Using Computer Model Uncertainty to Inform the Design of Physical Experiments: An Application in Glaciology

by

Sonja Surjanovic

B.Sc., Simon Fraser University, 2015

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Statistics and Actuarial Science
Faculty of Science

SIMON FRASER UNIVERSITY
Summer 2016

Approval

Name:	**Sonja Surjanovic**
Degree:	**Master of Science (Statistics)**
Title:	***Using Computer Model Uncertainty to Inform the Design of Physical Experiments: An Application in Glaciology***
Examining Committee:	**Chair:** Tim Swartz Professor

Derek Bingham
Senior Supervisor
Professor _____

Gwenn Flowers
Supervisor
Associate Professor
Earth Sciences _____

Dave Campbell
Internal Examiner
Associate Professor _____

Date Defended: 9 August 2016

Abstract

Computer models are used as surrogates for physical experiments in many areas of science. They can allow the researchers to gain a better understanding of the processes of interest, in situations where it would be overly costly or time-consuming to obtain sufficient physical data. In this project, we give an approach for using a computer model to obtain designs for a physical experiment. The designs are optimal for modelling the spatial distribution of the response across the region of interest. An additional consideration is the presence of several tuning parameters to the computer model, which represent physical aspects of the process but whose values are not precisely known. In obtaining the optimal designs, we account for this uncertainty in the parameters governing the system.

The project is motivated by an application in glaciology, where computer models are often used to model the melt of snow and ice across a glacier surface. The methodology is applied to obtain optimal networks of stakes, which researchers use to obtain measurements of summer mass balance (the difference between the amount of snow/ice before and after the melt season).

Keywords: Experimental design; design and analysis of computer experiments; Gaussian process regression; glacier stake networks

Acknowledgements

I would first like to thank my senior supervisor, Dr. Derek Bingham, for guiding me throughout my graduate and undergraduate studies. His advice, guidance and support have helped me be where I am today.

I would also like to thank all of my committee members for their suggestions and advice relating to this project. I am thankful to Dr. Gwenn Flowers of the Earth Sciences department, for all of our discussions and research regarding this interesting application.

Many thanks to my friends and family for their continuing help and understanding during my graduate studies. I am grateful for the encouragement and guidance I have received from them.

Finally, I would like to thank everyone in the Simon Fraser University Department of Statistics and Actuarial Science, for their friendly support and wonderful sense of community. I have always felt that the faculty, staff and graduate students alike have been ready to give a helping hand whenever it was needed. These last few years at SFU have truly been a memorable experience.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Experimental design plays an important role in gathering information in all areas of scientific discovery. Many experiments can be very costly to conduct, and thus the input values, or sampling locations, must be carefully chosen. An improvement in the experimental design can decrease the level of uncertainty in predicting a response value or estimating a parameter of interest, without increasing the cost to the experimenter. As a result, statistical experimental design plays a critical role in most investigations.

In the field of glaciology, physical measurements of ablation (loss of mass due to processes such as melting, evaporation, sublimation and calving) and accumulation (increase in mass mostly due to precipitation) are often taken at a network of stakes placed on the glacier surface (e.g. Østrem and Brugman, 1991). An important question facing scientists is how many stakes are required to obtain a reliable estimate of the spatial distribution of ablation or accumulation across the glacier and how these stakes should be arranged (e.g. Fountain and Vecchia, 1999). The common approach for selecting locations is to either arrange the stakes somewhat uniformly across the glacier surface (e.g. Kuhn et al., 1999; Hock and Jensen, 1999) or to place them along a centre line following the longitudinal axis of the glacier (e.g. Østrem and Brugman, 1991; Kaser et al., 2003; Thibert et al., 2008). Aside from the work done by Fountain and Vecchia (1999), where the standard error of estimating the total mass balance (using regression and piecewise linear spline models) was investigated for several different stake network configurations, obtaining a design that optimizes statistical properties (e.g. minimizes prediction error) has not been the focus of these standard approaches.

For many scientific problems, computer models can be used to gain information on the processes of interest while avoiding costly physical experiments. In glaciology, computer models are often used to model the melt across the glacier. Results from the computer model can be used in place of, or together with, physical data from the stakes (e.g. Wheler et al., 2014; Pellicciotti et al., 2005).

In this project, we undertake the novel approach of using the computer model to aid in the design of the physical experiment. The methodology will be generally applicable to settings where a computational model is available to act as a surrogate for observations in the field. In the glaciology context, this amounts to using a computer model for melt to obtain a stake network design that is optimal for modelling the spatial distribution of melt across the glacier surface. An issue that is met in the methodology is that several tuning parameters to the computer model are not precisely known. Indeed, these parameters can vary from glacier to glacier and also from year to year. Consequently, a design that is optimal for one glacier may not be optimal for another. Additionally, a design that is optimal for a specific glacier in one year may not be optimal for another year. In this project, a new design criterion is proposed that attempts to incorporate the uncertainty in the parameters governing the system. The resulting designs will have good properties (here we focus on prediction properties) across values of the uncertain parameters.

The project is organized as follows. The motivating problem, stemming from mass balance studies on glaciers, is described in Chapter 2. In addition, the computer model for glacier melt is presented. Chapter 3 describes kriging, or Gaussian process regression, which is the statistical model that will be used to obtain predictions of the response across the region of interest. In Chapter 4, we introduce a new design criterion that quantifies the average uncertainty of prediction across the region of interest, taking into account the uncertainty in the computer model parameters. Chapter 5 presents the overall approach used to obtain the optimal designs. Results for several test functions and for the glacier application are given in Chapters 6 and 7, respectively. Finally, a discussion of the results and further work is given in Chapter 8.

Chapter 2

Motivating Problem

Understanding the mass balance (the difference between accumulation and ablation) for the Earth's glaciers is important for prediction of sea-level changes and managing water resources (e.g. Allen et al., 2014; Lemke et al., 2007; Kaser et al., 2010; Moore et al., 2009). This is particularly important in the age of global warming and will have important consequences of interest to scientists and policy makers. Measurements of snow and ice melt, acquired through field studies, are needed to constrain the models that estimate and project glacier contribution to water resources and sea level (e.g. Radić et al., 2013).

The difference between accumulation and ablation that occurs during the summer season is referred to as *summer mass balance*. A common method for obtaining field measurements of (summer) mass balance is the *glaciological*, or *in situ*, method, whereby a network of stakes is drilled into the glacier surface, and the height of snow/ice is measured on the stake at the beginning and end of the melt season. The difference in heights, together with measurements of snow density obtained at several snow pits, is used to obtain the summer mass balance measurement at each stake (measured in metres water equivalent, i.e. m w.e.) (Østrem and Brugman, 1991; Kaser et al., 2003).

However, obtaining the physical data is a labour-intensive process. Thus, a judicious choice of stake locations is required to use the available sampling resources efficiently. In many studies, stakes are distributed somewhat uniformly across the entire surface of the glacier (e.g. Kuhn et al., 1999; Hock and Jensen, 1999). However, due to the spatial nature of the data and the high correlation between observations taken at nearby stakes, there might be some redundancy involved with this approach (Kuhn et al., 1999). The most common alternative approach is to arrange the stakes along a centre line, following the longitudinal axis of the glacier. Several transverse lines may also be added at right angles to the centre line (e.g. Østrem and Brugman, 1991; Kaser et al., 2003; Thibert et al., 2008). Fountain and Vecchia (1999) and Kaser et al. (2003) also suggest that stakes should be arranged so as to intermittently sample as much of the elevation as possible.

The main goal of this project, in terms of the glaciology application, is to formally address the question of optimal stake network designs. Specifically, for any fixed number of stakes n, we seek to obtain the stake network design that minimizes the uncertainty associated with predicting the spatial distribution of melt across the surface of the glacier. A physically-based computer model is used to aid the design search, by representing the "true" melt across the glacier, up to experimental error. Evaporation and sublimation are neglected in the computer model, and melt is treated as a good approximation to ablation.

The model is given in § 2.2. However, it has several unknown parameters. The uncertainty in the parameters is accounted for using the methodology described in Chapters 4-5.

A secondary goal is to decide on the number of stakes n that is required to predict the spatial distribution of melt across the glacier with a satisfactory level of precision. The question of stake number can be of considerable significance to the researcher due to the cost of each additional stake. This issue has been addressed by several researchers, such as Fountain and Vecchia (1999). Using the methodology given in this project, by visually analyzing the relationship between n and the prediction uncertainty of the optimal n-stake network design, the researcher can decide on a satisfactory n at which the decrease in prediction uncertainty becomes less significant than the cost of an additional stake.

2.1 Study Area and Field Data

For our stake network design study, we focus on a glacier in the Donjek Range of the St. Elias Mountains in southwestern Yukon, Canada, located at 60°50′N, 139°10′W (Wheler et al., 2014). Figure 2.1 shows the study area[1]. The glacier of interest has an area of 5.3 km^2 and ranges in elevation from 1970 to 2960 m above sea level (Flowers et al., 2011). This glacier is not yet named. However, in previous studies it has been referred to as "South Glacier", distinguishing it from other nearby sites that were studied by the same authors (e.g. Wheler et al., 2014; MacDougall et al., 2011; Schoof et al., 2014). We adopt this name here as well.

Wheler et al. (2014) give a detailed description of the procedure used to acquire the relevant field data. For this design study, we model the melt using data obtained during the 2012 melt season (May through September). Temperature records were obtained every five minutes during the melt season, from an automatic weather station (AWS) in the ablation area of South Glacier. The elevation of the weather station is approximately 2300 m above sea level. The measurements were averaged to give values at one-hour intervals.

A Digital Elevation Model (DEM) was constructed on a discretized version of the glacier surface, defined by a 30 m × 30 m grid. The DEM was constructed using real-time kine-

[1]Figure 2.1 is Figure 1a of Flowers et al. (2011) (`http://www.the-cryosphere.net/5/299/2011/`), published in *The Cryosphere* under the Creative Commons Attribution 3.0 License (`https://creativecommons.org/licenses/by/3.0/legalcode`).

