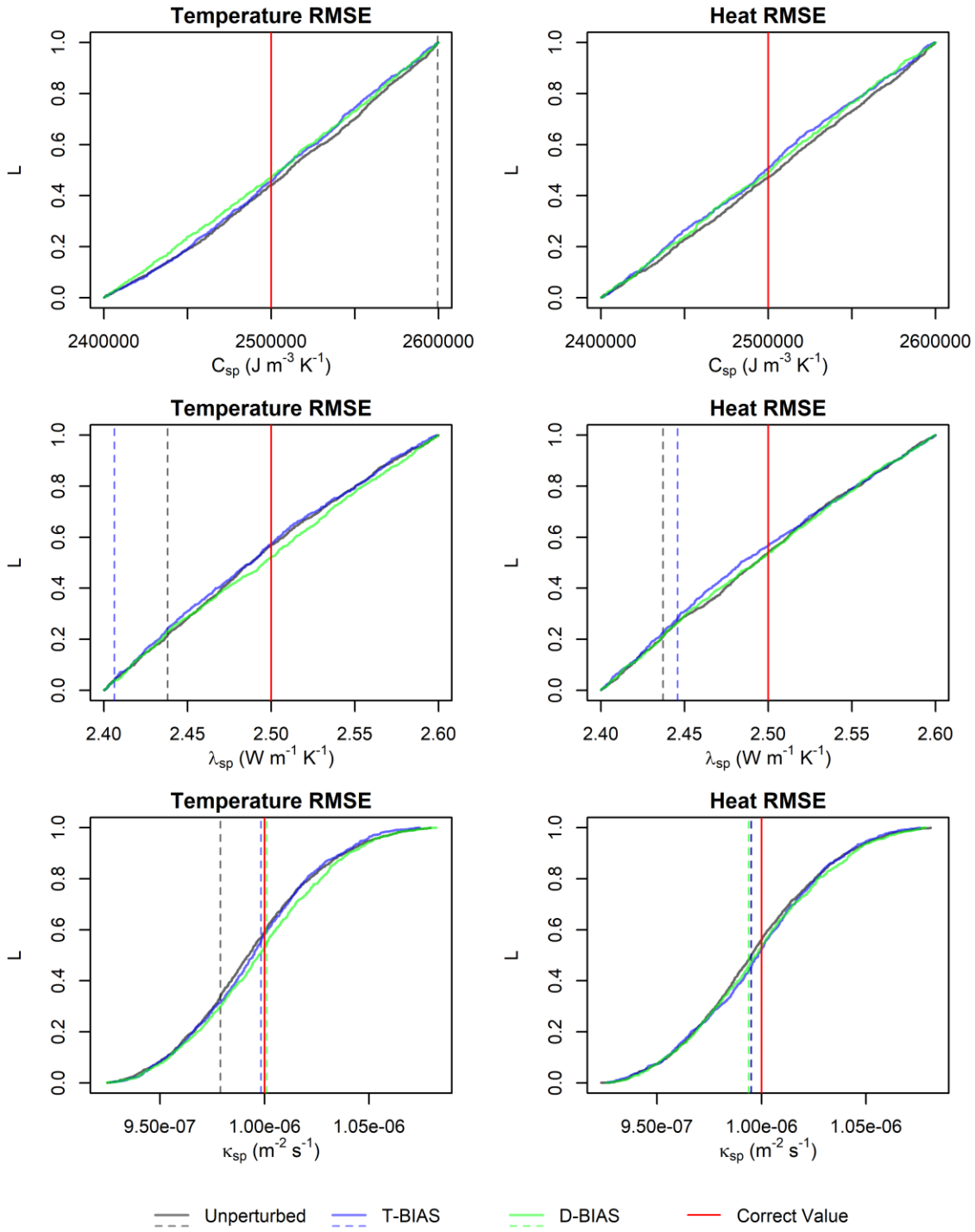

***Figure 22**: GLUE cumulative parameter likelihoods for thermal conductivity (λ), heat capacity (c) and thermal diffusivity (κ) based on narrow parameter bounds and synthetic data with noise added to temperature (T-BIAS) or depth (D-BIAS). Likelihood function retains best 1% of simulations. Dashed line signifies modal value of distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

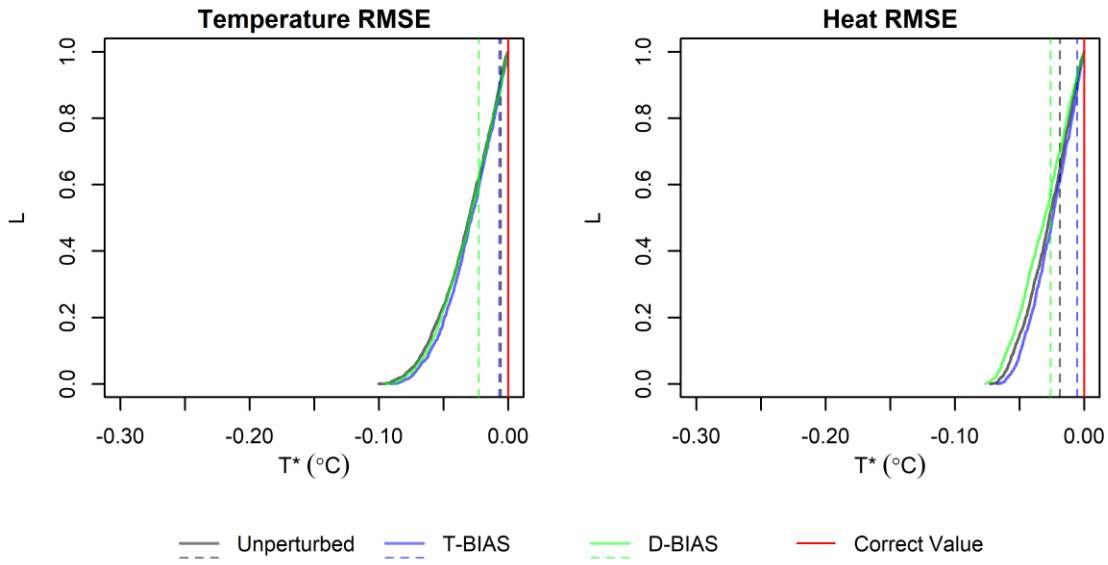

***Figure 23**: GLUE cumulative parameter likelihoods for T* based on narrow parameter bounds and synthetic data with noise added to temperature (T-BIAS) or depth (D-BIAS). Likelihood function retains best 1% of simulations. Dashed line signifies modal value of distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01*

.

The uncertainty associated with the GLUE 5-95% likelihood bounds for synthetic data with narrow parameter bounds using biased and unbiased data are shown in Figure 24. The heat-based objective function consistently produces wider uncertainty envelopes both for the fitted temperature data as well as the calculated energy and water content. The effect of adding noise to the observations is negligible. At the plotted depth of 5 m, there is no change to the heat or water uncertainty envelopes when using the temperature-based objective function. For the heat-based objective function, the noise lowers the upper uncertainty limit of the liquid water estimation, and shifts the envelope down for the estimation of energy content. However, this change is slight.

4.3 Parameter estimation with DE using synthetic data

Each optimization took between 12,510 and 17,550 GEOtop iterations to complete; this required between 77 and 110 hours. An example of the progression of two parameter estimates during optimization is shown for reference in Figure 25.

The results of using DE to estimate parameter values for unperturbed and biased synthetic data is shown in Figure 26. For unbiased synthetic data (PURE), the differences between all parameter estimates and the correct values are negligibly small. Furthermore, the variability between different realizations of the final estimates due to the stochastic nature of the technique is also very small.

Some early attempts at optimizing parameters also yielded some information about how best to carry out the procedure; if the θ_r parameter is allowed to vary instead of being set to a fixed value, it causes the estimated parameters to converge to incorrect values, even

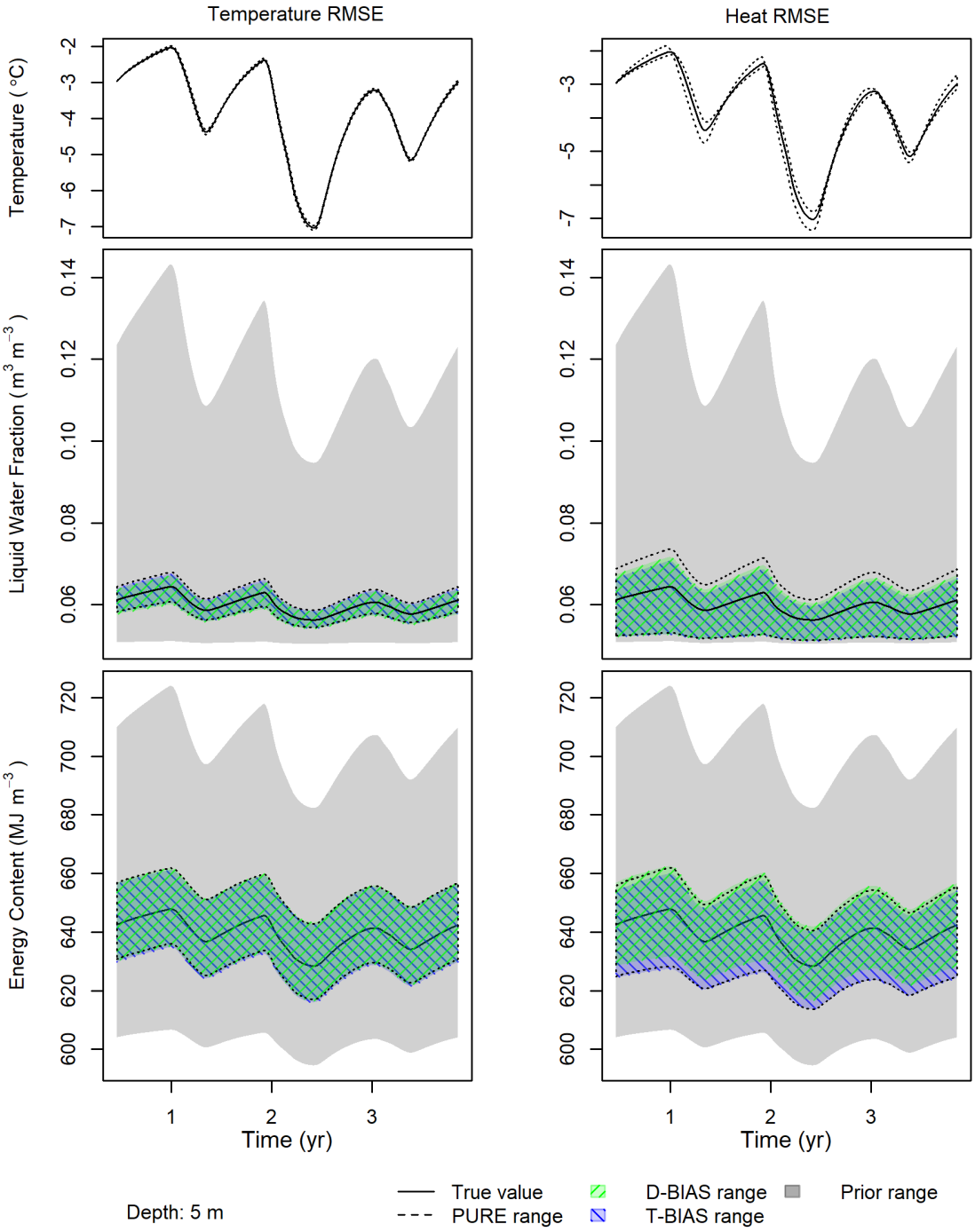

***Figure 24**: Uncertainty envelopes for temperature (middle panels) and liquid water content (bottom panels) for a 3-year synthetic dataset at 5 m depth. The 5-95% likelihood uncertainty envelopes are calculated based on the GLUE technique using either the temperature-based (left) or the heat-based (right) objective functions. The prior distribution for GLUE assumes that good-quality data are available for the site.*

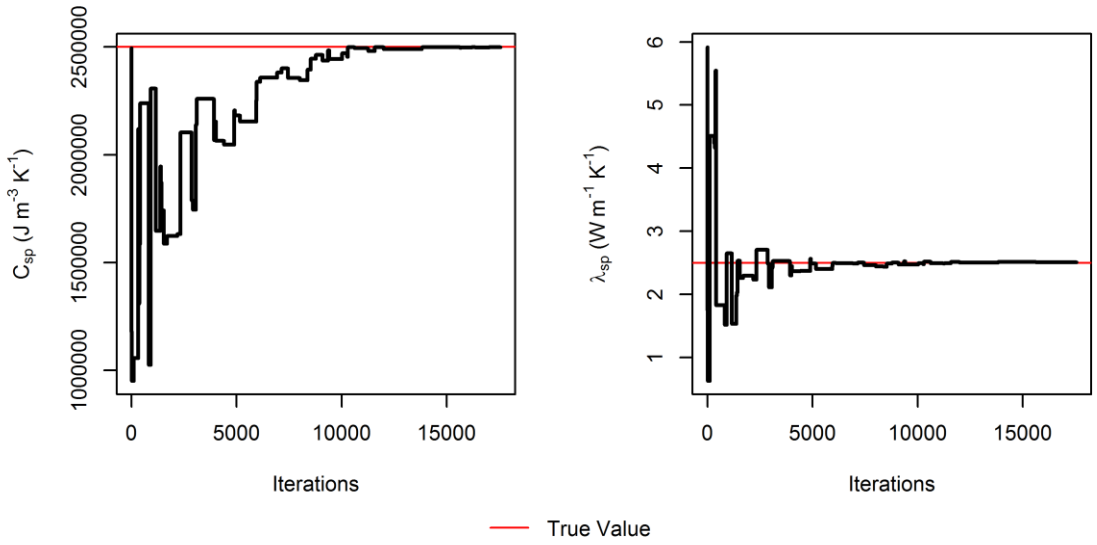

***Figure 25**: Convergence of parameter estimates. The black line indicates the value of a given parameter associated with the best parameter set at a given number of iterations. In both examples, the DE algorithm converges on the true value of the parameter.*

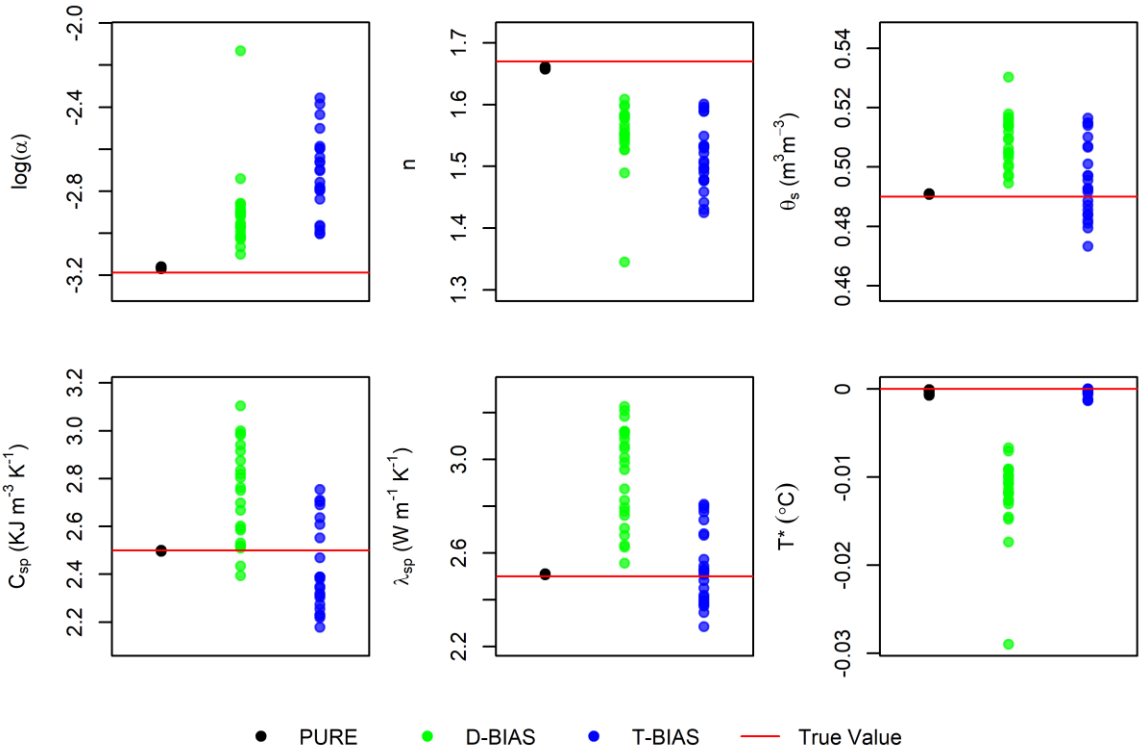

***Figure 26**: Estimated parameters using DE search with pure synthetic data as well as temperature- and depth-biased data.*

with unbiased synthetic data. Choosing random values of α from a logarithmic uniform rather than a standard uniform distribution also improved the results.

When noise is added to the synthetic data, the resulting parameter estimates are no longer tightly clustered on the correct value (Figure 26). Although the results using the PURE data represent 20 realizations of the same synthetic observations (to capture the inherent variability of the algorithm), the perturbed DE experiments using D-BIAS and T-BIAS are conducted with 20 different realizations of random noise. This differs from the GLUE experiments which use only one realization of the random noise. The range of estimates increases compared to the unperturbed data and the estimates are further away from the correct value.

Some parameters are more strongly affected by the noise than others. The range of variability for λ_{sp} in the noisy model runs (1.0 W m^{-1} K^{-1}) represents 40% of value of the absolute parameter value. C_{sp} is also strongly affected, with the range of variability representing 28% of the absolute parameter value. For log(α), n and θ_s, these values are 18%, 12% and 10%. The spread of the θ_s parameter is increased to 0.05 for T-BIAS and to just over 0.3 for T-BIAS. The T* parameter is the least affected by the biased datasets; estimated parameters are within 0.03 °C of the correct value, and there is little difference between the results for T-BIAS and PURE data. Overall, the range of parameter estimates associated with T-BIAS more frequently contain the correct values than those associated with D-BIAS.

Despite the spread in parameter estimates, the resulting SFCC is reasonably close to the correct one (Figure 27). In the case of D-BIAS, the total error away from the rising limb of the curve (temperatures less than -0.10 °C) is typically less than 0.06. For the T-BIAS, this is within 0.03. Within a narrow temperature range, in this case between -0.10 °C and 0 °C, the uncertainty becomes very large due to the strong nonlinearity near the freezing point.

The ranges of estimated heat and liquid water contents subject to depth and temperatures biases are shown in Figure 28 (3 year) and Figure 29 (100 year). For the 3-year dataset (Figure 28), the shape of the uncertainty envelopes generally resembles the true value. In the case of the liquid water content, the error range overestimates the liquid water content when the true value is less than 0.1. This is likely a consequence of the upper bound for T*: since the parameter can never achieve a value greater than 0°C, the estimated freezing point is always lower than the true freezing point and the water content is higher because the SFCC is shifted to the left. For a real site, it is unlikely that the true freezing point would be exactly 0°C, and so this problem of overestimation should not be expected to be a systemic problem with the technique presented. Overall, despite the added noise, estimates of liquid water content are within a few percent of the correct value. Error ranges for total ground energy content are also very accurate, and consistently contain the true value.