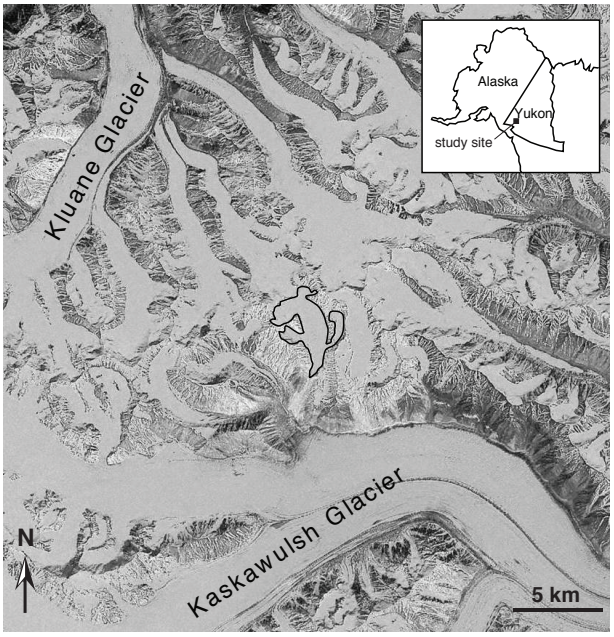

Figure 2.1: Donjek Range, with an outline of South Glacier.

matic Global Positioning System measurements and digitized map contours (De Paoli and Flowers, 2009; Wilson et al., 2013). Using the DEM, the temperature measurements were extrapolated from the station to all points on the grid using the constant temperature-elevation lapse rate -6.5 °C km^{-1}, which is a common value adopted in the literature (Wheler et al., 2014; Minder et al., 2010). On the same grid, values of potential direct solar radiation were computed at one-hour intervals, using the Solar Analyst tool in ArcGIS (Wheler et al., 2014; Fu and Rich, 2000).

The initial snow depth (in m w.e.) was calculated at the beginning of the melt season using field data at a network of stakes. At each stake location, the *winter mass balance* (the difference between accumulation and ablation that occurs during the winter season) was calculated using the snow depth measured at the stake and the snow densities observed in several snow pits. These observations were then extended to each grid point i on the glacier surface using the DEM and the linear relationship

$$b_{w,i} = \beta_0 + \beta_1 z_i + \beta_2 m_i + \beta_3 c_i, \tag{2.1}$$

where $b_{w,i}$ is the winter mass balance at point i, and z_i, m_i and c_i are the elevation, slope and curvature at point i, respectively (Wheler et al., 2014).

Field data were also collected to account for summer accumulation, since summer snowfall can affect melt rates (Wheler et al., 2014). Accumulation events were measured using an ultrasonic depth gauge at the weather station, and were then extrapolated to all points on

the discrete glacier grid, under the assumption of a constant accumulation-elevation lapse rate of 2 cm w.e. km^{-1} (Wheler et al., 2014).

The firn line of a glacier is defined to be the line that "marks the transition between exposed glacier ice and the snow-covered surface of a glacier" (U.S. Department of the Interior, 2013). For South Glacier, the firn line is set at an elevation of 2450 m above sea level, based on field observations (Wheler et al., 2014).

2.2 Melt Model

To aid in obtaining the optimal designs, a physically-based computer model is used to act as a surrogate for the melt across the glacier surface. A commonly used model for glacier snow- and ice-melt is the Hock (1999) enhanced temperature-index melt model, which utilizes the relationship between melt and temperature while incorporating a term to account for radiation (e.g. Schuler et al., 2002; Huss et al., 2007; Zhang et al., 2008; Wheler et al., 2014). The model relies on the field data described in § 2.1.

The computer code for the model was provided by the Simon Fraser University Glaciology Group. On the 30 m × 30 m grid, cumulative melt (i.e. total melt over the course of the melt season) is obtained at each grid point by calculating melt at one-hour intervals and taking the sum of these values. The melt rate M (mm w.e. h^{-1}) is calculated according to the model

$$M = \begin{cases} \left(\frac{1}{n} MF + a_{\text{snow/ice}} I\right) T & \text{if } T > 0 \\ 0 & \text{if } T \leq 0, \end{cases} \tag{2.2}$$

where $n = 24$ is the number of time-steps per day, MF is the temperature melt factor (mm w.e. day^{-1} $^\circ\text{C}^{-1}$), a_{snow} and a_{ice} are the radiation coefficients for snow and ice surfaces (mm w.e. h^{-1} $^\circ\text{C}^{-1}$ m^2 W^{-1}), I is the potential direct clear-sky solar radiation (W m^{-2}), and T is the air temperature ($^\circ$C) (Hock, 1999).

Hourly records of temperature T and potential direct clear-sky solar radiation I at each grid point were obtained as described in § 2.1. At every time-step, the model takes into account the amount of snow present at each grid point, based on the initial snow depth at the beginning of the melt season, the amount of previously melted snow, and the amount of new snow resulting from summer accumulation. At points in time when there is snow remaining at a given grid point, the radiation coefficient a_{snow} is used in (2.2). When there is no snow remaining at a grid point, the coefficient a_{ice} is used instead. However, the snow is taken to be arbitrarily deep above the firn line (2450 m above sea level) (Wheler et al., 2014).

The model in (2.2) contains several tuning parameters whose values are not known, and which may vary from glacier to glacier: MF, a_{snow} and a_{ice}. A joint distribution is proposed on these parameters, based on values used for various glaciers in the literature. Details are given in § 7.1.

2.3 Focus on a Rectangle

One of the complicating factors that must be taken into consideration is the non-convexity of the glacier. In this setting, when constructing a spatial model for the response surface of glacier melt, the Euclidean distance may not provide an accurate representation of the true closeness of two locations.

To simplify the problem, we assume for the purpose of this project that the region of interest, on which we wish to model the response (which is the melt in this case), is a rectangular region. With several slight modifications, the methodology can be applied to any arbitrary convex shape. In future work, as discussed in Chapter 8, it can be applied to the full glacier using the methods presented in Pratola et al. (2015).

The rectangular region of interest on South Glacier is defined by UTM Easting values between 601 783 m and 602 563 m, and UTM Northing values between 6 743 283 m and 6 744 423 m. This region was chosen so as to contain a significant portion of the central part of the glacier surface, while preserving the overall direction of elevation increase. The glacier, along with its elevation contours and the rectangular region of interest, is shown in Figure 2.2. The rectangular region lies entirely below the firn line.

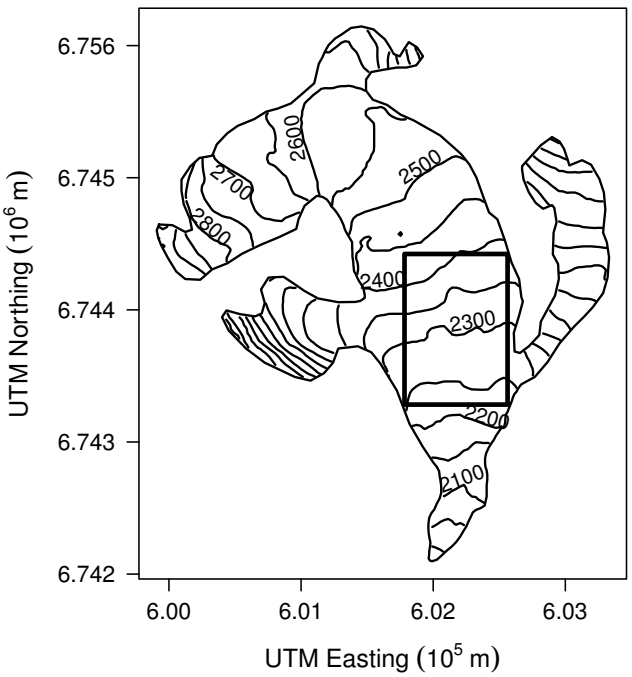

Figure 2.2: South Glacier, with elevation contours (m above sea level) and the rectangular region of interest.

Chapter 3

Kriging

The primary goal of this project is to propose optimal experimental designs using a computer model with uncertain inputs. In finding the optimal design, we must consider the modelling approach that will be used to predict response values across the region of interest.

We assume that the data, consisting of evaluations of the computer model at a set of design locations, are used to fit a *kriging* model to the surface. Kriging, or Gaussian process regression, is commonly used for modelling geostatistical data and output from computer experiments (e.g. Diggle and Ribeiro Jr., 2007; Sacks et al., 1989a,b; Cressie, 1993).

In terms of the glacier application, response values on a glacier surface, such as melt, ablation, accumulation and mass balance are often spatially correlated (e.g. Cogley, 1999), and thus this spatial modelling approach is commonly used in glacier studies. Several examples of studies that have used kriging to model values such as summer balance, winter balance, ablation and accumulation include Hock and Jensen (1999), Jansson and Pettersson (2007), Holmlund et al. (2005), Bales et al. (2001) and Jansson (1999).

3.1 Notation

Throughout the project, it is assumed that the region of interest is a d-dimensional hyperrectangle. For the glacier application, this is assumed to be the rectangle on the glacier surface defined by the coordinates in § 2.3 (see Figure 2.2), and thus there are $d = 2$ spatial dimensions: Easting and Northing. For simplicity, the region of interest is scaled to $\mathcal{A} = [0, 1]^d$. Then, the spatial location of a point in the domain is denoted by $\boldsymbol{x} = [x_1, \ldots, x_d]^\top$. At any $\boldsymbol{x}$ in the region of interest $\mathcal{A} = [0, 1]^d$, $Z(\boldsymbol{x})$ is the response value of interest.

A design $\boldsymbol{X}$ is defined to be a set of n design points in $\mathcal{A}$,

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1^\top \\ \vdots \\ \boldsymbol{x}_n^\top \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nd} \end{bmatrix}, \tag{3.1}$$

where x_{ij} is the j^{th} dimension of the i^{th} observed location (for $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, d\}$). The vector of response values observed at the design points in $\boldsymbol{X}$ is denoted by $\boldsymbol{z} = [Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n)]^\top$. Overall, the aim is to build a predictive model for $Z(\boldsymbol{x})|\boldsymbol{z}$ at any location $\boldsymbol{x} \in \mathcal{A}$. The predictor is denoted by $\hat{Z}(\boldsymbol{x})$, and is dependent on the data $(\boldsymbol{X}, \boldsymbol{z})$.

In some applications, it may be beneficial to incorporate a small additive error term in the observed response values, as a way to account for measurement error in the data. This error term is referred to as a *nugget*. Throughout this project, synthetic data are obtained using a deterministic computer model, and thus it is assumed that there is no measurement error. However, a very small nugget is incorporated for purely computational reasons because it can reduce the numerical instability caused by the inversion of an otherwise nearly singular covariance matrix (Santner et al., 2003; Peng and Wu, 2014).

3.2 The Model

The response $Z(\boldsymbol{x})$ is assumed to be a second-order stationary Gaussian process. In other words, for any set of locations $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \subset \mathcal{A}$, the corresponding responses $\{Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n)\}$ are jointly multivariate Gaussian (Santner et al., 2003), and for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{A}$,

$$\mathbb{E}\left[Z(\boldsymbol{x})\right] = \mu, \tag{3.2}$$

$$\text{Var}\left[Z(\boldsymbol{x})\right] = \sigma^2, \text{ and} \tag{3.3}$$

$$\text{Corr}\left[Z(\boldsymbol{x}), Z(\boldsymbol{x}')\right] = \rho(\boldsymbol{x} - \boldsymbol{x}'), \tag{3.4}$$

for a mean μ, variance σ^2 and positive-definite correlation function $\rho(\cdot)$ (Cressie, 1993). Note that, under the second-order stationarity assumption, the correlation between the response values at two locations is only a function of the difference vector $\boldsymbol{x} - \boldsymbol{x}'$. In many cases, it is further assumed that the correlation is *isotropic*, meaning that it is only a function of the distance $||\boldsymbol{x} - \boldsymbol{x}'||$ between the locations (Cressie, 1993); however, this assumption is not required.

There are many valid correlation functions to choose from. For this project, the Matérn correlation function is used:

$$\rho(\boldsymbol{x} - \boldsymbol{x}') = \prod_{j=1}^{d} \frac{2^{1-\kappa}}{\Gamma(\kappa)} \left(\frac{\sqrt{2\kappa}|x_j - x_j'|}{\theta_j} \right)^{\kappa} K_\kappa \left(\frac{\sqrt{2\kappa}|x_j - x_j'|}{\theta_j} \right), \tag{3.5}$$

for correlation parameters $\theta_1, \ldots, \theta_d > 0$, where $K_\kappa(\cdot)$ is the modified Bessel function (Rasmussen and Williams, 2006). A commonly used value of the Bessel function parameter is $\kappa = 3/2$, and it can be shown that for this value the correlation function is

$$\rho(\boldsymbol{x} - \boldsymbol{x}') = \prod_{j=1}^{d} \left(1 + \frac{\sqrt{3}|x_j - x_j'|}{\theta_j}\right) \exp\left(-\frac{\sqrt{3}|x_j - x_j'|}{\theta_j}\right) \tag{3.6}$$

(Rasmussen and Williams, 2006). Thus, a larger value of θ_j in any dimension generally implies a larger degree of spatial correlation between response values in that dimension. Other correlation functions (such as the Gaussian correlation function), can also be used to suit the application. However, the Matérn correlation function is chosen here due to its flexibility on the smoothness assumptions for the response surface (Rasmussen and Williams, 2006; Diggle and Ribeiro Jr., 2007).

3.3 Parameter Estimation

The aim is to find optimal designs that are helpful when using the kriging model for prediction. To use the model, the parameters that govern the model must be estimated. The kriging parameters that need to be estimated are the mean μ, variance σ^2 and correlation parameters $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_d]^\top$. The parameters are estimated using maximum likelihood (e.g., see Diggle and Ribeiro Jr., 2007; Jones et al., 1998).

For the Gaussian process specified in § 3.1-3.2, the likelihood function for the observations is

$$L(\mu, \sigma^2, \boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{n}{2}} \det(\sigma^2 \boldsymbol{V})^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{z} - \mu\mathbf{1})^\top (\sigma^2 \boldsymbol{V})^{-1} (\boldsymbol{z} - \mu\mathbf{1})\right), \tag{3.7}$$

where $\det(\cdot)$ denotes the determinant of a matrix, and $\mathbf{1}$ is an n-vector of 1's. In (3.7), $\sigma^2\boldsymbol{V}$ is the covariance matrix for the data $\boldsymbol{z} = [Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n)]^\top$. If there is no nugget term (i.e. random error) added to the responses, then $\boldsymbol{V}$ is simply the n-by-n matrix whose $(i,j)^{th}$ element is $\rho(\boldsymbol{x}_i - \boldsymbol{x}_j)$. However, even though there is no measurement error in the model we are using, a small nugget term is added purely for computational reasons, as discussed in § 3.1. In this case, the value ν^2 is added to the diagonal of $\boldsymbol{V}$, where ν^2 is a small value needed for numerical stability.

The log-likelihood function can thus be expressed as

$$\ell(\mu, \sigma^2, \boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log(\det(\boldsymbol{V})) - \frac{1}{2\sigma^2}(\boldsymbol{z} - \mu\mathbf{1})^\top \boldsymbol{V}^{-1}(\boldsymbol{z} - \mu\mathbf{1}). \tag{3.8}$$

Note that $\boldsymbol{V}$ is a function of $\boldsymbol{\theta}$. Assuming $\boldsymbol{\theta}$ is known, the maximum likelihood estimates of μ and σ^2 can be written as

$$\hat{\mu}(\boldsymbol{V}) = \left(\mathbf{1}^\top \boldsymbol{V}^{-1}\mathbf{1}\right)^{-1} \mathbf{1}^\top \boldsymbol{V}^{-1}\boldsymbol{z}, \text{ and} \tag{3.9}$$

$$\hat{\sigma}^2(\boldsymbol{V}) = \frac{1}{n}\left(\boldsymbol{z} - \hat{\mu}(\boldsymbol{V})\mathbf{1}\right)^\top \boldsymbol{V}^{-1}\left(\boldsymbol{z} - \hat{\mu}(\boldsymbol{V})\mathbf{1}\right). \tag{3.10}$$

Substituting these maximum likelihood estimates into (3.8) yields the *concentrated log-likelihood function*, which is a function of only $\boldsymbol{\theta}$.

Using a numerical optimization procedure, the concentrated log-likelihood function is maximized to obtain an estimate of $\boldsymbol{\theta}$ (numerical details of the optimization procedure are provided below). Then, by substituting the estimate of $\boldsymbol{\theta}$ back into (3.9) and (3.10), the estimates of μ and σ^2 are obtained.

3.3.1 Optimization of the Concentrated Log-Likelihood

Optimization of the concentrated log-likelihood function is treated as a constrained optimization problem. For the correlation parameter in the j^{th} dimension, denoted by θ_j ($j \in \{1, \ldots, d\}$), the constraint is

$$\theta_j \in \left[10^{-10}, 2\left(\max_{i=1,\ldots,n}(x_{ij}) - \min_{i=1,\ldots,n}(x_{ij})\right)\right], \tag{3.11}$$

where x_{ij} is the value of the i^{th} observation in the j^{th} dimension (as in Roustant et al., 2012).

The optimization is conducted using functions in the R package `stats` (R Core Team, 2015). For $d \geq 2$, we use the function `constrOptim()`, using the Nelder-Mead optimization algorithm. For cases where $d = 1$ (such as in one of the examples in Chapter 6), the function `optimize()` is used instead, since a warning message states that the Nelder-Mead method may be unreliable in one dimension. This function uses a "combination of golden section search and successive parabolic interpolation" (R Core Team, 2015). In either case, the optimization is randomly restarted three times, and the overall best set of kriging parameter values is taken to be the optimal set.

3.4 Prediction

We are interested in predicting the value of the response at any new location in the region of interest $\mathcal{A} = [0, 1]^d$. The prediction uncertainty is measured by the Mean Squared Prediction Error (MSPE). The MSPE for predicting the response $Z(\boldsymbol{x})$ at any point $\boldsymbol{x} \in \mathcal{A}$, using a predictor $\hat{Z}(\boldsymbol{x})$, is

$$\text{MSPE}(\boldsymbol{x}) \triangleq \mathbb{E}\left[\left(Z(\boldsymbol{x}) - \hat{Z}(\boldsymbol{x})\right)^2\right], \tag{3.12}$$

where the expectation is taken over the joint distribution of $Z(\boldsymbol{x})$ and $\hat{Z}(\boldsymbol{x})$, i.e. the joint distribution of $Z(\boldsymbol{x})$ and the data $\boldsymbol{z} = [Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n)]^\top$. Result 3.1, obtained from Diggle and Ribeiro Jr. (2007), gives a result on obtaining the predictor that minimizes the MSPE at any given location $\boldsymbol{x}$.

Result 3.1. *At any $\boldsymbol{x}$, the predictor that minimizes the MSPE is $\hat{Z}(\boldsymbol{x}) = \mathbb{E}[Z(\boldsymbol{x})|\boldsymbol{z}]$. The MSPE for this predictor is $\mathbb{E}\left[Var\left[Z(\boldsymbol{x})\middle|\boldsymbol{z}\right]\right]$.*

Using the result above for the kriging model, the MSPE-optimal predictor of the response at any $\boldsymbol{x} \in \mathcal{A}$ is

$$\hat{Z}(\boldsymbol{x}) = \mathbb{E}[Z(\boldsymbol{x})|\boldsymbol{z}] = \mu + \boldsymbol{r}^\top(\boldsymbol{x})\boldsymbol{V}^{-1}(\boldsymbol{z} - \mu\boldsymbol{1}), \tag{3.13}$$

and its MSPE is

$$\text{MSPE}(\boldsymbol{x}) = \text{Var}[Z(\boldsymbol{x})|\boldsymbol{z}] = \sigma^2\left(1 - \boldsymbol{r}^\top(\boldsymbol{x})\boldsymbol{V}^{-1}\boldsymbol{r}(\boldsymbol{x})\right), \tag{3.14}$$

where (3.13) and (3.14) can be shown to be the conditional mean and variance, respectively, of $Z(\boldsymbol{x})|\boldsymbol{z}$ (Diggle and Ribeiro Jr., 2007). In the above equations, $\boldsymbol{V}$ is defined as in § 3.3, and $\boldsymbol{r}(\boldsymbol{x}) = [\rho(\boldsymbol{x} - \boldsymbol{x}_i)]_{i=1,\ldots,n}$ is the vector of correlations between $Z(\boldsymbol{x})$ and each of the response values in the design. Note that the outer expectation has been dropped from the MSPE expression in (3.14), since the expression is a constant with respect to $\boldsymbol{z}$ (Diggle and Ribeiro Jr., 2007).

Thus, since the MSPE is a decreasing function of $\boldsymbol{r}^\top(\boldsymbol{x})\boldsymbol{V}^{-1}\boldsymbol{r}(\boldsymbol{x})$, and the correlation function is a decreasing function of the distance between spatial locations, the uncertainty of predicting the response at any location $\boldsymbol{x} \in \mathcal{A}$ is smaller when $\boldsymbol{x}$ is close to the points in the design. This property is explored further in Chapter 4, where optimal designs are discussed.