In the 100-year dataset, at 4 m depth, predictions of energy and water content are consistent with the true value. At this depth, annual temperature fluctuations weaken and disappear at about 40 years after warming begins, but the calculated phase change at that

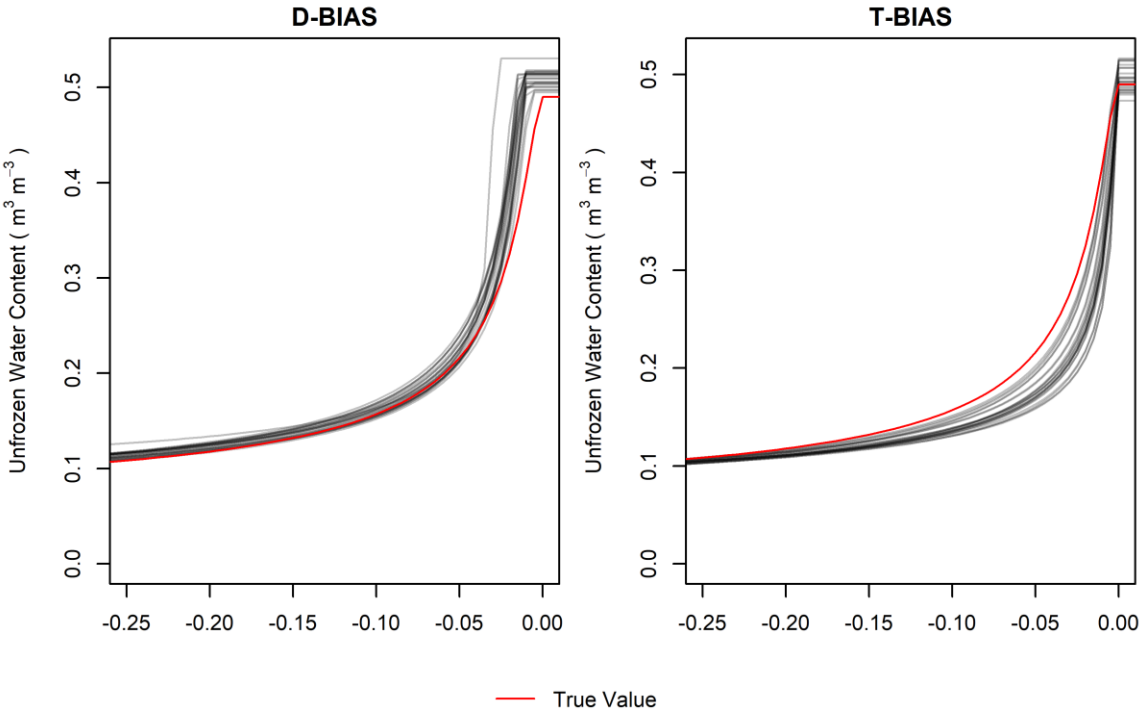

***Figure 27**: Effect of observational depth (D-BIAS) and temperature (T-BIAS) bias on fitted SFCC for synthetic data using the DE algorithm.*

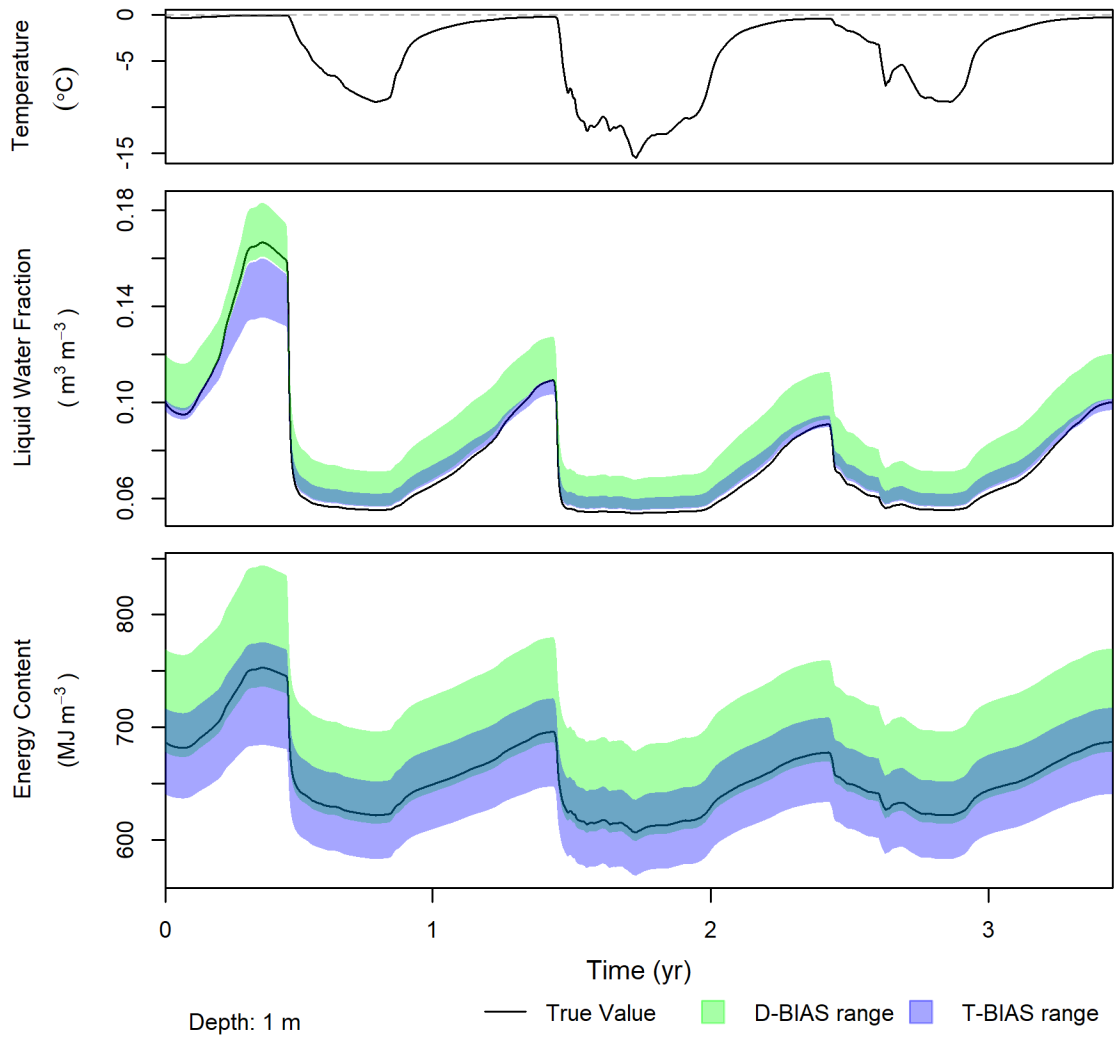

***Figure 28**: Effect of depth and temperature bias on estimates of liquid water and content for a 3-year synthetic data time series at 1 m. Coloured envelopes represent the range of estimates produced by 20 different realizations of perturbed simulations.*

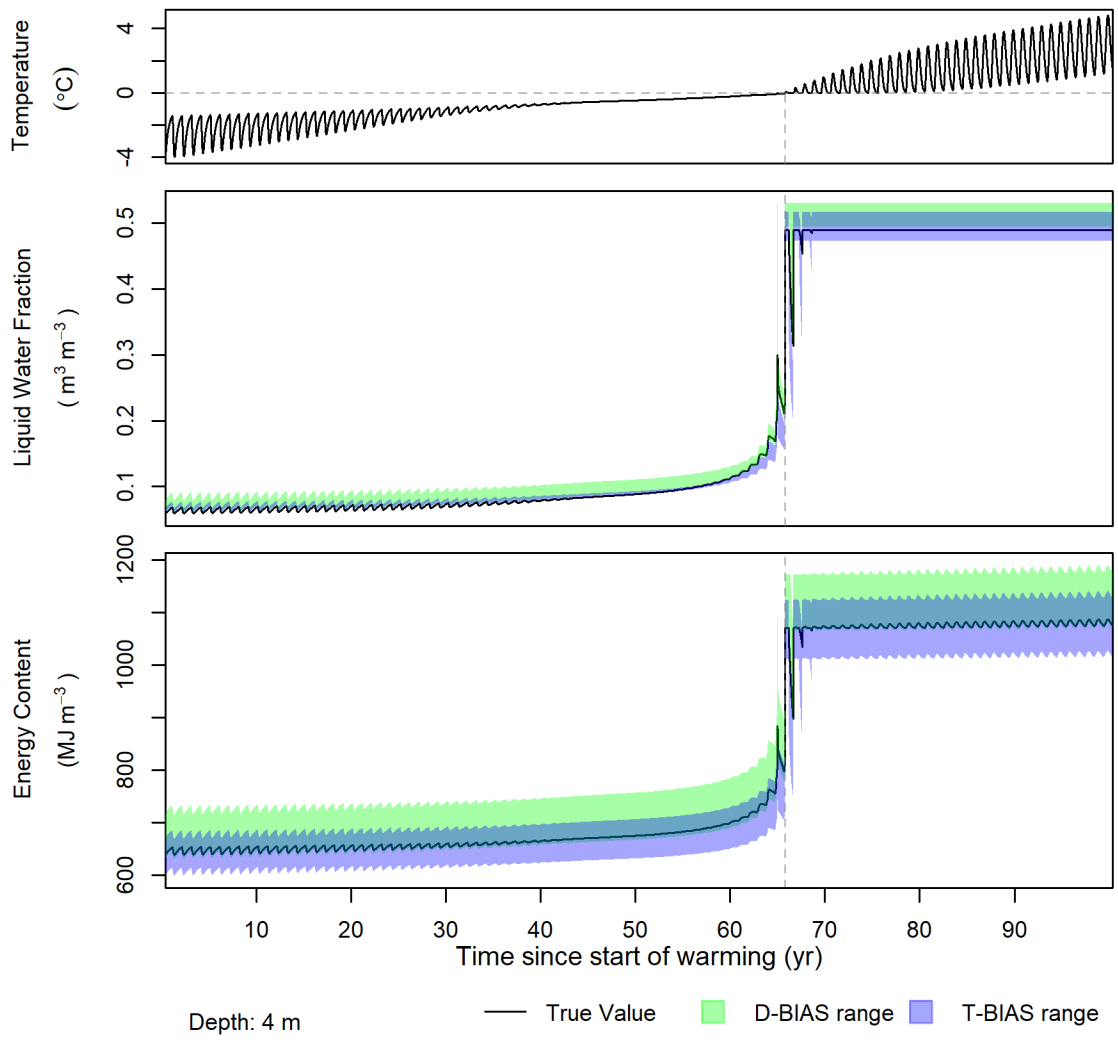

***Figure 29**: Effect of depth and temperature bias on estimates of liquid water and content for a 100-year warming synthetic data time series at 4 m. Coloured envelopes represent the range of estimates produced by 20 different realizations of perturbed simulations.*

depth does not occur many years later, around 65 years after warming begins. The range of estimates for liquid water content associated with the noisy data are approximately 2% when the temperature is away from the freezing point and increases slightly when the bulk of the phase change is occurring.

4.4 Application of GLUE and DE to borehole data

4.4.1 Estimated soil parameters from ROSETTA

The initial GLUE sampling distributions for the Chick Lake and Gibson Lake were estimated using soil textural data and the pedotransfer function software ROSETTA. The mean and standard deviation for soil parameters predicted by ROSETTA are plotted in Figure 30, and listed in Table 6 and Table 7. Note that the ranges used for the GLUE prior distributions represent a range of four standard deviations. For each site, soil units are defined by intervals whose soil parameter estimates are based on the same soil texture sample. For the Gibson Lake site, the estimated parameters for the soil at 4, 7 and 12 m (units 2 & 3) are quite similar. However, estimated parameters for soil at 1.5 and 2 m (unit 1) appear distinctly different. At the Chick Lake site, parameter estimates are also typically more similar between the second and third soil unit than they are to the first, although the first soil unit is more similar than at Gibson Lake. Parameter uncertainties are relatively similar between sites.

4.4.1 Gibson Lake

The results of GLUE and DE calibration for the Gibson Lake borehole are shown in Figure 31 and Figure 32. For GLUE, both the α and n distributions are positively skewed

***Table 6**: Mean ($\bar{x}$) and standard deviation (σ) for estimated soil properties from ROSETTA for the Gibson Lake permafrost monitoring site.*

Sensor depths (m)	Log(α) ($\bar{x}$, σ)	Log(n) ($\bar{x}$, σ)	θ_s ($\bar{x}$, σ)	θ_r ($\bar{x}$, σ)
1.5, 2	-2.00, 0.07	0.17, 0.01	0.49, 0.01	0.10, 0.01
4	-2.27, 0.09	0.21, 0.01	0.43 0.01	0.07, 0.01
7, 12	-2.30, 0.06	0.21, 0.01	0.42, 0.01	0.07, 0.01

***Table 7**: Mean ($\bar{x}$) and standard deviation (σ) for estimated soil properties from ROSETTA for the Chick Lake permafrost monitoring site.*

Sensor depths (m)	Log(α) ($\bar{x}$, σ)	Log(n) ($\bar{x}$, σ)	θ_s ($\bar{x}$, σ)	θ_r ($\bar{x}$, σ)
1.5, 2, 4	-2.04, 0.08	0.08, 0.02	0.49, 0.01	0.09, 0.01
7	-1.87, 0.11	0.13, 0.02	0.51, 0.02	0.11, 0.02
12	-1.92, 0.11	0.11, 0.02	0.51, 0.02	0.10, 0.01

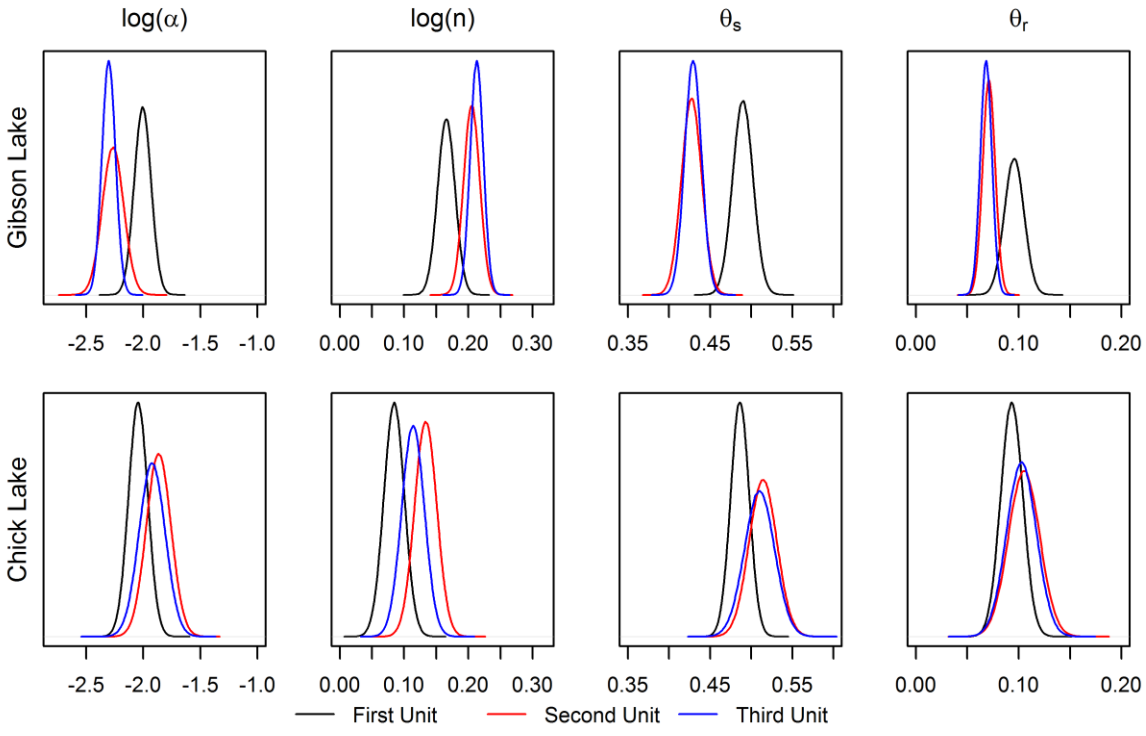

***Figure 30**: Comparison of estimated parameters from ROSETTA for Chick Lake and Gibson Lake. The depths corresponding to each unit are shown in Table 6 and Table 7.*

and clustered at the lower boundary. Neither distribution is affected by the choice of objective function. The parameter range of the *n* distribution has been reduced somewhat, with the upper boundary being reduced to 1.58 and 1.68 for the temperature and heat RMSE respectively. The behaviour of the θ_s distribution is strongly dependent on the choice of objective function. For the temperature RMSE, the distribution is slightly negatively skewed, but distributed over the entire parameter range. The distribution associated with the heat RMSE is much more strongly positively skewed, clustered at the lower parameter bound and the range has been reduced somewhat. The T* parameter is also dependent on the objective function. The temperature RMSE is positively skewed and clustered towards the lower parameter bound, whereas the heat RMSE is approximately normally distributed about -0.14 °C.

The posterior c_{sp} parameter (Figure 32) is only slightly affected by the choice of objective function. In both cases, the likelihood is distributed over the entire parameter range, but the temperature RMSE is biased towards higher values. The distribution for λ_{sp} is strongly affected by the choice of objective function. For the temperature RMSE, it is tightly clustered near the lower parameter bound of 1.0 and the total range is greatly reduced whereas for the heat RMSE it is approximately normally distributed around 2.1 and the total parameter range is slightly reduced. Both distributions are also clustered near the lower bound. The distribution for the heat RMSE is slightly wider and has a slightly more normal–albeit positively skewed–shape. Parameter estimates from the DE algorithm tend to only be coincident with the modal value for the temperature RMSE

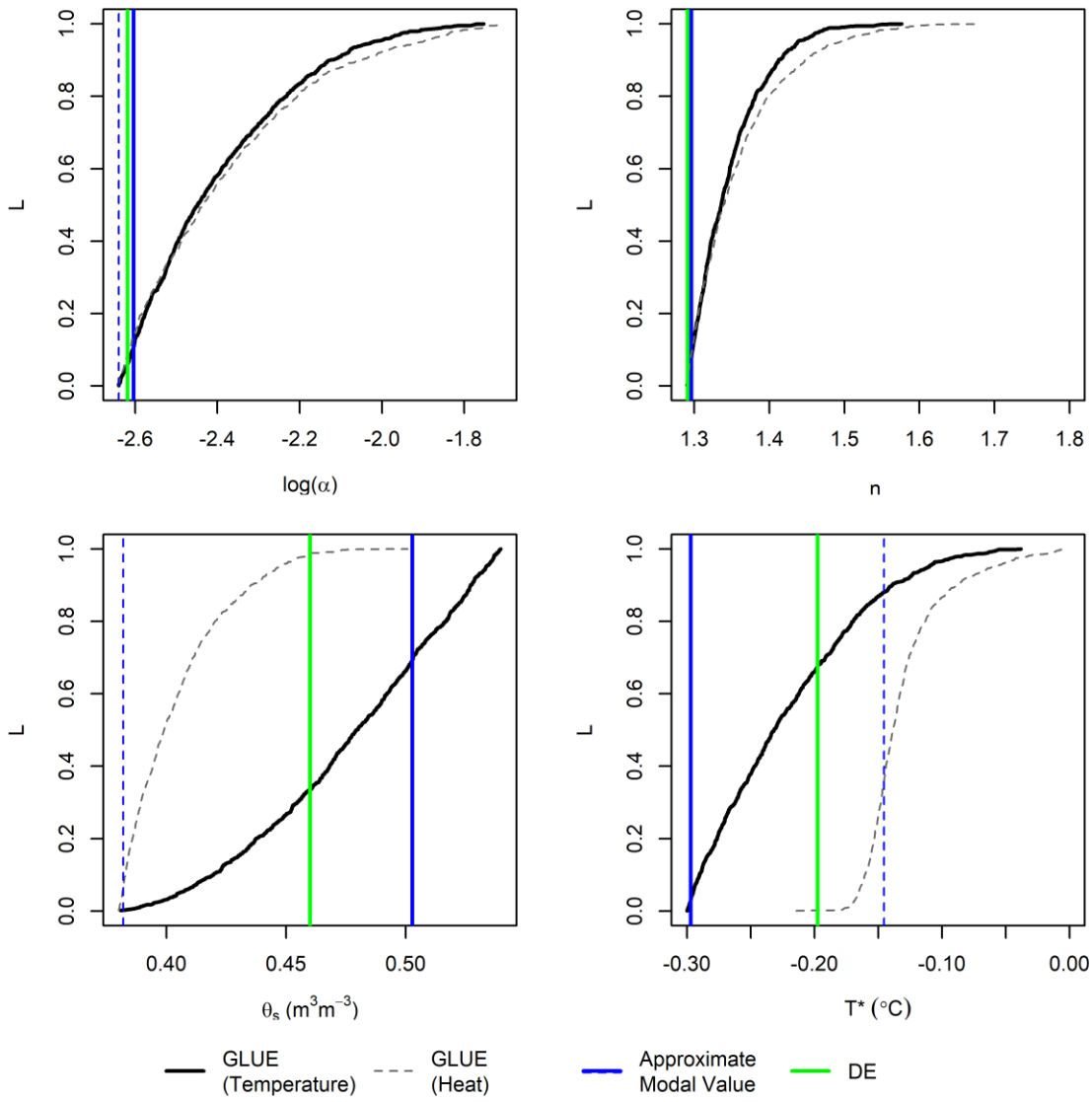

***Figure 31**: DE parameter estimates and GLUE cumulative parameter likelihoods for α n, θ_s and T* for Gibson Lake borehole. The GLUE likelihood function retains the best 1% of simulations. Blue lines signify the approximate modal value of the distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