Finally, we note that the minimum-MSPE predictor given in Result 3.1 is unbiased. That is,

$$\mathbb{E}_{\boldsymbol{z}}\left[\hat{Z}(\boldsymbol{x})\right] = \mathbb{E}_{\boldsymbol{z}}\left[\mathbb{E}_{Z(\boldsymbol{x})|\boldsymbol{z}}\left[Z(\boldsymbol{x})\middle|\boldsymbol{z}\right]\right] = \mathbb{E}_{Z(\boldsymbol{x})}\left[Z(\boldsymbol{x})\right], \tag{3.15}$$

from the law of iterated expectations (Wasserman, 2013). Thus, the predictor in (3.13) is the best linear unbiased predictor (BLUP) for $Z(\boldsymbol{x})$.

Chapter 4

Optimal Design Criteria

Finding an optimal experimental design amounts to finding a design that optimizes some criterion. There are several optimization criteria that can be considered. When predictions are based on the distances between a point and the points in the design, as is the case with kriging using a distance-based correlation function, space-filling designs are often implemented to ensure that the design points are well-spread throughout the region of interest (e.g. Johnson et al., 1990).

Since the goal for this project is optimal prediction, we propose a design criterion that is closely related to the Integrated Mean Squared Prediction Error of Sacks et al. (1989a,b). The criterion is discussed in detail in § 4.1. We also consider the optimal design obtained from a commonly-used space-filling criterion, and we compare the results. The numerical details for obtaining the optimal designs are given in Chapter 5.

A key new innovation for the proposed methodology is that the computer model is used to aid in the design of the physical experiment. Also, the criterion used for obtaining the optimal design takes into account the uncertainty in the computer model parameters, which represent physical factors governing the system.

4.1 Average Integrated Mean Squared Prediction Error

The spatial distribution of glacier melt that is given by the model in § 2.2 contains three tuning parameters: the temperature melt factor (MF), the radiation coefficient for snow (a_{snow}) and the radiation coefficient for ice (a_{ice}). The values of these parameters vary across different glacier environments, as well as across individual glaciers and under different weather conditions. In most studies involving the Hock (1999) melt model, the authors estimate the values of these parameters using field observations (e.g. Huss et al., 2008; Wheler et al., 2014; Gabbi et al., 2014). This amounts to being a type of inverse problem.

We are interested in obtaining a design (i.e. a stake network in the glaciology application) that can provide good predictions of the spatial distribution of the response for a variety of

values of the computer model parameters. In particular, we use a design criterion that is an extension of the Integrated Mean Squared Prediction Error (IMSPE) defined in Sacks et al. (1989a,b). The criterion, which we refer to as Average Integrated Mean Squared Prediction Error (ave-IMSPE), is the expected value of the IMSPE over a proposed joint distribution of the unknown computer model parameters $\boldsymbol{\phi}$.

4.1.1 General Definition of ave-IMSPE

As described in § 3.1, we continue under the assumption that the region of interest is a d-dimensional hyperrectangle, scaled to $\mathcal{A} = [0,1]^d$. Following previous notation, $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$ is the n-by-d design matrix containing n locations in $\mathcal{A}$.

A common measure of prediction error at any new location is the Mean Squared Prediction Error (MSPE). The MSPE at a location $\boldsymbol{x} \in \mathcal{A}$, for a design $\boldsymbol{X}$, is

$$\text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi}) \triangleq \mathbb{E}\left[\left(Z(\boldsymbol{x}) - \hat{Z}(\boldsymbol{x})\right)^2\right], \tag{4.1}$$

where $\boldsymbol{\phi}$ is a given set of input parameters to the computer model (Diggle and Ribeiro Jr., 2007). In the glaciology application, the computer model parameters are $\boldsymbol{\phi} = [MF, a_{\text{snow}}, a_{\text{ice}}]^\top$. The expectation in (4.1) is taken over the joint distribution of the response $Z(\boldsymbol{x})$ and the prediction $\hat{Z}(\boldsymbol{x})$, i.e. the joint distribution of $Z(\boldsymbol{x})$ and the data $\boldsymbol{z} = [Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n)]^\top$ observed using this design. When a kriging model is used to obtain $\hat{Z}(\boldsymbol{x})$, the analytical expression of the MSPE is given in (3.14), for any given design $\boldsymbol{X}$ and computer model parameters $\boldsymbol{\phi}$.

For a given set of computer model parameters $\boldsymbol{\phi}$, we would like the optimal design to have a low value of $\text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi})$ at all (or most) values of $\boldsymbol{x} \in \mathcal{A}$. This can be formulated using the IMSPE criterion of Sacks et al. (1989a,b). For a given $\boldsymbol{\phi}$, the IMSPE-optimal design is the one that minimizes

$$\text{IMSPE}(\boldsymbol{X}; \boldsymbol{\phi}) \triangleq \int_{\mathcal{A}} \text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi}) d\boldsymbol{x} = \int_{\mathcal{A}} \mathbb{E}\left[\left(Z(\boldsymbol{x}) - \hat{Z}(\boldsymbol{x})\right)^2\right] d\boldsymbol{x}. \tag{4.2}$$

Thus, an IMSPE-optimal design is one that minimizes the average uncertainty of prediction (Zhu and Evangelou, 2015).

However, the true values of the computer model parameters are not known. If it is assumed that they are random variables with some joint distribution $\pi(\boldsymbol{\phi})$, an optimal design is one that minimizes

$$\begin{aligned} \text{ave-IMSPE}(\boldsymbol{X}) &\triangleq \int_{\boldsymbol{\phi}} \text{IMSPE}(\boldsymbol{X}; \boldsymbol{\phi}) \pi(\boldsymbol{\phi}) d\boldsymbol{\phi} \\ &= \int_{\boldsymbol{\phi}} \int_{\mathcal{A}} \text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi}) \pi(\boldsymbol{\phi}) d\boldsymbol{x} d\boldsymbol{\phi}. \end{aligned} \tag{4.3}$$

Thus, the ave-IMSPE optimal design minimizes the expected value of the IMSPE over the assumed joint distribution of the computer model parameters. The design is thus optimal in terms of prediction over the region of interest, and robust to changes in the computer model parameters. For a given number of stakes n, the ave-IMSPE optimal design is denoted by $\boldsymbol{X}^*_{(n)}$.

The proposed distribution $\pi(\boldsymbol{\phi})$ on the computer model parameters can be estimated using expert knowledge or data from previous studies. Section 7.1 gives details on how $\pi(\boldsymbol{\phi})$ is obtained for the glacier application.

4.1.2 ave-IMSPE for the Kriging Model

Within the ave-IMSPE expression in (4.3), the MSPE is dependent on the model used to construct the predictions $\hat{Z}(\boldsymbol{x})$. Recall that for the kriging predictor in (3.13), the MSPE of predicting $Z(\boldsymbol{x})$ at any $\boldsymbol{x} \in \mathcal{A}$ is

$$\text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi}) = \sigma^2 \left(1 - \boldsymbol{r}^\top(\boldsymbol{x}) \boldsymbol{V}^{-1} \boldsymbol{r}(\boldsymbol{x})\right), \tag{4.4}$$

where σ^2, $\boldsymbol{r}(\boldsymbol{x})$ and $\boldsymbol{V}$ are defined as in Chapter 3.

Thus, for a kriging model, the MSPE at any point $\boldsymbol{x} \in \mathcal{A}$ depends on the design through $\boldsymbol{r}(\boldsymbol{x})$ and $\boldsymbol{V}$; in particular, it decreases with $\boldsymbol{r}^\top(\boldsymbol{x}) \boldsymbol{V}^{-1} \boldsymbol{r}(\boldsymbol{x})$. Since the correlation function chosen for the kriging model is a decreasing function of the distance between spatial locations, locations $\boldsymbol{x}$ which are farther from the design points will have a higher level of uncertainty in prediction due to a higher value of MSPE. In the design optimization problem, this results in a type of space-filling property, since designs that have nearby points to all locations in $\mathcal{A}$ are generally more favourable.

The MSPE expression in (4.4) also depends on the computer model parameters $\boldsymbol{\phi}$, through the estimation of the Gaussian process variance σ^2 and the correlation parameters $\boldsymbol{\theta}$. Thus, the ave-IMSPE design criterion for kriging is

$$\text{ave-MSPE}(\boldsymbol{X}) = \int_{\boldsymbol{\phi}} \int_{\mathcal{A}} \sigma^2 \left(1 - \boldsymbol{r}^\top(\boldsymbol{x}) \boldsymbol{V}^{-1} \boldsymbol{r}(\boldsymbol{x})\right) \pi(\boldsymbol{\phi}) d\boldsymbol{x} d\boldsymbol{\phi}. \tag{4.5}$$

4.1.3 Approximating ave-IMSPE

In this project, the integrals in the expression for ave-IMSPE are approximated numerically. The inner integral, which evaluates the IMSPE, is approximated by taking the average MSPE over a regular grid of $\boldsymbol{x}$-locations in the region of interest $\mathcal{A}$. Formally, for any given

set of computer model parameters $\boldsymbol{\phi}$,

$$\begin{aligned}\text{IMSPE}(\boldsymbol{X};\boldsymbol{\phi}) &= \int_{\mathcal{A}} \text{MSPE}(\boldsymbol{x};\boldsymbol{X},\boldsymbol{\phi})d\boldsymbol{x} \\ &\approx \frac{|\mathcal{A}|}{N_{\text{grid}}} \sum_{\boldsymbol{x}\in\boldsymbol{X}_{\text{grid}}} \text{MSPE}(\boldsymbol{x};\boldsymbol{X},\boldsymbol{\phi}),\end{aligned} \tag{4.6}$$

where $|\mathcal{A}| = (1-0)^d = 1$ is the measure of the hyperrectangle $\mathcal{A}$, $\boldsymbol{X}_{\text{grid}}$ is a regular d-dimensional grid of points in $\mathcal{A}$, and N_{grid} is the number of grid points. Recall that at any location $\boldsymbol{x}$, the MSPE of predicting the response $Z(\boldsymbol{x})$ is dependent on the model used, and the analytical expression for kriging is given in (4.4).

The outer integral in the expression for ave-IMSPE gives the expected value of the IMSPE over the assumed joint distribution of the computer model parameters. It is approximated using Monte Carlo. Values of $\boldsymbol{\phi}$ are sampled from the proposed joint distribution $\pi(\boldsymbol{\phi})$, and the ave-IMSPE is approximated as follows, for a given design:

$$\begin{aligned}\text{ave-IMSPE}(\boldsymbol{X}) &= \int_{\boldsymbol{\phi}} \text{IMSPE}(\boldsymbol{X};\boldsymbol{\phi})\pi(\boldsymbol{\phi})d\boldsymbol{\phi} \\ &= \mathbb{E}_{\boldsymbol{\phi}}\left[\text{IMSPE}(\boldsymbol{X};\boldsymbol{\phi})\right] \\ &\approx \frac{1}{N_{\text{params}}} \sum_{j=1}^{N_{\text{params}}} \text{IMSPE}(\boldsymbol{X};\boldsymbol{\phi}_j) \\ &\approx \frac{|\mathcal{A}|}{N_{\text{params}} N_{\text{grid}}} \sum_{j=1}^{N_{\text{params}}} \sum_{\boldsymbol{x}\in\boldsymbol{X}_{\text{grid}}} \text{MSPE}(\boldsymbol{x};\boldsymbol{X},\boldsymbol{\phi}_j),\end{aligned} \tag{4.7}$$

where the fourth line follows from (4.6), and $\{\boldsymbol{\phi}_1,\ldots,\boldsymbol{\phi}_{N_{\text{params}}}\}$ are the samples from the proposed distribution $\pi(\boldsymbol{\phi})$. The proposed distribution for the glacier application is estimated empirically using data from the literature. Details on the procedure are given in § 7.1.

4.2 Space-Filling Design Criteria

Space-filling designs are commonly used for spatial models such as kriging. Due to the spatial dependence between response values, it is beneficial to spread out the design points throughout the region of interest, ensuring a low uncertainty of prediction at most locations. Several design criteria exist for this purpose, each defining the space-filling property in different ways.

One such criterion, which is considered herein, ensures that the design points are as spread out as possible from each other. Specifically, the *maximin distance* design (Johnson

et al., 1990) is defined as the design $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$ that maximizes

$$mindist(\boldsymbol{X}) \triangleq \min_{i \neq i'} dist(\boldsymbol{x}_i, \boldsymbol{x}_{i'}), \tag{4.8}$$

where $dist(\boldsymbol{x}_i, \boldsymbol{x}_{i'})$ is a distance function. The Euclidean norm is used here, given by

$$dist(\boldsymbol{x}_i, \boldsymbol{x}_{i'}) = ||\boldsymbol{x}_i - \boldsymbol{x}_{i'}|| = \sqrt{\sum_{j=1}^{d} \left(x_{ij} - x_{i'j}\right)^2}, \tag{4.9}$$

where x_{ij} is the j^{th} dimension of the i^{th} design point.

Thus, a maximin distance design $\boldsymbol{X}^*$ is one that maximizes the minimum distance between any two points in the design. As a result, it ensures that the points in the design are as spread apart as possible, since no two points in the design are closer together than $mindist(\boldsymbol{X}^*)$. This property is useful for a modelling approach such as kriging. Since the prediction at a point $\boldsymbol{x}$ is based on the distances between $\boldsymbol{x}$ and the points in the design, design points that are very close together provide redundant information (Zhu and Evangelou, 2015). In this project, the prediction performance of the maximin design is compared to that of the minimum ave-IMSPE design.

Chapter 5

Finding the Optimal Design

In this chapter, we give details of the methodology for finding minimum ave-IMSPE designs. The first section describes how ave-IMSPE is evaluated for any given design. This requires a preliminary modelling step to be conducted before the design search takes place. In the second section, Particle Swarm Optimization (PSO) is described, which is the design search algorithm used to actually obtain the optimal designs. A final summary of the overall approach is given in the last section.

5.1 Evaluating the Criterion for a Fixed Design

Ideally, for each candidate design considered during the design search, the ave-IMSPE would be evaluated as follows: First, N_{params} values of the computer model parameters $\boldsymbol{\phi}$ would be sampled from their proposed distribution, and these would be used to obtain the corresponding responses from the computer model. For each set of responses, the data would be used to fit a kriging model, and these estimates of the kriging model parameters would be used to evaluate the values of MSPE at a grid of locations. The resulting values of MSPE would then be used to evaluate the ave-IMSPE, as in (4.7).

However, this approach is computationally expensive, since it requires obtaining computer model response values and fitting a new kriging model at each evaluation of ave-IMSPE. To speed up the computation, a preliminary modelling step is introduced, during which N_{params} kriging models are fit before the design search takes place. The same kriging models are then used throughout the entire design search. The step is described in detail below.

Preliminary Modelling Step:

1. N_{params} values of the computer model parameters $\boldsymbol{\phi}$ are sampled from their proposed distribution $\pi(\boldsymbol{\phi})$. For the glaciology application, the proposed distribution is based on values in the literature (see § 7.1).

2. For each of the sampled $\boldsymbol{\phi}$-values, the computer model is then evaluated at a dense grid of points $\boldsymbol{X}_{\text{dense}}$. The resulting responses are $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(N_{\text{params}})}\}$, where each $\boldsymbol{z}^{(j)}$ is the vector of responses observed on the grid for the j^{th} value of the computer model parameters.

3. Each dataset $\left(\boldsymbol{z}^{(j)}, \boldsymbol{X}_{\text{dense}}\right)$ is used to fit a kriging model, as in § 3.3, giving the j^{th} set of kriging parameter estimates, $\{\hat{\mu}^{(j)}, \hat{\sigma}^{2(j)}, \hat{\boldsymbol{\theta}}^{(j)}\}$.

Because the preliminary N_{params} kriging fits have been obtained on a dense grid, they should be close to the appropriate values for each response surface. As a result, they can be treated as surrogates to the computer model response surface throughout the remainder of the algorithm. Then, for every candidate design considered during the design search, the ave-IMSPE is evaluated as follows.

Evaluation of ave-IMSPE:

1. For each of the N_{params} kriging fits that were constructed in the preliminary modelling step, the MSPE is evaluated at a grid of locations $\boldsymbol{x} \in \boldsymbol{X}_{\text{grid}}$. Since the kriging fits are treated as surrogates to the computer model response surface, the kriging parameters $\{\mu^{(j)}, \sigma^{2(j)}, \boldsymbol{\theta}^{(j)}\}$ are assumed to be known and equal to the estimates obtained in the preliminary modelling step (for $j \in \{1, \ldots, N_{\text{params}}\}$). Thus, the MSPE for the design $\boldsymbol{X}$ can be evaluated at any $\boldsymbol{x} \in \boldsymbol{X}_{\text{grid}}$ using the analytical expression (4.4).

2. The values of MSPE, which have been obtained for each fit $j \in \{1, \ldots, N_{\text{params}}\}$ and at each location $\boldsymbol{x} \in \boldsymbol{X}_{\text{grid}}$, are substituted into (4.7) to give the approximation to the ave-IMSPE for this design.

Following the above procedure results in a less computationally expensive design search step. Since N_{params} kriging fits are all obtained on a dense grid in the preliminary modelling step, they are treated as surrogates for the computer model, where each fit corresponds to a different value of the computer model parameters. The value of ave-IMSPE obtained for each candidate design during the design search is then an approximation to the expected uncertainty of prediction, averaged over the computer model parameter values and the region of interest. Although the preliminary modelling step may be computationally expensive, it only needs to be done once, and no further kriging fits or computer model evaluations are required.

In order for an optimal design to actually be found, a design search algorithm must be chosen and implemented. The design search algorithm that is used for this application is PSO, described in the following section.

5.2 Particle Swarm Optimization

The design search algorithm that is used to minimize ave-IMSPE in terms of the design $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$, is Particle Swarm Optimization (PSO), a metaheuristic algorithm originally proposed by Kennedy and Eberhart (1995). The algorithm gets its name from its similarity to the movement of swarms of organisms, such as birds in a flock or fish in a school.

An advantage of PSO is that it requires no assumptions about the function to be minimized (such as monotonicity, convexity or unimodality). It also requires few tuning parameters to be defined to make the algorithm work reasonably well. Several such parameters have been proposed in the literature (Mandal et al., 2015). PSO has been used to obtain maximin designs (e.g. Leatherman et al., 2014), IMSPE-optimal designs (e.g. Leatherman et al., 2014), minimax designs (e.g. Chen et al., 2015), and optimal Latin Hypercube Designs (e.g. Chen et al., 2013).

Throughout the PSO algorithm, every candidate design $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$, containing n points in d dimensions, is reshaped into a vector of length nd and referred to as a *particle*. More specifically, a particle is a vectorized version of the design $\boldsymbol{X}$, constructed by defining the first n elements to be the first column of $\boldsymbol{X}$, the next n elements to be the second column of $\boldsymbol{X}$, etc. The design search algorithm then becomes an optimization problem for finding the particle that, when reshaped back into matrix form, gives the design that minimizes the ave-IMSPE. As before, it is assumed for simplicity that the region of interest for design points is $\mathcal{A} = [0, 1]^d$ (in the glaciology application, this is the scaled rectangle on the glacier, defined in § 2.3). Thus, particles (i.e. vectorized designs) can take on values in $\mathcal{A}_{\text{part}} = [0, 1]^{nd}$.

The main idea behind the PSO algorithm is that multiple particles, referred to as a *swarm*, move simultaneously through the search space $\mathcal{A}_{\text{part}} = [0, 1]^{nd}$, searching for an optimal position, where a position is a point in the nd-dimensional space that corresponds to a design. At each iteration of the algorithm, the position of each particle changes by an additive *velocity* term. The velocity is governed partly by a movement in the direction of the best position of the individual particle thus far, and partly by a movement in the direction of the best global position of any particle in the swarm thus far.

The pseudocode for PSO is given in Algorithm 1. It has been generalized for an arbitrary objective function $f(\cdot)$. For this problem, the objective function is ave-IMSPE. The values of two tuning parameters must be specified to implement the algorithm: the number of particles P that will simultaneously move through the search space, and the maximum number of iterations T that will take place. The algorithm begins by obtaining the initial positions of the P particles. In this project, the initial positions are taken to be randomly generated points in $\mathcal{A}_{\text{part}} = [0, 1]^{nd}$. The initial particle positions (i.e. candidate designs) are denoted by $\boldsymbol{p}_1^{(0)}, \ldots, \boldsymbol{p}_P^{(0)}$. The particle velocities are then initialized and denoted by $\boldsymbol{v}_1^{(0)}, \ldots, \boldsymbol{v}_P^{(0)}$. These initial velocities are set to zero, as in Leatherman et al. (2014).