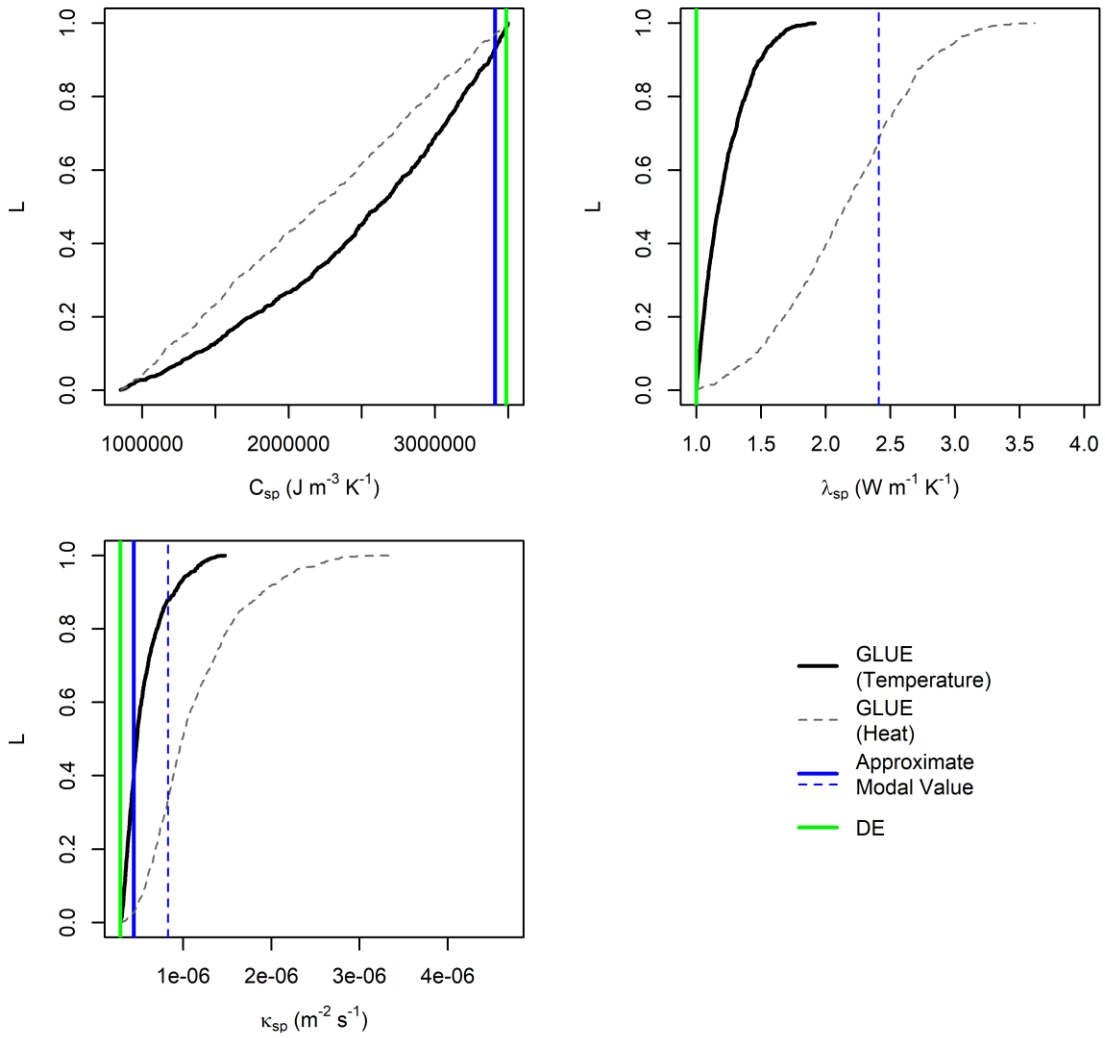

***Figure 32**: DE parameter estimates and GLUE cumulative parameter likelihoods for thermal conductivity (λ), heat capacity (c) and thermal diffusivity (κ) for Gibson Lake borehole. The GLUE likelihood function retains the best 1% of simulations. Blue lines signify the approximate modal value of the distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

when both are at the very edge of the parameter bound. For θ_s and T*, where GLUE distributions are not clustered at the edge of the parameter bounds, the estimates are distinctly different. Note that while GLUE is being evaluated with two different objective functions (heat and temperature), the DE algorithm is using only a temperature-based objective function.

The uncertainties at 1 m depth associated with the GLUE technique are shown in Figure 33. The uncertainty envelope for the temperature-based objective function is typically associated with greater liquid water contents. This is not surprising considering that the final distribution of the T* parameter is heavily weighted towards lower temperatures. That means that at the same temperature, the freezing point is more likely to have been surpassed.

Neither objective function produces uncertainty envelopes that enclose the observed temperature for the entire observational period. At this depth, the temperature-based objective function shows more fidelity to the observed temperature data. However, for temperatures near the freezing point, the heat-based objective function yields temperatures equally consistent with observations.

At 1.5 and 2.0 m, the heat-based objective function produces uncertainty envelopes that more closely match the observed temperatures near the freezing point (Figure 34), although again the modeled temperatures are overpredicted in the summer and underpredicted in the winter. The temperature-based objective function tends to underpredict temperature near the freezing point. Again, this is a consequence of the lower

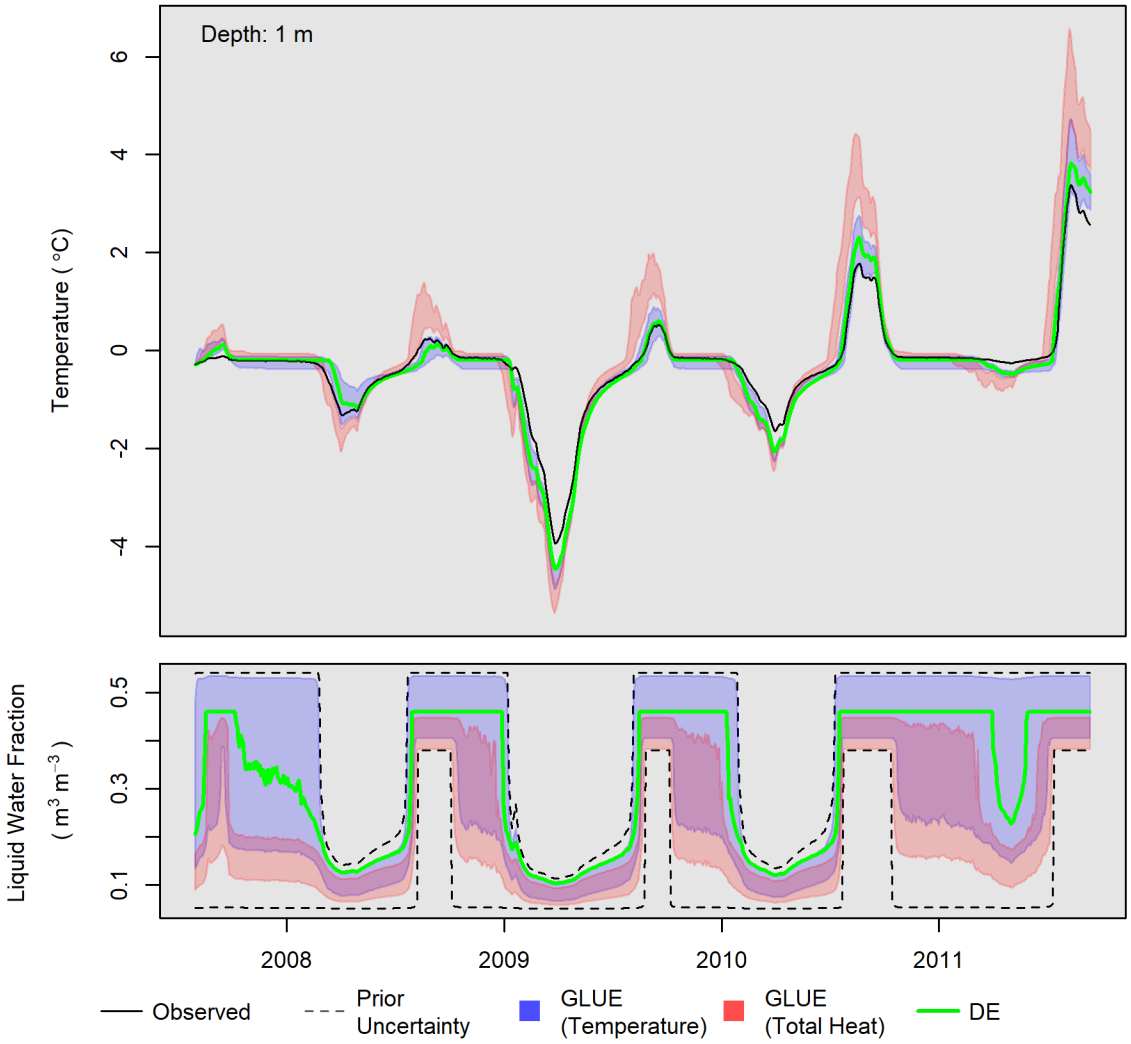

***Figure 33**: Modeled temperature (top) and liquid water content (bottom) at the Gibson Lake borehole, 1 m depth. The 5-95% likelihood uncertainty envelopes are calculated based on the GLUE technique using either the heat-based (red) or the temperature-based (blue) objective functions. Estimates using DE are shown in green.*

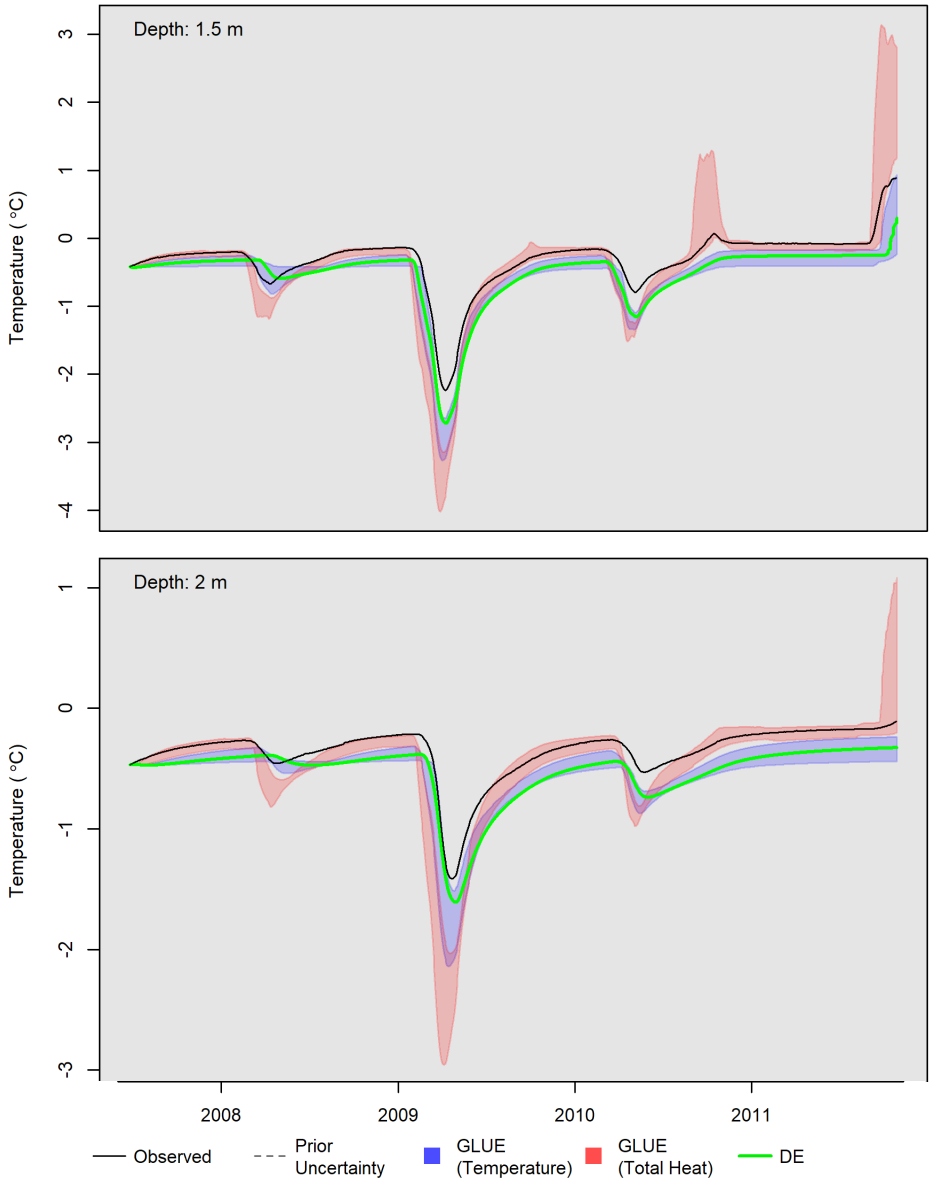

***Figure 34**: Modeled temperature for Gibson Lake at 1.5 (top) and 2.0 m (bottom). GLUE 5-95% uncertainty envelopes are calculated based on the GLUE technique using either the heat-based (red) or the temperature-based (blue) objective functions. Estimates using DE are shown in green.*

values for the soil freezing temperature; the zero curtain occurs at a lower temperature because the predicted freezing point is lower.

The uncertainties at 4 m depth associated with the GLUE technique are shown in Figure 35. The temperature increases very slightly, decreases in mid-2009 and then increases gradually again. The temperature values range between -0.5 °C and -0.35 °C. Neither temperature nor heat objective functions are able to fully recreate the observed temperature curve, although the error is less than 0.2 °C in either case. The modeled temperatures typically overpredict the temperature for both objective functions.

4.4.2 Chick Lake

The results of GLUE and DE calibration for the Chick Lake borehole CL01 are shown in Figure 36 and Figure 37. The α distribution is positive skew-uniform and clustered somewhat at the lower bound. The heat-based objective function is slightly more uniform. The *n* distribution is strongly clustered at the lower bound, particularly for the temperature-based RMSE. The upper parameter bound has been reduced to between 1.35 and 1.48. The θ_s distribution is uniform for the temperature RMSE and positive skew-uniform for the heat RMSE. The parameter range has not been reduced at all. The distribution of the T* parameter is strongly affected by the choice of objective function; for the temperature RMSE, it is positively skewed and clustered around the lower bound of -0.3 °C whereas the heat RMSE is approximately normally distributed around -0.15 °C.

As with the Gibson Lake site, parameter estimates from the DE algorithm at Chick Lake tend to only be coincident with the modal value for the temperature RMSE when both

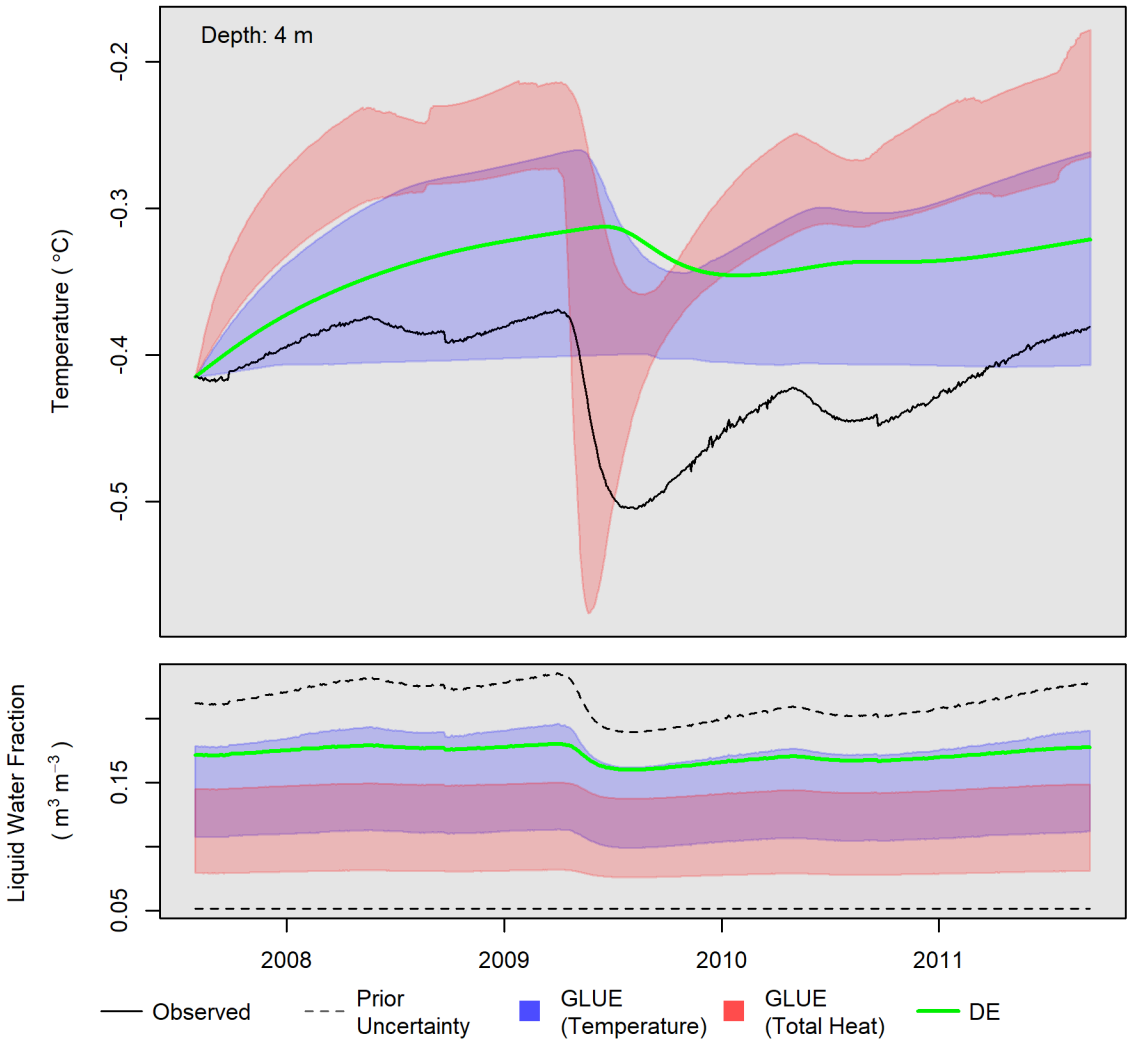

***Figure 35**: Modeled temperature (top) and liquid water content (bottom) at the Gibson Lake borehole, 4 m depth. The 5-95% likelihood uncertainty envelopes are calculated based on the GLUE technique using either the heat-based (red) or the temperature-based (blue) objective functions. Estimates using DE are shown in green.*