At any iteration t of the PSO algorithm, the *personal best* position for the i^{th} particle, $\boldsymbol{b}_i^{(t)}$ (for $i \in \{1, \ldots, P\}$), is defined to be the best position that has been encountered by the i^{th} particle thus far, in terms of ave-IMSPE. The *global best*, $\boldsymbol{b}_g^{(t)}$, is the best position that has been obtained by any particle in the swarm thus far. Thus, the next step of the algorithm is to evaluate ave-IMSPE for each of the initial particle positions, and then obtain $\boldsymbol{b}_i^{(0)}$ (for $i \in \{1, \ldots, P\}$) and $\boldsymbol{b}_g^{(0)}$ as described in Algorithm 1.

The algorithm then begins its first iteration. For $t \in \{1, \ldots, T\}$, the velocity term for each particle is calculated as follows:

$$\boldsymbol{v}_i^{(t)} = w\boldsymbol{v}_i^{(t-1)} + \gamma_1 \boldsymbol{\alpha}_i^{(t)} \odot \left(\boldsymbol{b}_i^{(t-1)} - \boldsymbol{p}_i^{(t-1)}\right) + \gamma_2 \boldsymbol{\beta}_i^{(t)} \odot \left(\boldsymbol{b}_g^{(t-1)} - \boldsymbol{p}_i^{(t-1)}\right), \tag{5.1}$$

and the particle moves according to

$$\boldsymbol{p}_i^{(t)} = \boldsymbol{p}_i^{(t-1)} + \boldsymbol{v}_i^{(t)}, \tag{5.2}$$

where the notation $\odot$ in (5.1) denotes the Hadamard product (Mandal et al., 2015). The second and third terms in the velocity expression define the movement of the particle in the direction of its personal best position and the global best position, respectively. The first term is an "inertia" term, which adds a degree of exploration of the search space (Shi and Eberhart, 1998).

There are three parameters in the PSO algorithm. The inertia weight, $w \in [0, 1]$, affects the exploration-exploitation trade-off in the optimization procedure (Bansal et al., 2011). It defines the extent to which the previous velocity affects the movement of the particle at each iteration. The intermediate value of $w = 0.5$ is chosen, as in Leatherman et al. (2014), although non-constant values can also be used, such as a linearly decreasing function of t (Bansal et al., 2011).

The remaining parameters are γ_1 and γ_2, the cognitive and social learning factors, respectively. The default values of $\gamma_1 = \gamma_2 = 2$ are used, thus equally weighing the influence of the personal and global best positions on average (Mandal et al., 2015; Leatherman et al., 2014). The variables $\boldsymbol{\alpha}_i^{(t)}$ and $\boldsymbol{\beta}_i^{(t)}$ are random vectors whose elements are independently distributed as Uniform$[0, 1]$. The random vectors are sampled independently at each iteration and for each particle. They provide a stochastic element in the design search algorithm, allowing for better exploration of the search space.

An additional consideration is the *confinement rule* for the position and velocity vectors. If a particle position $\boldsymbol{p}_i^{(t)}$ falls outside of the particle domain $\mathcal{A}_{\text{part}} = [0, 1]^{nd}$ in any dimension, the value in this dimension is set to the nearest domain point. Thus, the particles that fall outside of the domain are projected onto the edges and corners. To prevent this from happening too frequently, it is required that the velocities do not become too large. Thus, the velocity $\boldsymbol{v}_i^{(t)}$ (for each particle i and at each iteration t) is confined component-wise by $[-0.25 \times \mathbf{1}, 0.25 \times \mathbf{1}]$, where $\mathbf{1}$ is an nd-vector of 1's.

Algorithm 1 Pseudocode for Particle Swarm Optimization.

Randomly initialize particle positions: $\boldsymbol{p}_1^{(0)}, \ldots, \boldsymbol{p}_P^{(0)}$.

for particle i in $1:P$ **do**

Initialize the particle's velocity: $\boldsymbol{v}_i^{(0)}$.

Evaluate the objective function: $f(\boldsymbol{p}_i^{(0)})$.

Set the personal best position to the initial position: $\boldsymbol{b}_i^{(0)} \leftarrow \boldsymbol{p}_i^{(0)}$.

end for

Define $\boldsymbol{b}_{i^*}^{(0)}$ such that $f(\boldsymbol{b}_{i^*}^{(0)}) = \min_i \left(f(\boldsymbol{b}_i^{(0)}) \right)$.

Obtain the global best position: $\boldsymbol{b}_g^{(0)} \leftarrow \boldsymbol{b}_{i^*}^{(0)}$.

for iteration t in $1:T$ **do**

for particle i in $1:P$ **do**

Generate $\boldsymbol{\alpha}_i^{(t)} \overset{iid}{\sim} \text{Unif}[0,1]$.

Generate $\boldsymbol{\beta}_i^{(t)} \overset{iid}{\sim} \text{Unif}[0,1]$.

Update the particle's velocity according to (5.1).

Enforce confinement rule on the velocity.

Update the particle's position according to (5.2).

Enforce confinement rule on the position.

Evaluate the objective function: $f(\boldsymbol{p}_i^{(t)})$.

if $f(\boldsymbol{p}_i^{(t)}) < f(\boldsymbol{b}_i^{(t-1)})$ **then**

Update the personal best position: $\boldsymbol{b}_i^{(t)} \leftarrow \boldsymbol{p}_i^{(t)}$.

else

$\boldsymbol{b}_i^{(t)} \leftarrow \boldsymbol{b}_i^{(t-1)}$

end if

end for

Define $\boldsymbol{b}_{i^*}^{(t)}$ such that $f(\boldsymbol{b}_{i^*}^{(t)}) = \min_i \left(f(\boldsymbol{b}_i^{(t)}) \right)$.

Update the global best position: $\boldsymbol{b}_g^{(t)} \leftarrow \boldsymbol{b}_{i^*}^{(t)}$.

end for

OUTPUT: $b_g^{(T)}$ and $f(\boldsymbol{b}_g^{(T)})$

After a maximum number of iterations T, the optimal particle position is taken to be the globally best position $\boldsymbol{b}_g^{(T)}$. This is then reshaped back into matrix form, to obtain the optimal design $\boldsymbol{X}^* = [\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_n^*]^\top$. We note that, due to the nature of the PSO algorithm, only one particle needs to find the optimum. The optimum, which is the minimum value of ave-IMSPE, is then ave-IMSPE($\boldsymbol{X}^*$).

5.3 Summary of Overall Approach

The following is an overview of all of the steps involved in finding the ave-IMSPE optimal design. Steps 1 and 2 have been described in detail in § 5.1-5.2, and are given here only in brief.

Step 1 (Preliminary Modelling Step): N_{params} values of the computer model parameters $\boldsymbol{\phi}$ are sampled, and the corresponding responses are obtained using the computer model. For each sampled value of $\boldsymbol{\phi}$, a kriging model is fit to act as a surrogate for the computer model response surface at this value of $\boldsymbol{\phi}$.

Step 2 (Design Search): The PSO design search algorithm (described in § 5.2) is conducted. For each candidate design $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$ during the design search, the ave-IMSPE is evaluated as follows:

> **Step 2.1:** The MSPE is obtained from the analytical expression at a grid of locations $\boldsymbol{x} \in \boldsymbol{X}_{\text{grid}}$, for each of the N_{params} kriging fits.
>
> **Step 2.2:** The values of MSPE are substituted into (4.7) to give the approximation to the ave-IMSPE for this design.

Step 3: The result of the design search algorithm is the optimal design $\boldsymbol{X}_{(n)}^*$, which is the n-point design that minimizes ave-IMSPE. Step 2 can then be repeated for various values of n, obtaining the optimal design $\boldsymbol{X}_{(n)}^*$ each time.

Step 4: A plot of ave-IMSPE($\boldsymbol{X}_{(n)}^*$) vs. n can help to decide on a value of n that provides an acceptable level of prediction uncertainty while being affordable to the experimenter.

Chapter 6

Simulation Results

The methodology of Chapters 4-5 is implemented on several simple computer models, in order to illustrate how ave-IMSPE optimal designs are obtained. The results are compared to those of a maximin design. All programming is done in R (R Core Team, 2015; Adler et al., 2016; Venables and Ripley, 2002; Sarkar, 2008; Wickham, 2011; Calaway et al., 2015).

6.1 Additive Two-Dimensional Function

Consider the following two-dimensional example of a computer model giving the response $Z(\boldsymbol{x})$ at any location $\boldsymbol{x} \in \mathcal{A} \subset \mathbb{R}^2$:

$$Z(\boldsymbol{x}) = \exp\left(\frac{ax_1}{2}\right) + 2\sin(2bx_2). \tag{6.1}$$

The parameters for this model are $\boldsymbol{\phi} = [a, b]^\top$, with $a, b \in [1, 4]$. This is a simple additive two-dimensional model, for which different combinations of the computer model parameters result in response surfaces with different degrees of variability in each dimension. The region of interest is assumed to be $\mathcal{A} = [0, 1]^2$.

The first step is defining the joint distribution of the computer model parameters $\boldsymbol{\phi} = [a, b]^\top$. For this example, the distributional assumption on the computer model parameters is that a and b are independent and uniformly distributed on $[1, 4]^2$. Then, $N_{\text{params}} = 20$ values of $\boldsymbol{\phi} = [a, b]^\top$ are sampled from this proposed distribution.

As described in Chapter 5, a preliminary modelling step is completed before beginning the design search. At each of the $N_{\text{params}} = 20$ sampled values of $\boldsymbol{\phi}$, response values are obtained from the computer model in (6.1) at a 25×25 grid of locations, and a kriging model is fit to each dataset, with a small value of $\nu^2 = 10^{-3}$ added to the diagonal of the correlation matrix for purely computational reasons, as described in § 3.1 and § 3.3.