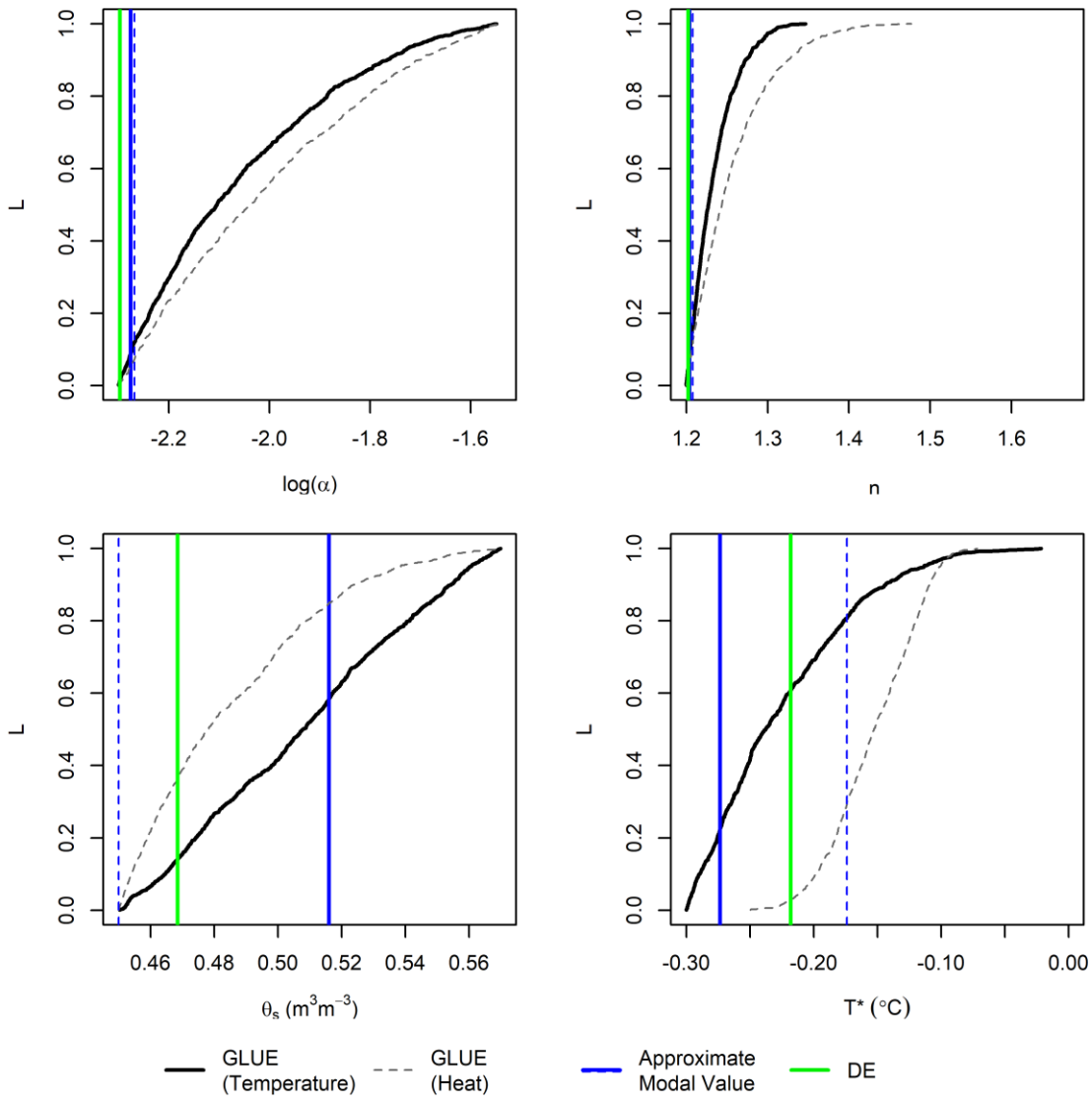

***Figure 36**: DE parameter estimates and GLUE cumulative parameter likelihoods for α n, θ_s and T* for Chick Lake borehole. Blue lines signify the approximate modal value of the distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

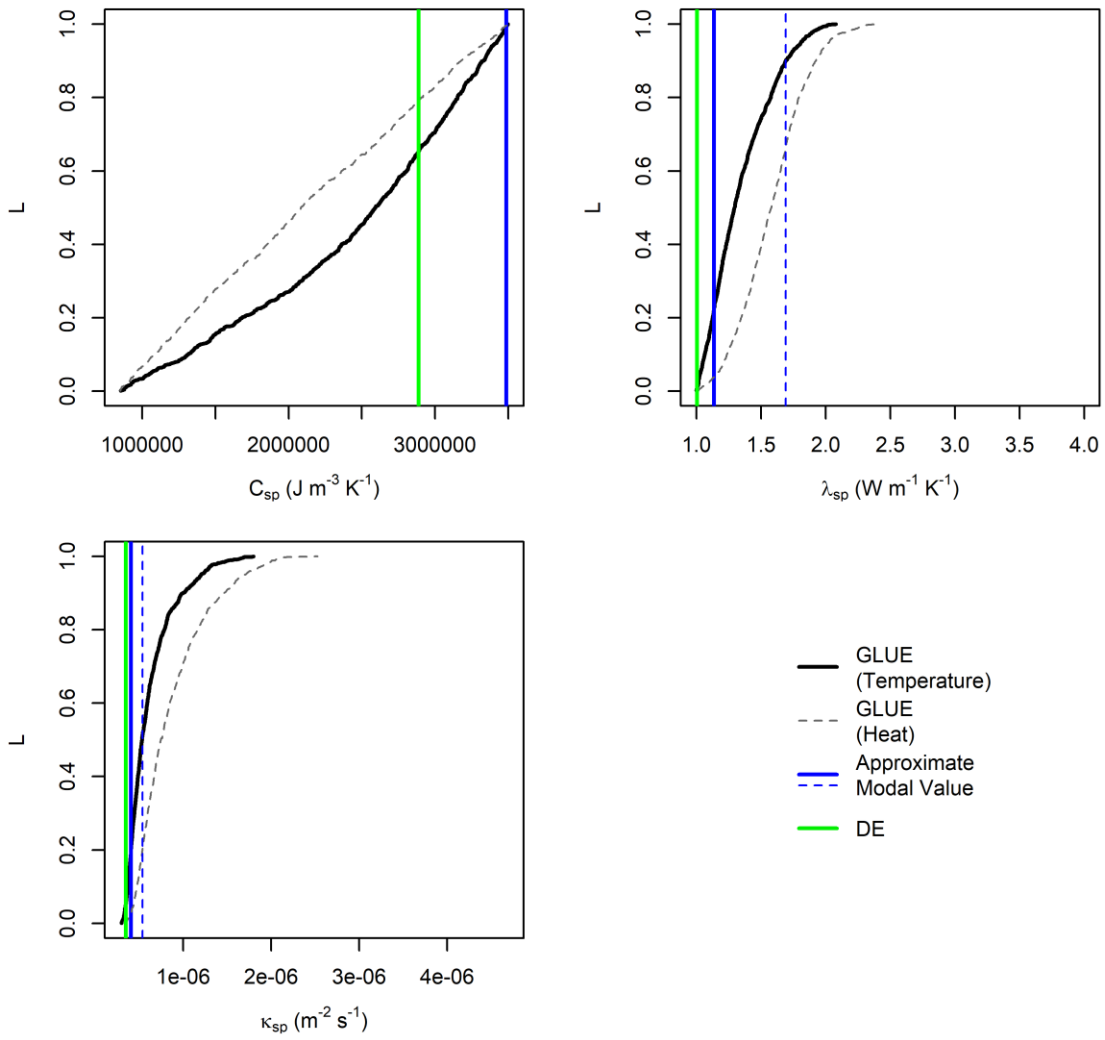

***Figure 37**: DE parameter estimates and GLUE cumulative parameter likelihoods for thermal conductivity (λ), heat capacity (c) and thermal diffusivity (κ) for the Chick Lake borehole. Blue lines signify the approximate modal value of the distribution. Modal values are only calculated for distributions significantly different from uniform at p<0.01.*

are at the very edge of the parameter bound. Again, θ_s and T* show the most distinctly different estimates between the two techniques.

The c_{sp} parameter distribution (Figure 37) is uniform in the case of the heat RMSE and slightly negatively skewed with the modal value at the upper bound for the temperature RMSE. The parameter range has not been narrowed. The λ_{sp} distribution is clustered at the lower bound of 1.0 W m^{-1} K^{-1} for the temperature RMSE and is approximately normally distributed around 1.5 W m^{-1} K^{-1} for the heat RMSE. Values of λ_{sp} larger than 2.5 W m^{-1} K^{-1} are rejected as non-behavioural by this analysis.

The uncertainties at 2 m depth associated with the GLUE technique are shown in Figure 38. At this depth, the annual amplitude of the temperature signal decreases each year. Both objective functions produce GLUE temperature uncertainty envelopes that resemble the observed temperature. However, the uncertainty bounds generated by the temperature-based objective function tend to contain the observed temperature more frequently than do those generated by the heat-based objective function. Furthermore, the temperature-based calibration resulted in a narrower uncertainty envelope for the temperature. Modeled temperatures from DE generally fall within the GLUE uncertainty bounds.

For all modeled temperature estimates, the upper limit of warming is predicted more accurately than the lower limit of cooling. For the winters of 2009 and 2010, the modeled temperatures overpredict the temperature by between 0.2 and 0.5 °C. The uncertainty in liquid water content associated with the prior distribution of the parameters ranges from 0.05 to 0.28, with annual fluctuations of approximately 0.05 at the

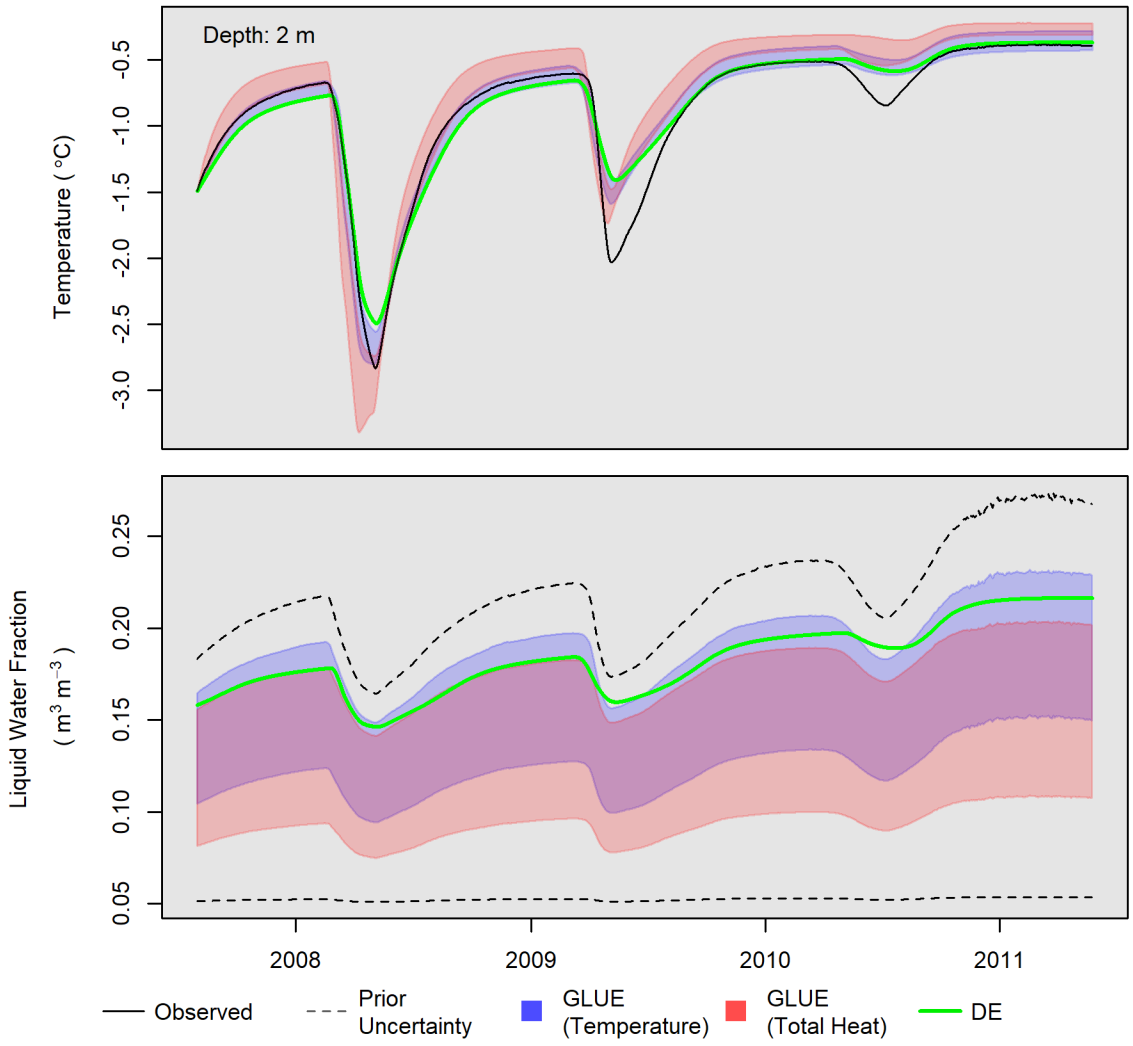

***Figure 38**: Uncertainty envelopes for temperature (top) and liquid water content (bottom) at the Chick Lake borehole, 2 m depth. The 5-95% likelihood uncertainty envelopes are calculated based on the GLUE technique using either the heat-based (red) or the temperature-based (blue) objective functions. Estimates using DE are shown in green.*

upper boundary. The GLUE procedure reduces the with of the uncertainty envelope by approximately one third. The estimates of liquid water content produced by the heat-based objective function are underpredicted relative to those produced by the temperature-based objective function, and the width of the uncertainty envelopes is slightly larger. The DE estimate of liquid water content is within the GLUE uncertainty bounds for the temperature objective function with the exception of two occasions during the winter.

The upper limit increases over the observation period, corresponding to an approximately 5% change in water content. Over the three-year observation period, the amplitude of temperature fluctuation decreases and the temperature gradually increases. One possible explanation is that as the temperature increases, more phase change takes place for the same relative change in temperature because the slope of the SFCC is steeper (Figure 3). Energy that would otherwise change the soil's temperature is instead used to melt or freeze ice. However, the observed annual fluctuations in liquid water content do not appear to increase, and even appear to decrease in 2010. This runs counter to the perhaps intuitive explanation that the decreasing temperature amplitude is caused by an increased latent heat budget. Below 2 m, the calibration is quite poor and the modeled temperature envelopes do not conform to observations (Figure 39).

4.5 Sensitivity of DE to calibration period

DE was run on 13 different calibration periods using data from boreholes at Janssonhaugen to determine the sensitivity of the technique to input data. The estimated parameters are shown in Figure 40. Parameter estimates exhibit a high degree of variability for different calibration periods. The one exception to this is θ_s which has a narrow range of between

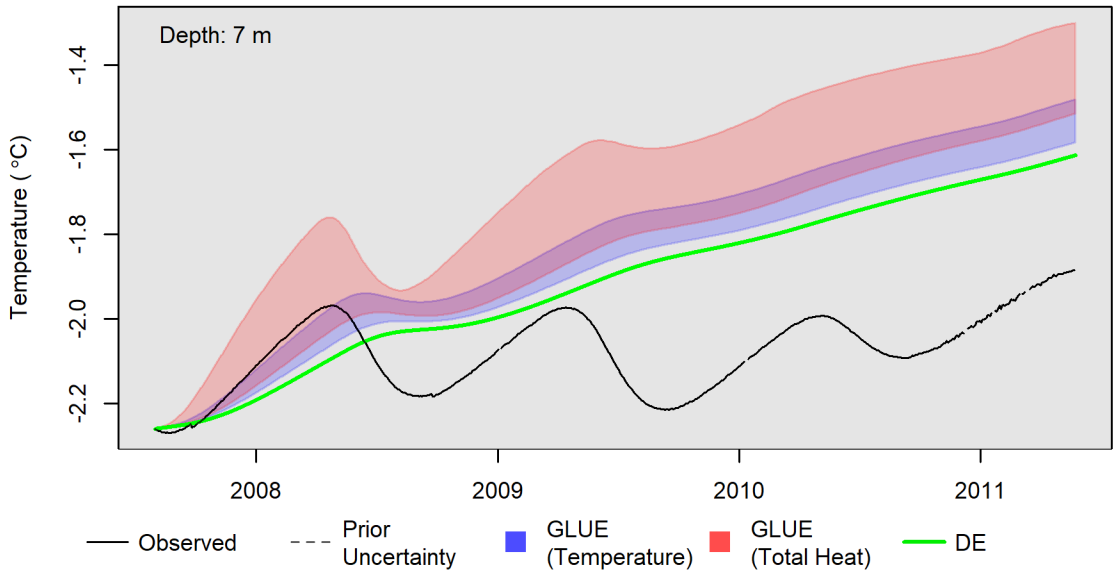

***Figure 39**: GLUE-calibrated temperature envelopes for the Chick Lake site at 7 m depth. Also shown is the estimated temperature using parameters produced by DE (shown in green).*

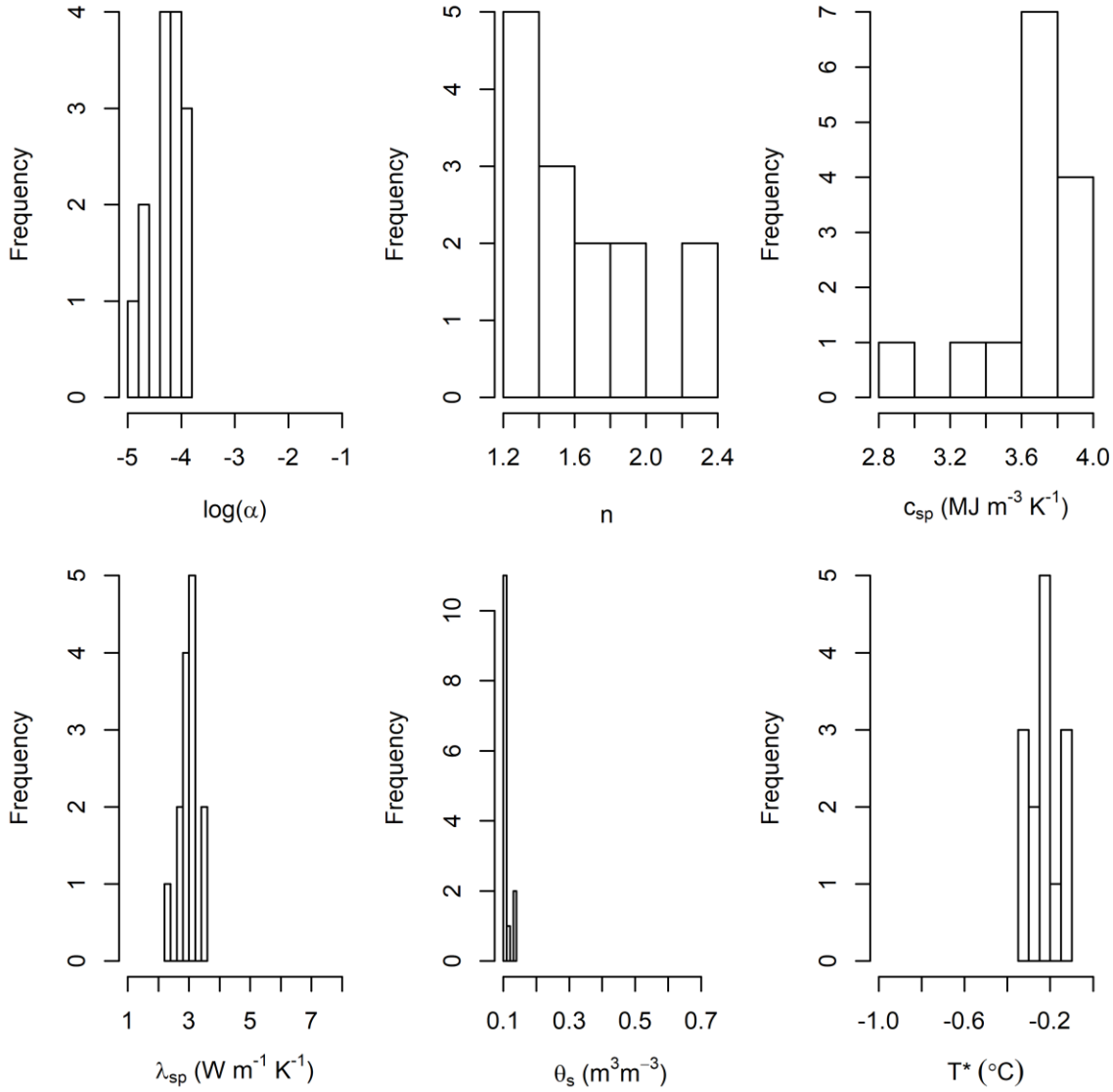

***Figure 40**: Histograms of parameter estimates for Janssonhaugen data using multiple calibration periods. The abscissa has been scaled to the allowable parameter range to demonstrate how well-constrained the collection of estimates is for different calibration data.*

0.10 and 0.14, however, these values are all clustered near the lower calibration limit, suggesting that the true optimal parameter lies outside the calibration bounds and the model structure is inadequate in some way.

Calibrated parameters generally do not match with published values (Isaksen et al. 2000). Measured porosity values are between 20 and 25%, whereas the fitted porosities (θ_s) are between 10 and 14%. Measured bulk thermal capacity is 800 J kg^{-1} K^{-1} which, using a measured density of 2280 kg m^{-3} gives a thermal capacity of 1.82×10^6 J m^{-3} K^{-1}. Using simple volumetric-fraction mixing model, bulk thermal capacities calculated from fitted parameters range from 2.96×10^6 to 4.01×10^6 J m^{-3} K^{-1}.

Using a weighted geometric mean, thermal diffusivities calculated from the fitted parameters range from 1.00×10^{-6} to 1.17×10^{-6} m^2 s^{-1} (frozen) and 0.91×10^{-6} to 1.02×10^{-6} m^2 s^{-1} (thawed). These ranges are consistent with the published mean value of 1.1×10^{-6} m^2 s^{-1}.

When the parameters from the 2001-2004 calibration period were used to drive the model for the full 16 years, the model fit is reasonably good, even outside of the calibrated range. The model error, calculated as the modeled temperatures subtracted from the observed temperatures, are plotted in Figure 41. 90% of model errors fall between -0.3 and 0.5 and 99% of model errors are between -0.7 and 0.9 °C.

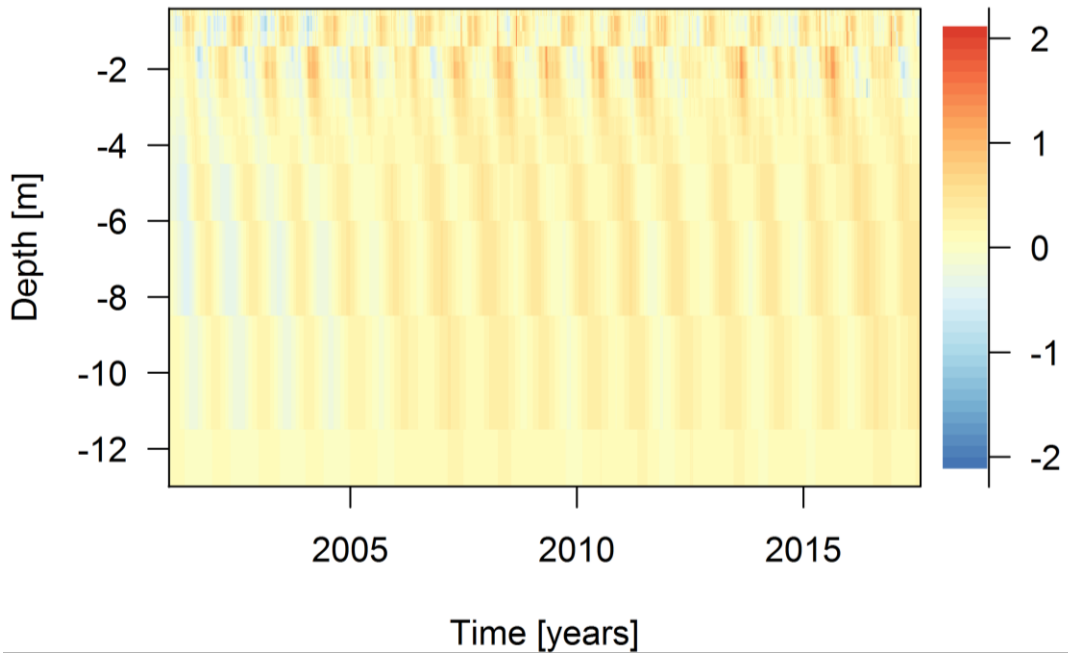

***Figure 41**: Model error for the Janssonhaugen borehole calibrated using DE. The calibrated period was 2001-2003.*

CHAPTER 5: INTERPRETATION

The results of the numerical experiments are interpreted in the context of the four main research questions:

5.1 Objective 1: Estimating model parameters

5.1.1 GLUE

Estimates of the van Genuchten α, θ_s and T* parameters are improved from prior distributions in nearly every simulation. The posterior parameter distributions are affected by changing the cut-off threshold and thereby including poorer-performing simulations which indicates that the model error is sensitive to the parameter values. Estimates of λ_{sp} can be narrowed down significantly when wide parameter bounds are used, with the temperature RMSE producing a slightly more accurate distribution than the heat RMSE. However, with narrow initial parameter bounds, the posterior distributions are uniform indicating that the model output is not sufficiently sensitive to this parameter to narrow down the range any further.

The model is even less sensitive to the c_{sp}. At wide parameter ranges, it can be narrowed somewhat broadly from [1,8] to [1.5, 4], but for narrow initial ranges, the GLUE technique is not able to reduce it any further. For synthetic data with added depth or temperature noise, the posterior distributions are also uniform. It is unclear whether the added noise obscures the estimate because the results are identical to those in the unperturbed case with narrow bounds.

The uniform distributions of λ_{sp} and c_{sp} suggest that the GLUE technique is not effective for reducing the uncertainty on these parameters. These results may seem to

suggest that in future experiments, one or both of these parameters can safely be set to a fixed reasonable value and then ignored. It is important to keep in mind however, that the synthetic data in this example is modeled after an ice-bonded silt, and that the relative importance of the thermophysical properties will be greater in a soil with lower ice content.

In all simulations using both pure and perturbed synthetic data, the van Genuchten *n* parameter was found to have a very low sensitivity. Even when the prior distribution is very wide, the posterior distribution is relatively unchanged.

At Chick Lake and Gibson Lake, the posterior parameter distributions for α and *n* are heavily skewed with the bulk of the distribution concentrated at the lower parameter boundary. This suggests that the global minimum lies outside the parameter bounds as defined by the ROSETTA analysis. This may also suggest that the model is not capable of accurately reproducing the physical behaviour of the system on a more fundamental level.

For both the Chick Lake and Gibson Lake sites, the calibration of temperature measurements at shallower depths was superior to that of deeper measurements. This can be explained by the greater proximity to the surface where the upper boundary condition drives the model. That is to say that, near the upper and lower boundaries of the model temperatures are specified using observational data and so model output nearer to these specified values should be expected to agree more closely with observations. However, another explanation should be considered: if the parameters are somewhat variable with depth, poor calibration near the surface will tend to be associated with higher RMSE values than poor calibration at depth simply due to the larger amplitudes at the surface. Furthermore, in the case of permafrost monitoring, sensors are typically installed in higher

density near the surface as well, meaning that the contribution of error from shallower depths is comparatively higher. Although it is not presented here, an attempt was made to take these problems into consideration by weighting the objective function inversely proportional to the sensor spacing. However, the resulting uncertainty bounds were not found to be significantly improved.

There is evidence to suggest that the model structure is responsible for the inability of estimated parameters to fully reconstruct the observations. At the Chick Lake site, although the model is able to simulate near-surface temperatures, at 7 m depth (Figure 39), all modeled temperatures warm at an average rate of approximately 0.2 °C yr^{-1} but observed temperatures do not show any transient behaviour until 2010 when a slight warming trend is perceptible. This behaviour would be consistent with the melting and runoff of excess ice at the top of permafrost. The melting of excess ice would absorb the energy that, in the modelled temperatures, causes warming at depth. Neither excess ice nor moisture transport were included in the model. In the simulated data, there is an observed dampening of the annual temperature while warming and cooling, likely a consequence of increased phase change at shallower depths. This dampening is not present in the observed data. Runoff of water melted from excess ice would be consistent with this observed lack of dampening because it would no longer be available to release latent heat and increase the apparent heat capacity of the soil. The overprediction of winter temperatures (Figure 38) in two of the three winters is also consistent with the removal of melted water. If melted water was removed from the soil column, this water would not have to re-freeze during the cooling period, permitting the temperature to drop below the freezing point much more

quickly than if it had remained. The warmer winter temperatures in the simulated results are consistent with the melted water refreezing and releasing latent heat during cooling, delaying the drop of temperature below the freezing point. Note that this melting and runoff of water is expected to occur at a shallower depth in the permafrost because at the plotted depth of 2 m, the soil temperatures do not reach high enough values to expect that excess ice would melt in any significant quantity. Furthermore, ice melt at the top of permafrost will be most readily transported.

The sharp distinction between soil parameter estimates in the upper and lower soil layers at Gibson Lake (Figure 30) suggests that a one-layer model may not be entirely appropriate to model the soil column here. For Chick Lake, the parameter estimates are more similar for all depths, so the one-layer model may be more appropriate.

5.1.2 Differential Evolution

Results from the DE experiments with synthetic data suggest that the algorithm is well-suited to finding the best-fitting parameters for the GEOtop model. All 6 parameters of interest were accurately estimated to within 1% of the correct value. Repeated experiments with different realizations of the random initial starting parameters show that the results are stable.

That the DE technique was able to estimate both the c_{sp} and λ_{sp} parameters is somewhat unexpected because temperature dynamics are controlled by the ratio of the two parameters (the thermal diffusivity, κ). In a homogenous medium, the two parameters cannot be distinguished by temperature fluctuations alone. In a soil, which comprises ice and water in addition to soil particles, the bulk thermal diffusivity is strongly temperature

dependent near the freezing point. However, it is not immediately obvious that c_{sp} and λ_{sp} should be able to be recovered from parameter fitting given that the temperature also controls the relative abundance of liquid vs solid water according to the SFCC whose parameters are not known.

It should be expected that for soils with lower ice contents, the ability to determine both c_{sp} and λ_{sp} will be decreased because of the weaker temperature dependence of κ. The synthetic soil profile presented here is modeled after a silt with a relatively high porosity of 49%.

Initially, experiments included θ_r as an estimated parameter but the resulting parameter sets did not converge to appropriate values, even with unperturbed synthetic data. However, the difference between the resulting θ_r and θ_s parameters was found to be approximately correct. This result suggests that it is the interaction between θ_r and θ_s that is being estimated and not necessarily the absolute value of the parameter. This difference corresponds to the total amount of freezable water in the ground and would control the total quantity of latent heat available between the completely frozen and thawed state. If this is the case, and the parameter estimation is only sensitive to changes in the quantity of freezable water and not the absolute water content, then true estimates of the liquid water histories at sites would require an independent estimate of porosity. Otherwise, meaningful interpretation would be limited to the change in liquid water over time but not the total quantity. This would still be useful for estimating ground heat uptake, but would be less informative for engineering applications where the soil strength depends on the absolute quantity of liquid water.

5.1.3 Summary

GLUE is capable of reducing parameter uncertainty for θ_s, and T*. However, the model is not sufficiently sensitive to variations in c_{sp} and *n* to estimate them. The sensitivity of α and λ_{sp} is somewhat intermediate. The sensitivity of the GLUE technique to parameters and the degree to which those parameters may be predicted is somewhat dependent on the ranges of the prior distributions. This means that if the initial level of prior site knowledge is sufficiently detailed, it may not be possible to appreciably improve estimates using GLUE. However, even with very limited knowledge of soil properties, the uncertainty in parameter values can still be improved. For synthetic data, the temperatures calculated from estimated parameters are visually indistinguishable from the input data, except when the heat-based objective function is used (discussed below). DE is able to accurately estimate soil characteristics from synthetic data, and is able to predict c_{sp} and λ_{sp} simultaneously.

Tests with real data suggest that parameter estimates for *n* and λ_{sp} are particularly sensitive to the input data and therefore the estimates likely do not represent site conditions. This may be solved by reducing model structural error.

5.2 Objective 2: Accuracy and stability of parameter estimates

5.2.1 GLUE

The GLUE technique seems robust with respect to noisy observations and forcing data. Only the θ_s parameter is significantly affected by the added noise. The effect of error in the depth is also relatively small with the exception of the θ_s parameter which is overpredicted in the case of the D-BIAS dataset. It is important not to over interpret the result because

this is only one possible realization of the noisy data. However, the total error is relatively small, corresponding to a shift of 0.04 $m^3 m^{-3}$ for the temperature RMSE and 0.06 $m^3 m^{-3}$ for the heat RMSE.

The effect of observational noise on the estimates for liquid water and energy content is almost imperceptible when using synthetic data (Figure 24). This is a likely consequence of the inherent tolerance in GLUE for some degree of model mis-specification. There are two caveats that must be kept in mind with this interpretation: first, the GLUE experiments are a result of only one realization of the noisy data, it is possible that alternate realizations would affect the parameters more severely. Second, the noisy data were only added to the experiment that used prior parameter ranges representative of a well-described site. For experiments with wider initial parameter ranges, it should be expected that the estimated parameter range will be more strongly affected.

5.2.2 Differential evolution

The addition of noise to the dataset results in a more dispersed set of parameter estimates. Visually, the impact of the temperature or depth error on the estimation of the SFCC (Figure 27) is relatively minor which suggests that the parameters are still suitable for the calculation of other variables. However, even for relatively small observation noise, the error in a small neighbourhood of the rising limb of the SFCC (approximately -0.1 to 0 °C in this case) becomes very large. Of course, this range of temperatures will be most relevant when attempting to measure rates of ice melt in isothermal soil columns, but even

for synthetic data, it becomes very difficult to constrain uncertainty over these temperature ranges.

Using the DE algorithm with observed data, the calibrated parameters produce a good fit to observed data, and the majority of the modeled temperatures are within only a few tenths of a degree of the observed temperatures (Figure 41). However, the parameter results depend strongly on the calibration period used and the SFCC curves that are produced do not show the same degree of similarity as with synthetic data. It is also worth mentioning that values for θ_s are clustered at the lower calibration bound of 0.1 whereas the measured porosity is between 0.2 and 0.25.

This technique may be especially useful in the context of measuring soil properties within a soil column experiment where the structure of the soil column and the sensor depth and calibration can be controlled very precisely. Because of the strong effect of slight perturbations on results in the face of model structural error, this technique may not yield accurate soil properties if the soil column stratigraphy is not well-constrained ahead of time (e.g. from borehole logs).

5.2.3 Summary

Parameter estimates from GLUE show a high accuracy when derived using synthetic data. For observed data, this is more difficult to assess because of a lack of measured values with which to compare, but the tendency for parameter estimates to be clustered at the edge of the allowable limits suggests that they may not be accurate for this simulation. Model structural error is believed to be the cause of the discrepancy between real and synthetic data.

The accuracy of parameter estimates from the DE search algorithm is not significantly reduced with the addition of small noise biases in the observational data, although this depends on the parameter. SFCCs estimated from noisy data resemble the correct SFCC for synthetic data. Using different periods of the data for calibration produce a range of different parameter estimates. Soil properties that demonstrate low sensitivity typically have higher variability when different calibration data are used.

5.3 Objective 3: Choice of objective function

Overall, the most significant differences between the two objective functions for synthetic data appear in the θ_s and α parameters. Notably, the heat RMSE tends to minimize the saturated porosity when the initial model parameters are very wide. It may be the case that when many parameters are bad estimates (as is the case when the bounds are so wide), one way to reduce the error is to minimize the latent heat penalty associated with temperature error near the freezing point by reducing the quantity of water available to undergo phase change. This idea will be discussed further below.

For both tests of the GLUE technique with real data, the posterior distribution of the θ_s parameter is strongly dependent on the choice of objective function (Figure 31 and Figure 36). In all cases, the heat RMSE produced distributions that were heavily weighted towards the lower limit of the parameter bounds whereas the temperature RMSE produced distributions that were uniform or slightly weighted towards the upper limit of the parameter bounds. This is interpreted to be an artefact of the calibration procedure. The total heat error associated with a temperature bias near the freezing point is proportional to the total quantity of ice (or water) available to undergo phase change. If θ_s is as small as

possible, the total error will be reduced. It should be noted that even for perturbed synthetic data, this behaviour was not observed, so it is possible that this behaviour arises when the model doesn't represent the system under consideration on a more fundamental level than small errors in depth or temperature values.

The T* parameter is also strongly affected by the choice of objective function. For both fine-grained soil columns (Chick Lake and Gibson Lake), the heat-based RMSE produces distributions that are approximately normally distributed and centered between -0.15 and -0.10 °C. In contrast, the temperature RMSE produces distributions that are heavily weighted to the lower parameter boundary of -0.3 °C. From visual inspection, the freezing point associated with the zero curtain appears to be no colder than -0.2 °C. These unrealistic results demonstrate that, in this instance, the heat RMSE outperforms the temperature RMSE. Although not included here for the sake of brevity, the results of a similar analysis performed on a sandy site produced distributions centered on 0 °C for both objective functions. Intuitively, this makes sense, as the freezing point depression effect in sand should be expected to be less than in finer-grained material assuming there is not significant difference in solute concentration.

5.3.1 Summary

The results in this study suggest that transforming the objective function to weight observations based on the error in total heat content may result in more accurate model behaviour near the freezing point in some circumstances. This is due to more realistic estimates of the freezing point temperature. However, the new objective function does not

provide an overall benefit to either the estimated parameters or the resulting predictions of liquid water or energy content and tends to underpredict the saturation water content.

5.4 Alternate methods of interpreting energy and water content from temperature data

5.4.1 Parameter estimation using pedotransfer functions

Where soil textural data were available, it was possible to provide first-order estimates of soil parameter ranges using pedotransfer functions. Although this technique does not provide an estimate of T*, this was visually estimated from zero curtain periods to within a few tenths of a degree.

Using parameters estimated from ROSETTA, it was possible to calculate uncertainty bounds for the value-added monitoring variables from existing temperature records. When the observed temperature is within the uncertainty of the freezing point (e.g. Figure 33), the liquid water content becomes difficult to quantify, depending mainly on the estimate for the saturation water content. However, for lower temperatures (e.g. Figure 35), these estimates are well-constrained.