Several of the resulting kriging model surfaces are shown in Figure 6.1. It is clear that for different values of the computer model parameters, the variability of the response surface

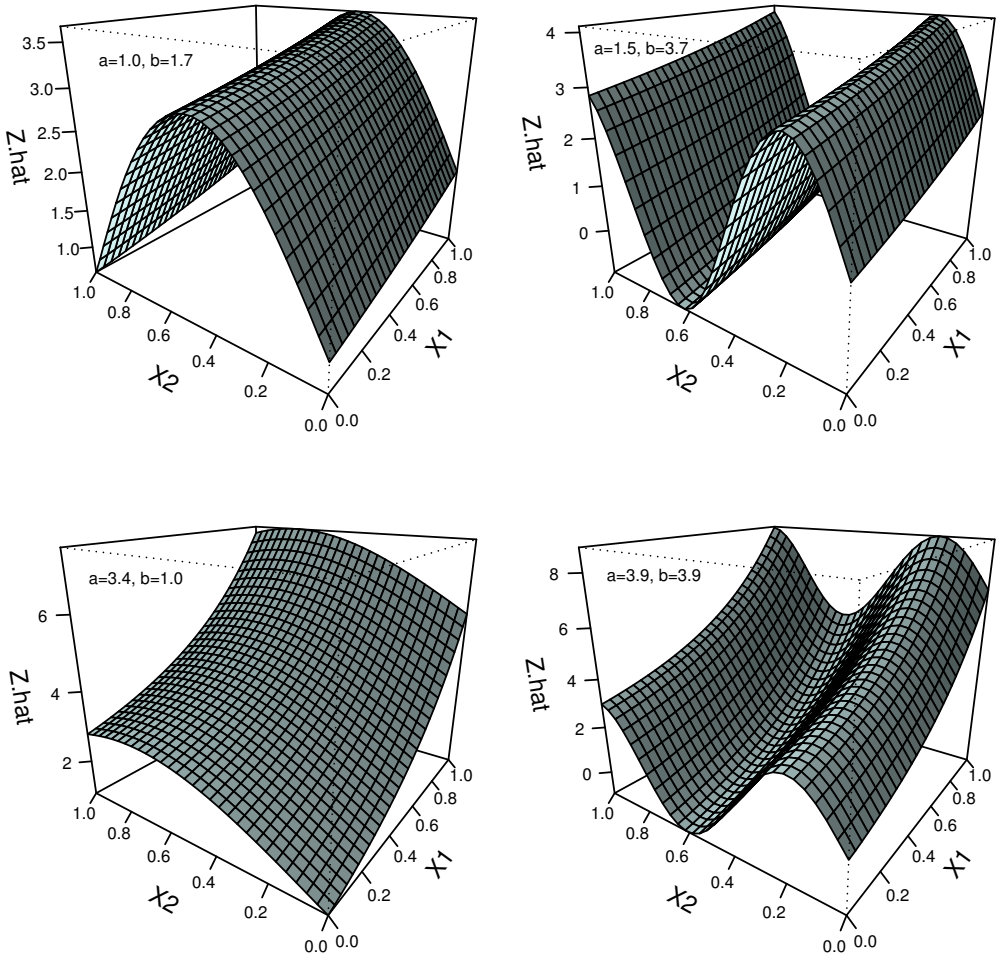

Figure 6.1: Kriging model fits corresponding to the response surfaces for the additive two-dimensional function, with $\boldsymbol{\phi}^\top = [1.0, 1.7]$, $[1.5, 3.7]$, $[3.4, 1.0]$ and $[3.9, 3.9]$, respectively.

in each dimension is quite different. Note that if the IMSPE-optimal designs were to be found for each of the 20 surfaces separately, it is most likely that surfaces with a higher variability in x_2 than in x_1 would require a larger number of points in the x_2-dimension, and vice versa.

Recall from § 4.1.3 that the ave-IMSPE criterion for any given design $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$ is approximated as

$$\text{ave-IMSPE}(\boldsymbol{X}) \approx \frac{|\mathcal{A}|}{N_{\text{params}} N_{\text{grid}}} \sum_{j=1}^{N_{\text{params}}} \sum_{\boldsymbol{x} \in \boldsymbol{X}_{\text{grid}}} \text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi}_j),$$

where $|\mathcal{A}| = 1$, $N_{\text{params}} = 20$, $\boldsymbol{X}_{\text{grid}}$ is a regular 10×10 grid of points in the region of interest, and thus $N_{\text{grid}} = 10^2 = 100$. Recall that the value of $\text{MSPE}(\boldsymbol{x}; \boldsymbol{X}, \boldsymbol{\phi}_j)$ for the j^{th}

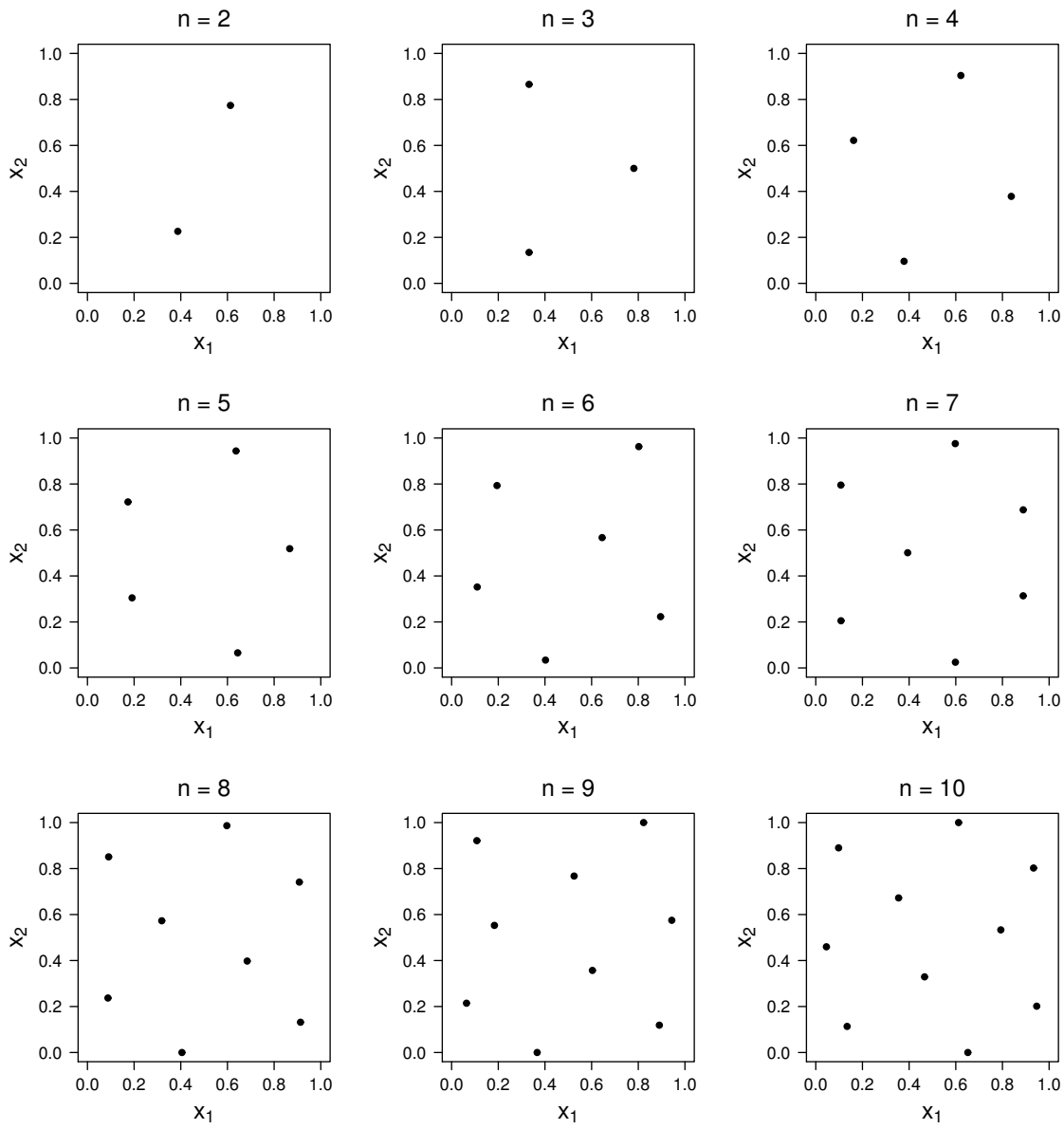

Figure 6.2: ave-IMSPE optimal designs for the additive two-dimensional function.

set of computer model parameters is found using the analytical expression

$$\sigma^2 \left(1 - \boldsymbol{r}^\top(\boldsymbol{x}) \boldsymbol{V}^{-1} \boldsymbol{r}(\boldsymbol{x})\right),$$

where the kriging parameters σ^2 and $\boldsymbol{\theta}$ are taken from the j^{th} kriging model obtained earlier.

The ave-IMSPE optimal designs are found for each of $n \in \{2, 3, \ldots, 10\}$, as described in Chapter 5. The PSO algorithm is run with $P = 80$ candidate particles moving simultaneously at each iteration, and $T = 150$ iterations. The algorithm is randomly restarted ten times, and the best of the ten results is taken to be the overall optimal design for a given value of n.

The optimal designs $\boldsymbol{X}^*_{(n)}$ for $n \in \{2, 3, \ldots, 10\}$ are shown in Figure 6.2. The designs all appear to fill the region of interest very well. This space-filling property makes intuitive sense, since it allows for a low uncertainty of prediction everywhere in the region of interest (see § 4.1.2).

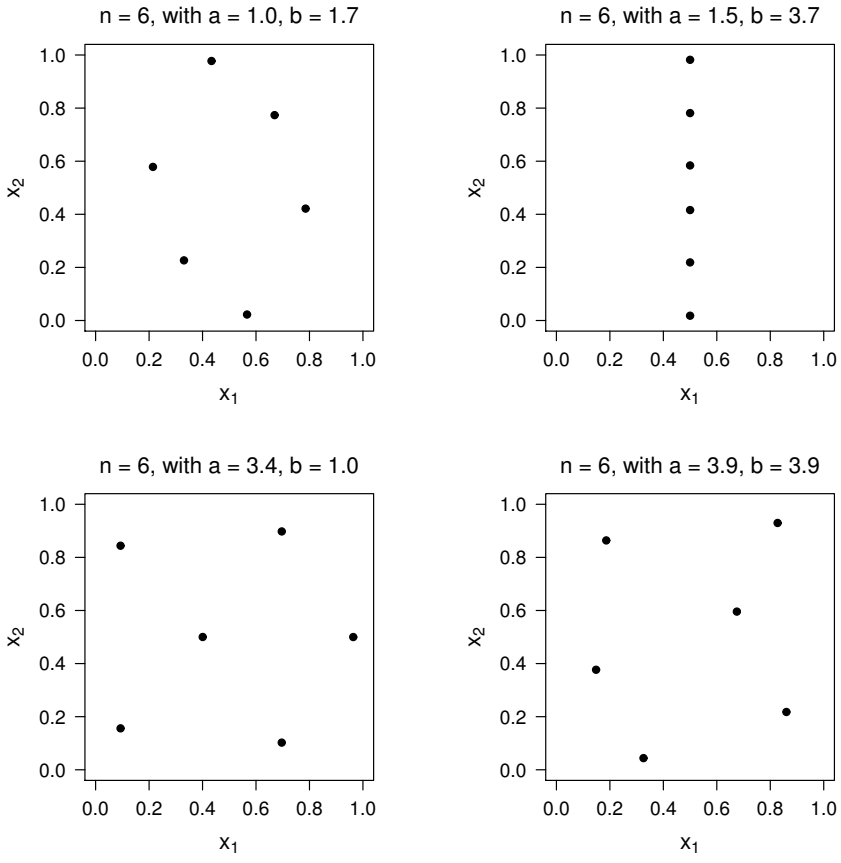

Figure 6.3: IMSPE-optimal designs for four individual fits for the additive two-dimensional function, with $n = 6$.