When using pedotransfer functions to set calibration bounds, both parameter estimation procedures occasionally yielded estimates that were at the upper or lower limit of these bounds. As discussed above, this is indicative of either model structural error, or inappropriate calibration limits. This result should not necessarily be interpreted as a failure of the pedotransfer functions to produce accurate parameters of a soil sample however. This is because of the scale mismatch between the soil column to which the estimates are being applied and the soil sample from which the textural data were obtained.

In the model, the point samples are being applied to the entire soil column. Consequently, the quantification of water and energy content derived from pedotransfer functions may only be suitable at near the depth of the soil sample or the sensor.

5.4.2 Visual inspection of the temperature profile

Although it remains a qualitative description, some information about energy and liquid water content can be obtained by visual examination of the temperature profile, but this must be done cautiously. In the absence of other information, the reduction in amplitude of the temperature signal at 2.0 m Chick Lake (Figure 38) could easily be explained by the effect of latent heat. One might expect that as the temperature increases and moves into the steeper part of the freezing characteristic curve, more phase change takes place to compensate for the reduction in temperature change. This does not appear to be the case given the estimate of soil water content using estimated soil parameters. Instead, the reduction in temperature amplitude at a shallower depth may be more responsible for the reduction in temperature amplitude at 2.0 m than the increased energy demands of phase change at 2.0 m. More specifically, increased phase change at one depth may reduce the annual temperature amplitude at that depth which subsequently causes a reduction in temperature amplitude at a greater depth even though there is no increased phase change taking place at the greater depth. This is particularly important for the interpretation of temperature time series, where any reduction in the annual temperature amplitude is often attributed to latent heat effects. It may be the case that the associated phase change is some nontrivial distance from the depth of the sensor under investigation.

This observation is also consistent with the estimates of liquid water content associated with the 100-year synthetic time series. Here, the amplitude of annual temperature fluctuations is reduced and eventually disappears entirely at the measured depth as the temperature nears the melting point. However, the bulk of the phase change that takes place at that depth happens long after the annual temperature signal has vanished. This effect may be a consequence of the amplitude dampening being an integrated effect of the column above, but the measured phase change representing only an infinitesimally thin model slice. In any case, it suggests that in the presence of a transient warming signal that induces ice loss the temperature signal must be interpreted carefully with respect to understanding where ice loss is actually occurring.

CHAPTER 6: DISCUSSION

6.1 Parameter sensitivity

In a study examining the effect of parameter uncertainty on permafrost temperature, Harp et al. (2016) found that the van Genuchten α and *n* parameters were highly insensitive and they were not included in subsequent analysis. This was attributed to the small window of time (annual freezing and thawing) during which they play a role. In this study, the *n* parameter was also found to exhibit low sensitivity in the GLUE analysis, making it difficult to calibrate using temperature measurements. However, for synthetic data it was shown that the α parameter exhibits some sensitivity and can be narrowed down using the GLUE technique. In future experiments, it may be worthwhile to conduct sequential parameter fitting similar to Nicolsky et al. (2007b) in which the SWFC parameters are fit using a subset of the data that comprises temperature data from the spring and fall when freezing is taking place. The higher sensitivity of the α parameter may also be attributed to the smaller system under consideration. In the study by Harp et al. (2016), the soil model included a peat layer immediately below the depth of the upper boundary condition. In this study, the upper boundary condition was always within the mineral soil to ensure that a one-layer soil profile was a reasonable assumption. Because of the exclusion of the influence of peat on ground temperatures, which can be very significant, the relative effect of the van Genuchten α parameter may be greater, allowing it to be estimated.

Marmy et al. (2016) used a combination of R^2 and mean error to calibrate permafrost models with ground temperature data using the GLUE method. The R^2 measure of model quality was used in order to capture the effect of latent heat by enforcing the

correct timing of freeze/thaw events. In contrast to the present study, Marmy et al. (2016) found that consideration of R^2 during calibration tended to result in higher estimates for porosity while consideration of the mean error tended to result in lower calibrated porosities. In this study, the use of a heat-based RMSE tended to produce calibrated porosities that were at the lower limit of their allowable bounds when tested against real data. Here, this effect was interpreted to be a consequence of the calibration procedure rather than a physically meaningful result.

This study found that when using GLUE to estimate parameters for Chick Lake and Gibson Lake, some of the resulting distributions were clustered at the upper- or lowermost boundary. This suggests that the optimal parameter set is outside the allowable calibration range and is therefore physically unrealistic. This sort of behaviour should be expected to arise as a result of model structural error. One way to solve this problem is by iteratively changing model structure as well as parameter sets: Atchley et al. (2015) found that by changing the parameterization of soil conductivity from a model that relies on bulk conductivities of dry or unfrozen soil to one that is based on the porosity and the conductivity of the mineral component of the soil, the parameter estimates were more likely to fall within allowable ranges. They also found that calibrating the soil saturation improved the distribution of the parameters for some locations, but worsened it for others. In this study, a model of soil thermal conductivity similar to the improved model of Atchley et al. (2015) was used. Although the saturation of the soil was not calibrated here, this would be possible to implement in future steps using the GEOtop model.

6.2 Equifinality

The concept of equifinality is central to the philosophy of the GLUE technique (Beven & Freer 2001). In short, this idea states that for environmental systems that are sufficiently complex, there can be many different parameter combinations and model structures that produce simulations equally resembling the observed system. A similar problem, often referred to as non-uniqueness, is encountered in geophysical inversion when an observed signal can be the outcome of infinitely many variations on the geometry and composition of the subsurface. For these reasons, some modelers clearly point out that the parameters they are estimating are not intended to be compared against field data, and that although they cause the model to conform to reality, the parameters themselves should not be overinterpreted.

On the other hand, some parameter-fitting research is explicitly designed to replace or augment costly and difficult field measurements by using an easily-obtained measurement to test hypothetical model scenarios (Nicolsky et al. 2007; Dong et al. 2016; Bateni et al. 2012). Whether the discrepancy in belief in the authenticity of fitted parameters is due to differences in the complexity of the systems under consideration or differences in the philosophies of the researcher remains unclear.

This thesis investigates whether or not calibrated soil parameters for the GEOtop model can be treated as alternatives to parameter measurements. In the case of the DE algorithm, the results suggest that although the fitted parameters produce temperatures that correspond well to observations even outside the period of calibration, the values of the parameters are likely not suitable for the independent calculation of other subsurface

properties because of their strong dependence on the calibration period and the large difference from measured values. In the case of the GLUE method, which implicitly accepts the notion of equifinality, the results with synthetic data demonstrate that estimated parameter values consistent with the observed temperature data are also able to reproduce the behaviour of the liquid water and total heat in the soil profile. In other words, the correct behaviour of the liquid water and total heat content are within the range of uncertainty created by the fitted parameter values. For measured data however, the modeled temperatures were not consistently able to reproduce the observed temperatures.

When errors arise in environmental modelling, Beven et al. (2001) argue that they may arise for a number of reasons. The parameterization may be incorrect, but in addition to this, there may be errors arising from model structure, errors in the boundary conditions that drive the model, or errors in the observations against which the model results are compared. In the experiments conducted in this thesis, the effect of error associated with faulty boundary conditions or observations were isolated through the use of synthetic data. These errors, although measurable, were modest; the SFCC profiles still resembled the true value for the DE experiment, and uncertainty envelopes still contained the correct value most if not all of the time in the GLUE experiments. This suggests that much of the error in the model when tested with real data can be attributed to either poor parameterization or model structural error.

6.3 Assumptions and uncertainties:

6.3.1 Model structure

The research presented here is limited in its applicability by a number of assumptions and uncertainties. The first category pertains to the construction of the model. In this study, the soil is assumed to be completely saturated with no excess ice present. This is quite a strong assumption and common for this type of modelling, but it is difficult to reconcile with observations. In reality, it is common for soil saturation to be variable within permafrost environments for a number of reasons. The middle of the active layer is known to be drier as moisture is drawn towards the upper and lower freezing fronts when freezing occurs from both the top and bottom (2-sided freezing) (Mackay 1983) and the top of permafrost may be supersaturated from ice lensing associated with this water movement (Shur et al. 2005). Furthermore, precipitation and infiltration can change the soil moisture content within the active layer quite significantly. As soil saturation decreases, the freezing point is lowered (Wen et al. 2011), so allowing for the calibration of both the depressed freezing point temperature and the saturation water content may still allow for a reasonable approximation of the SFCC by producing a best-fitting curve. However, any error that is 'accommodated' in this way should be expected to affect the estimates of bulk thermal conductivity and heat capacity.

The model presented here also uses a one-layer model for the soil column. This differs from the work presented in other permafrost parameter calibration studies (Nicolsky et al. 2007; Marmy et al. 2016) and this limits the applicability of the results to soil columns that are relatively homogenous over the depths of interest. There are no theoretical

restrictions that prohibit the use of a multi-layer model with the methods presented here, but because of the sampling strategy associated with GLUE samples from all areas of the parameter space, the addition of even a single soil layer would double the number of parameters and decrease the sampling density by several orders of magnitude for the same computational effort. The length of time required to run the DE algorithm also increases non-linearly with increasing numbers of parameters. This computational burden could be eased with the addition of more site information. Measurements of thermal capacity for instance, which are relatively easy to collect from drill samples, could be used instead of allowing it to be a free parameter.

In the case of the borehole data from Chick Lake and Gibson Lake that was used to test the techniques, drilling logs suggest that a one-layer model is appropriate for the depths modeled based on textual unit descriptions and quantitative soil texture measurements where available. In general, visual descriptions of ice content also support this assumption except in the case of Chick Lake where higher ice contents are described below 9 m. Furthermore, the logs do not contain information on soil organic content or volumetric water content which could affect the reliability, particularly for organic-rich horizons near the surface.

The synthetic soil profile used in these experiments was created using parameters typically associated with fine-grained soils. This was done intentionally because of the relatively greater importance of the SFCC in influencing the behaviour of temperature dynamics in such a soil. Furthermore, both of the observed datasets from the Mackenzie valley were collected from boreholes in fine-grained soils. This somewhat limits the

applicability of the results to a broader context. For instance, in the GEOtop model, the sensitivity of the hydraulic parameters has been shown to depend on the soil type (Gubler et al. 2013) and so the ability for the model to estimate those parameters from temperature time series may be stronger or weaker depending on the nature of the substrate.

When considering the sampling of parameter distributions to produce test SFCCs, it is important to note that not all parameter combinations are equally likely. Furthermore, certain combinations of parameters are also not physically realistic. For example, combinations of small n and α can produce SFCCs that are nearly linear. However, allowing for these small values of n may be necessary to represent the full range of behaviour at larger α values.

6.4 Beyond the thermal state

The aim of the soil parameter estimations described here is to better constrain the changes in liquid water content or subsurface heat gain in warming permafrost, particularly at greater depths where the warming signal is less affected by seasonal fluctuations. Unfortunately, it is not possible to estimate SFCC parameters in layers that do not undergo some phase change (Nicolsky et al. 2007). The use of a one-layer soil model, while not without its limitations, does allow the information gained near the surface to inform changes at greater depths.

Although the results with measured data are somewhat discouraging, the results done with synthetic data are hopeful, and suggest that improved monitoring of subsurface properties such as heat and liquid water are feasible to within some margin of uncertainty provided that model structural errors can be minimized. Some parameters were more

difficult to estimate than others. For instance, c_{sp}, λ_{sp} and n were the least sensitive in the GLUE tests, and were most affected by the addition of noise in DE. On the other hand, GLUE was able to reliably improve estimated distributions for α, θ_s and T*, and the effect of noise on the estimates for DE was relatively slight. Priority should be given to measuring those variables that are difficult to estimate. In the case of c_{sp} and λ_{sp}, it may be possible to achieve improvements by measuring only one of them given that their ratio, κ_{sp}, could be estimated somewhat more reliably.

Other studies that estimate parameters with temperature data highlight the utility of soil moisture data against which the calibrated properties can be tested (Marmy et al. 2016; Harp et al. 2016), but note that in many cases these data are not available. Indirect methods of measuring soil moisture such as multi frequency GPR (Westermann et al. 2010) could also be used but would require additional computation. The release of a 20-year record of permafrost ground temperature along with soil moisture content (Boike et al. 2017) may allow for a more thorough test of the methods outlined here as well as new techniques to estimate liquid water contents from temperature time series. Unfortunately, even in this dataset the soil liquid water content data are limited to the upper 1.2 m of the soil column.

CHAPTER 7: CONCLUSION

The aim of this work was to investigate a widely-applicable method of estimating soil characteristics from temperature time series in order to quantify changes to soil liquid water and energy content. The results have demonstrated the utility of GLUE and DE for estimating soil characteristics, and have identified some of the challenges of applying them. The influence of different objective functions to measure model error has also been demonstrated. The main conclusions of this work are summarized below.

7.1 Objective 1: Estimating soil properties

The first objective of this thesis was to use GLUE and DE to fit model parameters in order to determine which of the model parameters could be accurately estimated. When using GLUE to estimate soil parameters for unbiased synthetic data, the estimates for the van Genuchten α, θ_s, and T* parameters are improved from the initial parameter ranges. The λ_{sp}, c_{sp} and van Genuchten *n* parameters show little to no improvement; here the ability of GLUE to reduce parameter uncertainty is somewhat dependent on choice of prior parameter distributions. This is consistent with other studies using GLUE.

The DE search algorithm is able to accurately estimate van Genuchten α and *n*, θ_s, T*, c_{sp} and λ_{sp} with a small error when run with pure synthetic data. This method is also able to simultaneously estimate both the c_{sp} and λ_{sp}, presumably because of the temperature dependence of bulk soil properties below zero. However, both saturated and residual water content cannot both be estimated. It is hypothesized that this is because the model is only sensitive to the total quantity of 'freezable water' i.e. the difference between θ_s and θ_r.

GLUE and DE give comparable results when used to estimate soil parameters from measured data when the same cost function is used; when GLUE results are not clustered at a limiting parameter boundary, DE estimates are consistently within the 5 and 95% likelihood bounds. When GLUE results are clustered at an upper or lower boundary, the approximate modal value is in agreement with the DE estimate.

7.2 Objective 2: Testing the stability of parameter estimates

The second objective of this thesis was to evaluate the stability of parameter estimates with the addition of noise, or when different calibration periods were used from the same site. Adding small measurement noise to the model has little effect on GLUE parameter estimates when tested with synthetic data.

When using the DE algorithm, the effect of added noise to the model shifts the recovered parameter sets away from the true values, but the resulting SFCCs still resemble those constructed from the true parameter values. When run with measured data, the DE search algorithm is able to produce parameters that give a good model fit, but these parameter values are dependent on the period of calibration data. Parameters that exhibit low sensitivity with GLUE, particuarly λ_{sp} and the van Genuchten n parameter, tend to have higher variability when the calibration data is changed. For this reason, more work is needed to determine under which circumstances the DE method is appropriate to use. Even then, parameters should be used cautiously to calculate heat or liquid water contents.

7.3 Objective 3: Comparison of objective functions

The final objective of this thesis was to investigate the effect of an objective function that weighted model error based on the calculated ground energy content instead of comparing

temperature measurements directly. Overall, the results suggest that temperature RMSE is the more reliable objective function. In almost every experiment, the temperature RMSE produced narrower uncertainty bounds that were more likely to contain the observed temperature values. However, the heat-based RMSE may be useful for estimating the freezing point depression, and because of this it produced more realistic modeled temperature values during zero-curtain periods at the Gibson Lake site.

The heat-based RMSE is prone to greatly underestimating θ_s, particularly when the prior distributions are too wide or when the model suffers from structural errors that render it unable to reproduce the observed system behaviour. This is interpreted to be a consequence of the algorithm minimizing error by reducing the freezable water content and therefore also reducing the possible error in heat content.

7.4 Final Thoughts

The overall aim of this thesis was to evaluate parameter estimation techniques GLUE and DE in order to estimate subsurface heat and liquid water contents. Results suggest that estimates are reasonably accurate provided that the model structure is sufficiently well-suited to the site. Because independent moisture content or soil parameter measurements were not available, the accuracy of this method for measured data remains equivocal. Fidelity of simulation results to observations does not guarantee that parameters are representative of site conditions. However, the inability of the model to reproduce observed temperatures may give some indication as to when parameter estimates should not be used. Future studies should test these techniques with simple observations such as those produced by a laboratory soil column experiment.

REFERENCES

Anderson, M.D. & Tice, A.R. 1972. Predicting unfrozen water contents in frozen soils from surface area measurements. *US Highway Research Record* **393**: 12–18.

Atchley, A.L., Painter, S.L., Harp, D.R., Coon, E.T., Wilson, C.J., Liljedahl, A.K. & Romanovsky, V.E. 2015. Using field observations to inform thermal hydrology models of permafrost dynamics with ATS (v0.83). *Geoscientific Model Development* **8**: 2701–2722. DOI: 10.5194/gmd-8-2701-2015

Aubertin, M., Mbonimpa, M., Bussière, B. & Chapuis, R.P. 2003. A model to predict the water retention curve from basic geotechnical properties. *Canadian Geotechnical Journal* **40**: 1104–1122. DOI: 10.1139/T03-054

Azmatch, T.F., Sego, D.C., Arenson, L.U. & Biggar, K.W. 2012. Using soil freezing characteristic curve to estimate the hydraulic conductivity function of partially frozen soils. *Cold Regions Science and Technology* **83**–**84**: 103–109. DOI: 10.1016/j.coldregions.2012.07.002

Banin, A. & Anderson, M.D. 1974. Effects of salt concentration changes during freezing on the unfrozen water content of porous materials. *Water Resources Research* **10**: 124–128. DOI: 10.1029/WR010i001p00124

Bateni, S.M., Jeng, D.S. & Mortazavi Naeini, S.M. 2012. Estimating soil thermal properties from sequences of land surface temperature using hybrid genetic algorithm-finite difference method. *Engineering Applications of Artificial Intelligence* **25**: 1425–1436. DOI: 10.1016/j.engappai.2012.02.017

Beven, K. 2012. *Rainfall Runoff Modelling* 2nd ed., John Wiley & Sons, Ltd: 457 pp. DOI: 10.1002/9781119951001

Beven, K. & Binley, A. 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes* **6**: 279–298. DOI: 10.1002/hyp.3360060305

Beven, K. & Freer, J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* **249**: 11–29. DOI: 10.1016/S0022-1694(01)00421-8

Biskaborn, B.K., Lanckman, J., Lantuit, H., Elger, K., Streletskiy, D.A., Cable, W.L. & Romanovsky, V.E. 2015. The new database of the Global Terrestrial Network for Permafrost (GTN-P). *Earth System Science Data* **7**: 245–259. DOI: 10.5194/essd-7-245-2015

Bittelli, M. 2011. Measuring soil water content: A review. *HortTechnology* **21**: 293–300.

Black, P.B. & Tice, A.R. 1989. Comparison of soil freezing curve and soil water curve data for Windsor sandy loam. *Water Resources Research* **25**: 2205–2210. DOI: 10.1029/WR025i010p02205

Boeckli, L., Brenning, A., Gruber, S. & Noetzli, J. 2012. Permafrost distribution in the European Alps: Calculation and evaluation of an index map and summary statistics. *The Cryosphere* **6**: 807–820. DOI: 10.5194/tc-6-807-2012

Boike, J., Juszak, I., Langue, S., Chadburn, S., Burke, E., Overduin, P.P., Roth, K., Ippisch, O., Bornemann, N., Stern, L., Gouttevin, I., Hauber, E. & Westermann, S. 2017. A 20-year record (1998-2017) of permafrost, active layer, and meteorological conditions at a High Arctic permafrost research site (Bayelva, Spitsbergen): an opportunity to validate remote sensing data and land surface, snow, and permafrost models. *Earth System Science Data Discussions*. DOI: 10.5194/essd-2017-100

Boike, J. & Roth, K. 1997. Time domain reflectometry as a field method for measuring water content and soil water electrical conductivity at a continuous permafrost site. *Permafrost and Periglacial Processes* **8**: 359–370. DOI: 10.1002/(SICI)1099-1530(199710/12)8:4<359::AID-PPP261>3.0.CO;2-S

Bommer, C., Phillips, M. & Arenson, L. 2010. Practical recommendations for planning, constructing and maintaining infrastructure in mountain permafrost. *Permafrost and Periglacial Processes* **21**: 97–104. DOI: 10.1002/ppp.679

Bonnaventure, P.P., Lewkowicz, A.G., Kremer, M. & Sawada, M.C. 2012. A permafrost probability model for the southern Yukon and northern British Columbia, Canada. *Permafrost and Periglacial Processes* **23**: 52–68. DOI: 10.1002/ppp.1733

Brown, J., Ferrians, O., Heginbottom, J.A. & Melnikov, E. 1998. *Circum-Arctic Map of Permafrost and Ground-Ice Conditions.* U.S. Geological Survey in Cooperation with the Circum-Pacific Council for Energy and Mineral Resources. Circum-pacific map series CP-45, scale 1:10,000,000, 1 sheet. Washington, D.C.

Brown, J., Hinkel, K.M. & Nelson, F.E. 2000. The circumpolar active layer monitoring (CALM) program: Research designs and initial results. *Polar Geography* **24**: 166–258. DOI: 10.1080/10889370009377698

Buri, P. 2013. *Simulation of cold-firn-temperatures at an Alpine site using the model GEOtop*. Masters Thesis. University of Zurich, Zurich, Switzerland: 98 pp.

Burn, C.R. 1998. The response (1958-1997) of permafrost and near-surface ground temperatures to forest fire, Takhini River valley, southern Yukon Territory. *Canadian Journal of Earth Sciences* **35**: 184–199. DOI: 10.1139/e97-105

Carslaw, H.S. & Jaeger, J.C. 1959. *The Conduction of Heat in Solids* 2nd ed., Clarendon Press: Oxford; 510 pp.

Cheng, G. & Wang, S. 1983. Distributive regularities of high ice-content permafrost along Qinghai–Xizang Highway. In *Proceedings of the Fourth International Conference on Permafrost, Volume 1*. National Academy Press: Washington D.C.; 142–146.

Christiansen, H.H., Etzelmuller, B., Isaksen, K., Juliussen, H., Farbrot, H., Humlum, O., Johansson, M., Ingeman-Nielsen, T., Kristensen, L., Hjort, J., Holmlund, P., Sannel, A.B.K., Sigsgaard, C., Aakerman, H.J., Foged, N., Blikra, L.H., Pernosky, M.A. & Oedegaard, R.S. 2010. The thermal state of permafrost in the nordic area during the International Polar Year 2007-2009. *Permafrost and Periglacial Processes* **21**: 156–181. DOI: 10.1002/ppp.687

Clein, J.S. & Schimel, J.P. 1995. Microbial activity of tundra and taiga soils at sub-zero temperatures. *Soil Biology and Biochemistry* **27**: 1231–1234. DOI: 10.1016/0038-0717(95)00044-F

Corte, A. 1963. Particle sorting by repeated freezing and thawing. *Science* **142**: 499–501. DOI: 10.1126/science.142.3591.499

Dalgaard, P. 2008. *Introductory Statistics with R* 2nd ed., Springer: New York; 363 pp. DOI: 10.1017/CBO9781107415324.004

Dall'Amico, M., Endrizzi, S., Gruber, S. & Rigon, R. 2011. A robust and energy-conserving model of freezing variably-saturated soil. *The Cryosphere* **5**: 469–484. DOI: 10.5194/tc-5-469-2011

Dallimore, S.R. & Collett, T.S. 1995. Intrapermafrost gas hydrates from a deep core-hole in the Mackenzie Delta, Northwest-Territories, Canada. *Geology* **23**: 527–530.

Davis, J.L. & Annan, A.P. 1989. Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophysical Prospecting* **37**: 531–551. DOI: 10.1111/j.1365-2478.1989.tb02221.x

Dong, J., Steele-Dunne, S.C., Ochsner, T.E. & van de Giesen, N. 2016. Estimating soil moisture and soil thermal and hydraulic properties by assimilating soil temperatures using a particle batch smoother. *Advances in Water Resources* **91**: 104–116. DOI: 10.1016/j.advwatres.2016.03.008

Doré, G., Niu, F. & Brooks, H. 2016. Adaptation methods for transportation infrastructure built on degrading permafrost. *Permafrost and Periglacial Processes* **27**: 352–364. DOI: 10.1002/ppp.1919

Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A. & Jackson, T. 2011. The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences* **15**: 1675–1698. DOI: 10.5194/hess-15-1675-2011

Duan, Q., Sorooshian, S., Gupta, H. V. & Gupta, V. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* **28**: 1015–1031. DOI: 10.1029/91WR02985

Duan, Q., Sorooshian, S. & Gupta, V.K. 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology* **158**: 265–284. DOI: 10.1016/0022-1694(94)90057-4

Dysli, M. 2003. Scaling of supersaturation by a simple test. In *Proceedings of the Eighth International Conference on Permafrost, Volume 1*. M. Phillips, S. M. Springman, & L. U. Arenson, eds. Swets & Zeitlinger: Zurich, Switzerland; 223–227.

Endrizzi, S. & Gruber, S. 2012. Investigating the effects of lateral water flow on the spatial patterns of thaw depth. In *Proceedings of the Tenth International Conference on Permafrost, Volume 1*. K. M. Hinkel, ed. The Northern Publisher (Severnoye Izdatelstvo): Salekhard, Russia; 91–96.

Endrizzi, S., Gruber, S., Dall'Amico, M. & Rigon, R. 2014. GEOtop 2.0: Simulating the combined energy and water balance at and below the land surface accounting for soil freezing , snow cover and terrain effects. *Geoscientific Model Development* **7**: 2831–2857. DOI: 10.5194/gmd-7-2831-2014

Etzelmüller, B., Schuler, T. V., Isaksen, K., Christiansen, H.H., Farbrot, H. & Benestad, R. 2011. Modeling the temperature evolution of Svalbard permafrost during the 20th and 21st century. *The Cryosphere* **5**: 67–79. DOI: 10.5194/tc-5-67-2011

Fiddes, J. & Gruber, S. 2014. TopoSCALE v.1.0: Downscaling gridded climate data in complex terrain. *Geoscientific Model Development* **7**: 387–405. DOI: 10.5194/gmd-7-387-2014

Flerchinger, G.N., Seyfried, M.S. & Hardegree, S.P. 2006. Using soil freezing characteristics to model multi-season soil water dynamics. *Vadose Zone Journal* **5**: 1143–1153. DOI: 10.2136/vzj2006.0025

Fortier, R., LeBlanc, A. & Yu, W. 2011. Impacts of permafrost degradation on a road embankment at Umiujaq in Nunavik (Quebec), Canada. *Canadian Geotechnical Journal* **48**: 720–740. DOI: 10.1139/t10-101

French, H. 2007. *The Periglacial Environment* 3rd ed., John Wiley & Sons, Ltd: Chichester, England; 458 pp.

French, H. & Shur, Y. 2010. The principles of cryostratigraphy. *Earth-Science Reviews* **101**: 190–206. DOI: 10.1016/j.earscirev.2010.04.002

Garcia, F., Folton, N. & Oudin, L. 2017. Which objective function to calibrate rainfall–runoff models for low-flow index simulations? *Hydrological Sciences Journal* **62**: 1149–1166. DOI: 10.1080/02626667.2017.1308511

Gavrilova, M.K. 1993. Climate and permafrost. *Permafrost and Periglacial Processes* **4**: 99–111. DOI: 10.1002/ppp.3430040203

van Genuchten, M.T. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* **44**: 892–898. DOI: 10.2136/sssaj1980.03615995004400050002x

Ghanbarian-Alavijeh, B., Liahgat, A., Huang, G.H. & van Genuchten, M.T. 2010. Estimation of the van Genuchten soil water retention properties from soil textural data. *Pedosphere* **20**: 456–465. DOI: 10.1016/S1002-0160(10)60035-5

Gilbert, G.L., Kanevskiy, M. & Murton, J.B. 2016. Recent Advances (2008–2015) in the Study of Ground Ice and Cryostratigraphy. *Permafrost and Periglacial Processes* **27**: 377–389. DOI: 10.1002/ppp.1912

Goodrich, L.E. 1982. The influence of snow cover on the ground thermal regime. *Canadian Geotechnical Journal* **19**: 421–432. DOI: 10.1139/t82-047

Gruber, S. 2012. Derivation and analysis of a high-resolution estimate of global permafrost zonation. *The Cryosphere* **6**: 221–233. DOI: 10.5194/tc-6-221-2012

Gubler, S., Endrizzi, S., Gruber, S. & Purves, R.S. 2013. Sensitivities and uncertainties of modeled ground temperatures in mountain environments. *Geoscientific Model Development* **6**: 1319–1336. DOI: 10.5194/gmd-6-1319-2013

Gubler, S., Fiddes, J., Keller, M. & Gruber, S. 2011. Scale-dependent measurement and analysis of ground surface temperature variability in alpine terrain. *The Cryosphere* **5**: 431–443. DOI: 10.5194/tc-5-431-2011

Gupta, H. V., Kling, H., Yilmaz, K.K. & Martinez, G.F. 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* **377**: 80–91. DOI: 10.1016/j.jhydrol.2009.08.003

Hansson, K. & Lundin, L.C. 2006. Equifinality and sensitivity in freezing and thawing simulations of laboratory and in-situ data. *Cold Regions Science and Technology* **44**: 20–37. DOI: 10.1016/j.coldregions.2005.06.004

Harp, D.R., Atchley, A.L., Painter, S.L., Coon, E.T., Wilson, C.J., Romanovsky, V.E. & Rowland, J.C. 2016. Effect of soil property uncertainties on permafrost thaw projections: a calibration-constrained analysis. *The Cryosphere* **10**: 341–358. DOI: 10.5194/tc-10-341-2016

Harris, C., Haeberli, W., Vonder Mühll, D. & King, L. 2001. Permafrost monitoring in the high mountains of Europe: The PACE Project in its global context. *Permafrost and Periglacial Processes* **12**: 3–11. DOI: 10.1002/ppp

Harris, S.., French, H.M., Heginbottom, J.A., H, J.G., Ladanyi, B., Sego, D.C. & van Everdingen, R.O. 1998. *Multi-language glossary of permafrost and related ground-ice terms*. National Research Council of Canada Technical Memorandum 142. NRC Research Press, Ottawa, Canada; 142 pp.

Heginbottom, J.A. & Dubreuil, M.A. 1993. A new permafrost and ground ice map for the national atlas of Canada. In *Proceedings of the Sixth International Conference on Permafrost, Volume 1*. G. Cheng, ed. South China University of Technology Press: Beijing, China; 255–260.

Henry, K. & Smith, M. 2001. A model-based map of ground temperatures for the permafrost regions of Canada. *Permafrost and Periglacial Processes* **12**: 389–398. DOI: 10.1002/ppp.399

Hinkel, K.M. & Hurd, J.K. 2006. Permafrost destabilization and thermokarst following snow fence installation, Barrow, Alaska, U.S.A. *Arctic, Antarctic, and Alpine Research* **38**: 530–539.

Hivon, E.G. & Sego, D.C. 1995. Strength of frozen saline soils. *Canadian Geotechnical Journal* **32**: 336–354. DOI: 10.1139/t95-034

Houska, T., Kraft, P., Chamorro-Chavez, A. & Breuer, L. 2015. SPOTting model parameters using a ready-made python package. *PLoS ONE* **10**: 1–22. DOI: 10.1371/journal.pone.0145180

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J.W., Schuur, E.A.G., Ping, C.L., Schirrmeister, L., Grosse, G., Michaelson, G.J., Koven, C.D., O'Donnell, J.A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J. & Kuhry, P. 2014. Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps. *Biogeosciences* **11**: 6573–6593. DOI: 10.5194/bg-11-6573-2014

Humlum, O., Christiansen, H.H. & Juliussen, H. 2007. Avalanche-derived rock glaciers in Svalbard. *Permafrost and Periglacial Processes* **18**: 75–88. DOI: 10.1002/ppp

Isaksen, K., Mühll, D. V., Gubler, H., Kohl, T. & Sollid, J.L. 2000. Ground surface temperature reconstruction based on data from a deep borehole in permafrost at Janssonhaugen Svalbard. *Annals of Glaciology* **31**: 287–294. DOI: 10.3189/172756400781820291

Jafarov, E.E., Marchenko, S.S. & Romanovsky, V.E. 2012. Numerical modeling of permafrost dynamics in Alaska using a high spatial resolution dataset. *The Cryosphere* **6**: 613–624. DOI: 10.5194/tc-6-613-2012

Jones, B.M., Stoker, J.M., Gibbs, A.E., Grosse, G., Romanovsky, V.E., Douglas, T.A., Kinsman, N.E.M. & Richmond, B.M. 2013. Quantifying landscape change in an Arctic coastal lowland using repeat airborne LiDAR. *Environmental Research Letters* **8**: 1–10. DOI: 10.1088/1748-9326/8/4/045025

Jones, E., Oliphant, T. & Peterson, P. 2001. *SciPy: Open Source Scientific Tools for Python*- http://www.scipy.org/. [accessed: 2017-11-02].

Kass, M.A., Irons, T.P., Minsley, B.J., Pastick, N.J., Brown, D.R.N. & Wylie, B.K. 2017. In situ nuclear magnetic resonance response of permafrost and active layer soil in boreal and tundra ecosystems. *The Cryosphere Discussions* **3**: 1–21. DOI: 10.5194/tc-2016-256

Keller, K., Blum, J.D. & Kling, G.W. 2010. Stream geochemistry as an indicator of increasing permafrost thaw depth in an Arctic watershed. *Chemical Geology* **273**: 76–81. DOI: 10.1016/j.chemgeo.2010.02.013

Kleinberg, R.L. 2006. Nuclear magnetic resonance pore-scale investigation of permafrost and gas hydrate sediments. In *New Techniques in Sediment Core Analysis*. R. G. Rothwell, ed. Geological Sociery of London: London; pp. DOI: 10.1144/GSL.SP.2006.267.01.13

Kokelj, S. V., Lantz, T.C., Kanigan, J., Smith, S.L. & Coutts, R. 2009. Origin and polycyclic behaviour of tundra thaw slumps, Mackenzie Delta region, Northwest Territories, Canada. *Permafrost and Periglacial Processes* **20**: 173–184. DOI: 10.1002/ppp

Koopmans, R.W.R. & Miller, R.D. 1966. Soil freezing and soil water characteristic curves. *Soil Science Society of America Journal* **30**: 680–685. DOI: 10.2136/sssaj1966.03615995003000060011x

Kozlowski, T. 2003a. A comprehensive method of determining the soil unfrozen water curves 1. Application of the term of convolution. *Cold Regions Science and Technology* **36**: 71–79. DOI: 10.1016/S0165-232X(03)00007-7

Kozlowski, T. 2003b. A comprehensive method of determining the soil unfrozen water curves 2. Stages of the phase change process in frozen soil-water system. *Cold Regions Science and Technology* **36**: 81–92. DOI: 10.1016/S0165-232X(03)00006-5

Kozlowski, T. 2007. A semi-empirical model for phase composition of water in clay-water systems. *Cold Regions Science and Technology* **49**: 226–236. DOI: 10.1016/j.coldregions.2007.03.013

Krause, P. & Boyle, D.P. 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* **5**: 89–97. DOI: 10.5194/adgeo-5-89-2005

Kurylyk, B.L. & Watanabe, K. 2013. The mathematical representation of freezing and thawing processes in variably-saturated, non-deformable soils. *Advances in Water Resources* **60**: 160–177. DOI: 10.1016/j.advwatres.2013.07.016

Lachenbruch, A.H., Sass, J.H., Marshall, B. V. & Moses, T.H. 1982. Permafrost, heat flow, and the geothermal regime at Prudhoe Bay, Alaska. *Journal of Geophysical Research* **87**: 9301–9316. DOI: 10.1029/JB087iB11p09301

Lamb, R., Beven, K. & Myrabø, S. 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Advances in Water Resources* **22**: 305–317. DOI: 10.1016/S0309-1708(98)00020-7

Lantz, T.C. & Kokelj, S. V. 2008. Increasing rates of retrogressive thaw slump activity in the Mackenzie Delta region, N.W.T., Canada. *Geophysical Research Letters* **35**: 1–5. DOI: 10.1029/2007GL032433

Liu, Z., Zhang, B., Yu, X., Zhang, B. & Tao, J. 2012. A new method for soil water characteristic curve measurement based on similarities between soil freezing and drying. *Geotechnical Testing Journal* **35**: 1–10. DOI: 10.1520/GTJ103653

Lovell, C.W. 1957. Temperature effects on phase composition and strength of partially-frozen soil. *Highway Research Board Bulletin* **168**: 74–95.

Lunardini, V.J. 1996. Climatic warming and the degradation of warm permafrost. *Permafrost and Periglacial Processes* **7**: 311–320. DOI: 10.1002/(SICI)1099-1530(199610)7:4<311::AID-PPP234>3.0.CO;2-H

Mackay, J.R. 1973. A frost tube for the determination of freezing in the active layer above permafrost. *Canadian Geotechnical Journal* **10**: 392–396. DOI: 10.1139/t73-033

Mackay, J.R. 1997. A full-scale field experiment (1978-1995) on the growth of permafrost by means of lake drainage, Western Arctic coast: A discussion of the method and some results. *Canadian Journal of Earth Sciences* **34**: 17–33. DOI: 10.1139/e17-002

Mackay, J.R. 1995. Active layer changes (1968 to 1993) following the forest-tundra fire near Inuvik, N.W.T. *Arctic and Alpine Research* **27**: 323–336. DOI: 10.2307/1552025

Mackay, J.R. 1993. Air temperature, snow cover, creep of frozen ground, and the time of ice-wedge cracking, western Arctic coast. *Canadian Journal of Earth Sciences* **30**: 1720–1729. DOI: 10.1139/e93-151

Mackay, J.R. 1983. Downward water movement into frozen ground, western Arctic coast, Canada. *Canadian Journal of Earth Sciences* **20**: 120–134. DOI: 10.1139/e83-012

Mackay, J.R. 1990. Seasonal growth bands in pingo ice. *Canadian Journal of Earth Sciences* **27**: 1115–1125. DOI: 10.1139/e90-116

Mackay, J.R. & Burn, C.R. 2002. The first 20 years (1978-1979 to 1998-1999) of active-layer development, Illisarvik experimental drained lake site, western Arctic coast, Canada. *Canadian Journal of Earth Sciences* **39**: 1657–1674. DOI: 10.1139/e02-068

Mackay, J.R. & MacKay, D.K. 1974. Snow cover and ground temperatures, Garry Island, N.W.T. *Arctic* **27**: 287–296. DOI: 10.14430/arctic2885

Manache, G. & Melching, C.S. 2004. Sensitivity analysis of a water-quality model using latin hypercube sampling. *Journal of Water Resources Planning and Management* **130**: 232–242. DOI: 10.1061/(ASCE)0733-9496(2004)130:3(232)

Marmy, A., Rajczak, J., Delaloye, R., Hilbich, C., Hoelzle, M., Kotlarski, S., Lambiel, C., Noetzli, J., Phillips, M., Salzmann, N., Staub, B. & Hauck, C. 2016. Semi-automated calibration method for modelling of mountain permafrost evolution in Switzerland. *The Cryosphere* **10**: 2693–2719. DOI: 10.5194/tc-10-2693-2016

Mathews, J.H. & Fink, K.D. 2004. *Numerical methods using MATLAB* 4th ed., Pearson Education Inc.: New York; 680 pp.

McGaw, R.W., Outcalt, S.I. & Ng, E. 1978. Thermal properties and regime of wet tundra soils at Barrow, Alaska. In *Proceedings of the Third International Conference on Permafrost, Volume 1*. National Research Council of Canada: Edmonton, Canada; 47–53.

Mckay, M.D., Beckman, R.J. & Conover, W.J. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**: 239–245. DOI: 10.1080/00401706.1979.10489755

Miller, R. 1966. Phase equilibria and soil freezing. In *Proceedings of the First International Conference on Permafrost*. National Academny of Science-National Research Council: Washington D.C.; 193–197.

Mitchell, J.K. & Soga, K. 2005. *Fundamentals of soil behaviour* 3rd ed., Wiley & Sons Ltd.: Hoboken, New Jersey; 592 pp.

Morishige, K. & Kawano, K. 1999. Freezing and melting of water in a single cylindrical pore: The pore-size dependence of freezing and melting behavior. *Journal of Chemical Physics* **110**: 4867–4872. DOI: 10.1063/1.478372

Moussa, R. & Chahinian, N. 2009. Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events. *Hydrology and Earth System Sciences* **13**: 519–535. DOI: 10.5194/hess-13-519-2009

Nelson, F.E. & Anisimov, O.A. 1993. Permafrost zonation in Russia under anthropogenic climatic change. *Permafrost and Periglacial Processes* **4**: 137–148. DOI: 10.1002/ppp.3430040206

Nelson, F.E., Anisimov, O.A. & Shiklomanov, N.I. 2001. Subsidence risk from thawing permafrost. *Nature* **410**: 889–890. DOI: 10.1038/35073746

Nicolsky, D.J., Romanovsky, V.E., Alexeev, V.A. & Lawrence, D.M. 2007. Improved modeling of permafrost dynamics in a GCM land-surface scheme. *Geophysical Research Letters* **34**: L08501. DOI: 10.1029/2007GL029525

Nicolsky, D.J., Romanovsky, V.E. & Panteleev, G.G. 2009. Estimation of soil thermal properties using in-situ temperature measurements in the active layer and permafrost. *Cold Regions Science and Technology* **55**: 120–129. DOI: 10.1016/j.coldregions.2008.03.003

Nicolsky, D.J., Romanovsky, V.E. & Tipenko, G.S. 2007. Using in-situ temperature measurements to estimate saturated soil thermal properties by solving a sequence of optimization problems. *The Cryosphere* **1**: 41–58. DOI: 10.5194/tc-1-41-2007

Nixon, J.F. 1986. Thermal simulation of subsea saline permafrost. *Canadian Journal of Earth Sciences* **23**: 2039–2046. DOI: 10.1139/e86-188

Osterkamp, T.E. & Romanovsky, V.E. 1997. Freezing of the active layer on the Coastal Plain of the Alaskan Arctic. *Permafrost and Periglacial Processes* **8**: 23–44. DOI: 10.1002/(SICI)1099-1530(199701)8:1<23::AID-PPP239>3.0.CO;2-2

Outcalt, S.I., Arbor, A., Nelson, F.E. & Hinkel, K.M. 1990. The zero-curtain effect: Heat and mass transfer across an isothermal region in freezing soil. *Water Resources Research* **26**: 1509–1516.

Painter, S.L. 2011. Three-phase numerical model of water migration in partially frozen geological media: Model formulation, validation, and applications. *Computational Geosciences* **15**: 69–85. DOI: 10.1007/s10596-010-9197-z

Pan, X., Li, Y., Yu, Q., Shi, X., Yang, D. & Roth, K. 2016. Effects of stratified active layers on high-altitude permafrost warming: A case study on the Qinghai-Tibet Plateau. *The Cryosphere* **10**: 1591–1603. DOI: 10.5194/tc-10-1591-2016

Patil, N.G. & Singh, S.K. 2016. Pedotransfer Functions for Estimating Soil Hydraulic Properties: A Review. *Pedosphere* **26**: 417–430. DOI: 10.1016/S1002-0160(15)60054-6

Patterson, D.E. & Smith, M.W. 1981. The measurement of unfrozen water content by time domain reflectometry: Results from laboratory tests. *Canadian Geotechnical Journal* **18**: 131–144. DOI: 10.1139/t81-012

Pellet, C., Hilbich, C., Marmy, A. & Hauck, C. 2016. Soil moisture data for the validation of permafrost models using direct and indirect measurement approaches at three alpine sites. *Frontiers in Earth Science* **3**: 1–21. DOI: 10.3389/feart.2015.00091

Rempel, A.W., Wettlaufer, J.S. & Worster, M.G. 2004. Premelting dynamics in a continuum model of frost heave. *Journal of Fluid Mechanics* **498**: 227–244. DOI: 10.1017/S0022112003006761

Rigon, R., Bertoldi, G. & Over, T.M. 2006. GEOtop: A distributed hydrological model with coupled water and energy budgets. *Journal of Hydrometeorology* **7**: 371–388. DOI: 10.1175/JHM497.1

Riseborough, D. 2007. The effect of transient conditions on an equilibrium permafrost-climate model. *Permafrost and Periglacial Processes* **18**: 21–32. DOI: 10.1002/ppp

Riseborough, D., Shiklomanov, N., Etzelmüller, B., Gruber, S. & Marchenko, S. 2008. Recent advances in permafrost modelling. *Permafrost and Periglacial Processes* **19**: 137–156. DOI: 10.1002/ppp

Romanovskii, N.N., Hubberten, H.W., Gavrilov, A. V., Eliseeva, A.A. & Tipenko, G.S. 2005. Offshore permafrost and gas hydrate stability zone on the shelf of East Siberian Seas. *Geo-Marine Letters* **25**: 167–182. DOI: 10.1007/s00367-004-0198-6

Romanovsky, V.E., Drozdov, D.S., Oberman, N.G., Malkova, G. V., Kholodov, A.L., Marchenko, S.S., Moskalenko, N.G., Sergeev, D.O., Ukraintseva, N.G., Abramov, A.A., Gilichinsky, D.A. & Vasiliev, A.A. 2010. Thermal state of permafrost in Russia. *Permafrost and Periglacial Processes* **21**: 136–155. DOI: 10.1002/ppp.683

Romanovsky, V.E. & Osterkamp, T.E. 2000. Effects of unfrozen water on heat and mass transport processes in the active layer and permafrost. *Permafrost and Periglacial Processes* **11**: 219–239. DOI: 10.1002/1099-1530(200007/09)11:3<219::AID-PPP352>3.0.CO;2-7

Romanowicz, R.J., Osuch, M. & Grabowiecka, M. 2013. On the choice of calibration periods and objective functions: A practical guide to model parameter identification. *Acta Geophysica* **61**: 1477–1503. DOI: 10.2478/s11600-013-0157-6

Schaap, M.G. & Leij, F.J. 1998. Database-related accuracy and uncertainty of pedotransfer functions. *Soil Science* **163**: 765–779. DOI: 10.1017/CBO9781107415324.004

Schaap, M.G., Leij, F.J. & van Genuchten, M.T. 2001. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology* **251**: 163–176. DOI: 10.1016/S0022-1694(01)00466-8

Schuur, E.A.G., McGuire, A.D., Grosse, G., Harden, J.W., Hayes, D.J., Hugelius, G., Koven, C.D. & Kuhry, P. 2015. Climate change and the permafrost carbon feedback. *Nature* **520**: 171–179. DOI: 10.1038/nature14338

Sellmann, P. V & Hopkins, D.M. 1983. Subsea permafrost distribution on the Alaskan Shelf. In *Proceedings of the Fourth International Conference on Permafrost, Volume 1*. National Academy Press: Fairbanks, Alaska; 75–82.

Sessa, R. & Dolman, H. 2008. *Terrestrial essential climate variables for climate change assessment, mitigation and adaptation*. Global Terrestrial Observing System Report 52, Food and Agriculture Organization of the United Nations. Rome, Italy; 24-25 pp.

Shen, Z.Y., Chen, L. & Chen, T. 2012. Analysis of parameter uncertainty in hydrological and sediment modeling using GLUE method: A case study of SWAT model applied to Three Gorges Reservoir Region, China. *Hydrology and Earth System Sciences* **16**: 121–132. DOI: 10.5194/hess-16-121-2012

Short, N., Brisco, B., Couture, N., Pollard, W., Murnaghan, K. & Budkewitsch, P. 2011. A comparison of TerraSAR-X, RADARSAT-2 and ALOS-PALSAR interferometry for monitoring permafrost environments, case study from Herschel Island, Canada. *Remote Sensing of Environment* **115**: 3491–3506. DOI: 10.1016/j.rse.2011.08.012

Shur, Y., Hinkel, K.M. & Nelson, F.E. 2005. The transient layer: Implications for geocryology and climate-change science. *Permafrost and Periglacial Processes* **16**: 5–17. DOI: 10.1002/ppp.518

Smith, M. & Tice, A. 1988. *Measurement of the unfrozen water content of soils: Comparison of NMR and TDR methods*. CRELL Report 88-18, U.S. Army Corps of Engineers. Hanover, USA; 11 pp.

Smith, M.W. 1975. Microclimatic influences on ground temperatures and permafrost distribution, Mackenzie Delta, Northwest Territories. *Canadian Journal of Earth Sciences* **12**: 1421–1438. DOI: 10.1139/e75-129

Smith, P., Beven, K.J. & Tawn, J.A. 2008. Informal likelihood measures in model assessment: Theoretic development and investigation. *Advances in Water Resources* **31**: 1087–1100. DOI: 10.1016/j.advwatres.2008.04.012

Smith, S. & Brown, J. 2009. *Essential Climate Variables: Permafrost and seasonally frozen ground*. Global Terrestrial Observing System Report 62: ECV-T17, Food and Agriculture Organization of the United Nations. Rome, Italy; 22 pp.

Smith, S.L., Burgess, M.M., Riseborough, D. & Nixon, F.M. 2005. Recent trends from Canadian permafrost thermal monitoring network sites. *Permafrost and Periglacial Processes* **16**: 19–30. DOI: 10.1002/ppp.511

Smith, S.L., Chartrand, J., Nguyen, T., Riseborough, D.W., Ednie, M. & Ye, S. 2009. *Geotechnical database and descriptions of permafrost monitoring sites established 2006-07 in the central and southern Mackenzie Corridor*. Geological Survey of Canada, Open File 6041, CD-ROM. Ottawa; 183 pp.

Smith, S.L., Romanovsky, V.E., Lewkowicz, A.G., Burn, C.R., Allard, M., Clow, G.D., Yoshikawa, K. & Throop, J. 2010. Thermal state of permafrost in North America: A contribution to the International Polar Year. *Permafrost and Periglacial Processes* **21**: 117–135. DOI: 10.1002/ppp.690

Spaans, E.J.A. & Baker, J.M. 1996. The soil freezing characteristic: Its measurement and similarity to the soil moisture characteristic. *Soil Science Society of America Journal* **60**: 13–19. DOI: 10.2136/sssaj1996.03615995006000010005x

Sreelash, K., Sekhar, M., Ruiz, L., Tomer, S.K., Guérif, M., Buis, S., Durand, P. & Gascuel-Odoux, C. 2012. Parameter estimation of a two-horizon soil profile by combining crop canopy and surface soil moisture observations using GLUE. *Journal of Hydrology* **456**–**457**: 57–67. DOI: 10.1016/j.jhydrol.2012.06.012

Storn, R. & Price, K. 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* 341–359. DOI: 10.1023/A:1008202821328

Streletskiy, D.A., Shiklomanov, N.I., Little, J.D., Nelson, F.E., Brown, J., Nyland, K.E. & Klene, A.E. 2017. Thaw subsidence in undisturbed tundra landscapes, Barrow, Alaska, 1962-2015. *Permafrost and Periglacial Processes* **28**: 566–572. DOI: 10.1002/ppp.1918

Stumpp, C., Engelhardt, S., Hofmann, M. & Huwe, B. 2009. Evaluation of pedotransfer functions for estimating soil hydraulic properties of prevalent soils in a catchment of the Bavarian Alps. *European Journal of Forest Research* **128**: 609–620. DOI: 10.1007/s10342-008-0241-7

Tice, A.R., Anderson, D.M. & Sterrett, K.F. 1981. Unfrozen water contents of submarine permafrost determined by nuclear magnetic resonance. *Engineering Geology* **18**: 135–146. DOI: 10.1016/0013-7952(81)90053-3

Tice, A.R., Anderson, M.D. & Banin., A. 1976. *The prediction of unfrozen water contents in frozen soils from liquid limit determinations*. CRREL Report 76-8, US Army Corps of Engineers. Hanover, USA; 9 pp.

Tice, A.R., Burrous, C.M. & Anderson, D.M. 1978. Determination of unfrozen water in frozen soil by pulsed nuclear magnetic resonance. In *Third International Conference on Permafrost*. National Research Council of Canada: Edmonton, Canada; 149–155.

Topp, G.C., Davis, J.L. & Annan, A.P. 1980. Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water Resources Research* **16**: 574–582. DOI: 10.1029/WR016i003p00574

Uhlenbrook, S. & Sieber, A. 2005. On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model. *Environmental Modelling and Software* **20**: 19–32. DOI: 10.1016/j.envsoft.2003.12.006

Vaughan, D.G., Comiso, J.C., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K. & Zhang, T. 2013. Observations: Cryosphere. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. T. F. Stocker, D. Qin, G. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgely, eds. Cambridge University Press: Cambridge, United Kingdom and New York, USA; pp. DOI: 10.1017/CBO9781107415324.012

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M.G. & van Genuchten, M.T. 2010. Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: a review. *Vadose Zone Journal* **9**: 795. DOI: 10.2136/vzj2010.0045

Vonder Mühll, D., Nötzli, J. & Roer, I. 2008. PERMOS – A comprehensive monitoring network of mountain permafrost in the Swiss Alps. In *Proceedings of the Ninth International Conference on Permafrost, Volume 2*. D. L. Kane & K. M. Hinkel, eds. Institute of Northern Engineering, University of Alaska: Fairbanks, USA; 1869–1874. DOI: 10.5167/uzh-4288

Watanabe, K. & Mizoguchi, M. 2002. Amount of unfrozen water in frozen porous media saturated with solution. *Cold Regions Science and Technology* **34**: 103–110. DOI: 10.1016/S0165-232X(01)00063-5

Wen, Z., Ma, W., Feng, W., Deng, Y., Wang, D., Fan, Z. & Zhou, C. 2011. Experimental study on unfrozen water content and soil matric potential of Qinghai-Tibetan silty clay. *Environmental Earth Sciences* **66**: 1467–1476. DOI: 10.1007/s12665-011-1386-0

Westermann, S., Wollschläger, U. & Boike, J. 2010. Monitoring of active layer dynamics at a permafrost site on Svalbard using multi-channel ground-penetrating radar. *The Cryosphere* **4**: 475–487. DOI: 10.5194/tc-4-475-2010

Williams, P.J. 1964. Unfrozen water content of frozen soils and soil moisture suction. *Géotechnique* **14**: 231–246. DOI: 10.1680/geot.1964.14.3.231

Williams, P.J. & Smith, M.W. 1989. *The Frozen Earth: Fundamentals of Geocryology*, Cambridge University Press: Cambridge; 306 pp.

Wolfe, S.A., Stevens, C.W., Gaanderse, A.J. & Oldenborger, G.A. 2014. Lithalsa distribution, morphology and landscape associations in the Great Slave Lowland, Northwest Territories, Canada. *Geomorphology* **204**: 302–313. DOI: 10.1016/j.geomorph.2013.08.014

Wösten, J.H.M., Pachepsky, Y.A. & Rawls, W.J. 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology* **251**: 123–150. DOI: 10.1016/S0022-1694(01)00464-4

Wu, M., Tan, X., Huang, J., Wu, J. & Jansson, P.-E. 2015. Solute and water effects on soil freezing characteristics based on laboratory experiments. *Cold Regions Science and Technology* **115**: 22–29. DOI: 10.1016/j.coldregions.2015.03.007

Yang, X. & You, X. 2013. Estimating parameters of van Genuchten model for soil water retention curve by intelligent algorithms. *Applied Mathematics and Information Sciences* **7**: 1977–1983. DOI: 10.12785/amis/070537

Zhang, T., Heginbottom, J.A., Barry, R.G. & Brown, J. 2000. Further statistics on the distribution of permafrost and ground ice in the Northern Hemisphere. *Polar Geography* **24**: 126–131. DOI: 10.1080/10889370009377692

Zhao, L., Wu, Q., Marchenko, S.S. & Sharkhuu, N. 2010. Thermal state of permafrost and active layer in Central Asia during the International Polar Year. *Permafrost and Periglacial Processes* **21**: 198–207. DOI: 10.1002/ppp.688

APPENDIX I: SYMBOLS USED

α	van Genuchten (1980) water retention parameter	mm^{-1}
n	van Genuchten (1980) water retention parameter	dimensionless
θ_r	Residual water content	dimensionless
θ_s	Saturation water content	dimensionless
L_f	Volumetric latent heat of freezing for water	$J\ m^{-3}$
λ_i	Thermal conductivity of ice	$W\ m^{-1}\ K^{-1}$
λ_i	Thermal conductivity of water	$W\ m^{-1}\ K^{-1}$
λ_{sp}	Thermal conductivity of soil particles	$W\ m^{-1}\ K^{-1}$
λ_a	Thermal conductivity of air	$W\ m^{-1}\ K^{-1}$
κ_{sp}	Thermal diffusivity of soil particles	$m^2\ s^{-1}$
c_i	Thermal capacity of ice	$J\ m^{-3}\ K^{-1}$
c_w	Thermal capacity of water	$J\ m^{-3}\ K^{-1}$
c_{sp}	Thermal capacity of soil particles	$J\ m^{-3}\ K^{-1}$
c_a	Thermal capacity of air	$J\ m^{-3}\ K^{-1}$
θ_i	Volumetric fraction of ice in soil	dimensionless
θ_w	Volumetric fraction of water in soil	dimensionless
θ_{sp}	Volumetric fraction of soil particles in soil	dimensionless
θ_a	Volumetric fraction of air in soil	dimensionless
U	Volumetric internal thermal energy of soil	$J\ m^{-3}$
H_s	Volumetric sensible heat content of soil	$J\ m^{-3}$
H_l	Volumetric latent heat content of soil H_2O	$J\ m^{-3}$
dz_i	Thickness of discretized model cell	m
z	Depth in soil column	m
Δ	Depth of midpoint of discretized soil cell	m
δ	Depth of temperature sensor	m
T*	Freezing temperature of soil water	°C
T	Temperature	°C
$\hat{T}$	Simulated (model) temperature	°C
τ	Spline-interpolated temperature	°C
t	Time	s
G	Number of layers in model discretization	dimensionless
J	Objective function	variable
$\mathcal{C}$	Control vector parameter set	variable
L	Likelihood function	Dimensionless
Q	Heat flux	$W\ m^{-2}$

APPENDIX II: ACRONYMS

ALT	Active Layer Thickness
CALM	Circumpolar Active Layer Monitoring
GSC	Geological Survey of Canada
GTNP	Global Terrestrial Network for Permafrost
NMR	Nuclear Magnetic Resonance
PACE	Permafrost and Climate in Europe
RMSE	Root Mean Squared Error
SFCC	Soil Freezing Characteristic Curve
SWCC	Soil Water Characteristic Curve
TDR	Time-Domain Reflectometry
TSP	Thermal State of Permafrost
GLUE	Generalized Likelihood Uncertainty Estimation