For comparison, the same PSO algorithm is used to obtain the IMSPE-optimal designs for the four kriging fits in Figure 6.1, for $n = 6$. The results are shown in Figure 6.3. It can be seen that most of the four IMSPE-optimal designs are very different. Perhaps the most characteristic result is the one for $a = 1.5$ and $b = 3.7$. The response surface for this kriging fit is very variable in x_2 and varies only slightly in x_1. As a result, the IMSPE-optimal design is one that is evenly-spaced along the x_2-dimension, with only a single value of x_1 being observed. As expected, the ave-IMSPE optimal designs in Figure 6.2 are, in a sense, "averages" of the designs that would have been obtained using any individual surface. They are thus robust to changes in the computer model parameters.

The PSO algorithm is also used to obtain the two-dimensional maximin designs, defined in § 4.2 (shown in Figure 6.4). For comparing the performance of the two designs, a measure of the proportion of unexplained variability is given by

$$\int \frac{\text{IMSPE}(\boldsymbol{X};\boldsymbol{\phi})}{\sigma^2(\boldsymbol{\phi})}\pi(\boldsymbol{\phi})d\boldsymbol{\phi}, \tag{6.2}$$

where $\sigma^2(\boldsymbol{\phi})$ is the variability of the response when the true value of the computer model parameters is $\boldsymbol{\phi}$. This is approximated using

$$\text{scaled ave-IMSPE}(\boldsymbol{X}) = \frac{\text{ave-IMSPE}(\boldsymbol{X})}{\int \sigma^2(\boldsymbol{\phi})\pi(\boldsymbol{\phi})d\boldsymbol{\phi}}, \tag{6.3}$$

where the denominator is approximated by the sample mean $\frac{1}{N_{\text{params}}}\sum_{j=1}^{N_{\text{params}}}\sigma^{2(j)}$.

Table 6.1 and Figure 6.5 show the resulting scaled ave-IMSPE values for the optimal designs and the maximin designs, for each n. As expected, the ave-IMSPE optimal designs perform better than the maximin designs, and the value of ave-IMSPE decreases with n.

Table 6.1: Scaled values of ave-IMSPE for the additive two-dimensional function.

n	ave-IMSPE Optimal Design	Maximin Design
2	0.435	0.666
3	0.305	0.445
4	0.223	0.500
5	0.174	0.280
6	0.136	0.174
7	0.107	0.178
8	0.088	0.127
9	0.074	0.199
10	0.060	0.094

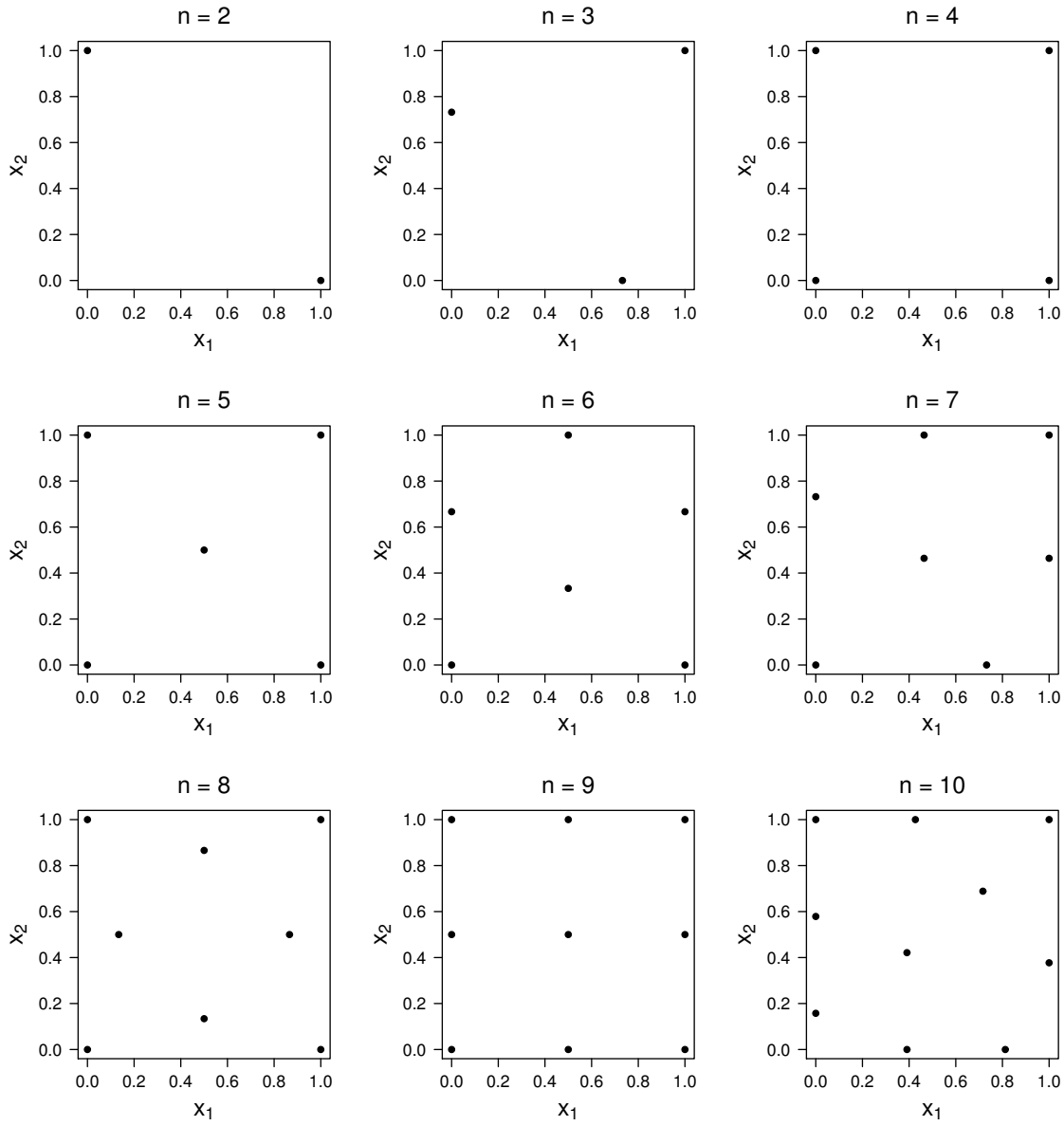

Figure 6.4: Two-dimensional maximin designs.

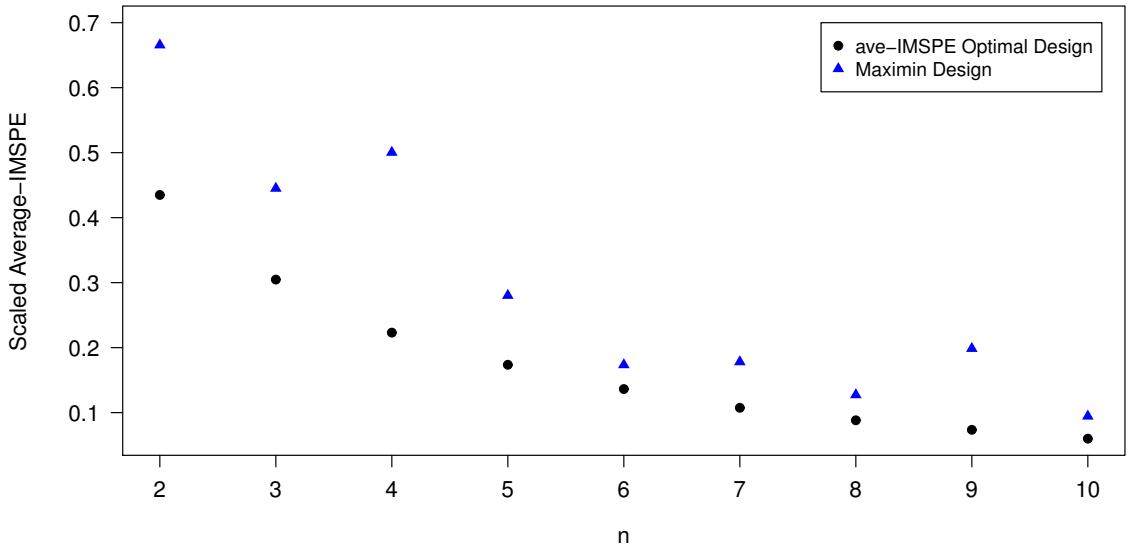

Figure 6.5: Scaled ave-IMSPE values at the optimal designs and maximin designs, for the additive two-dimensional function.

6.2 One-Dimensional Damped Cosine Function

The next example computer model that is considered is

$$Z(x) = \exp(-ax)\cos(b\pi x). \tag{6.4}$$

This is a modified version of the Santner et al. (2003) one-dimensional damped cosine function, obtained from the Virtual Library of Simulation Experiments (Surjanovic and Bingham, 2013). The parameters for this model are $\boldsymbol{\phi} = [a, b]^\top$, with $a \in [0, 3]$ and $b \in [2, 5]$. The region of interest is $\mathcal{A} = [0, 1]$.

We again assume that the computer model parameters are independent and uniformly distributed on their respective domains. As in the previous example, $N_{\text{params}} = 20$ values of $\boldsymbol{\phi} = [a, b]^\top$ are sampled from this proposed distribution, responses are obtained for each sample at 25 equally-spaced locations in $\mathcal{A}$, and a kriging model is fit to each dataset (with a small value of $\nu^2 = 10^{-5}$ added to the diagonal of the correlation matrix). Several of the kriging model surfaces are shown in Figure 6.6.

Following the same methodology as in § 6.1, the ave-IMSPE optimal designs are found for each of $n \in \{2, 3, \ldots, 10\}$. As before, the PSO algorithm is run with $P = 80$ candidate particles, $T = 150$ iterations, and ten random restarts. The optimal designs are shown in Figure 6.7. They all consist of fairly equally-spaced points in $\mathcal{A} = [0, 1]$.

Figure 6.8 shows the one-dimensional maximin designs, obtained using the PSO algorithm. As in § 6.1, the scaled value of ave-IMSPE is used to compare the performance

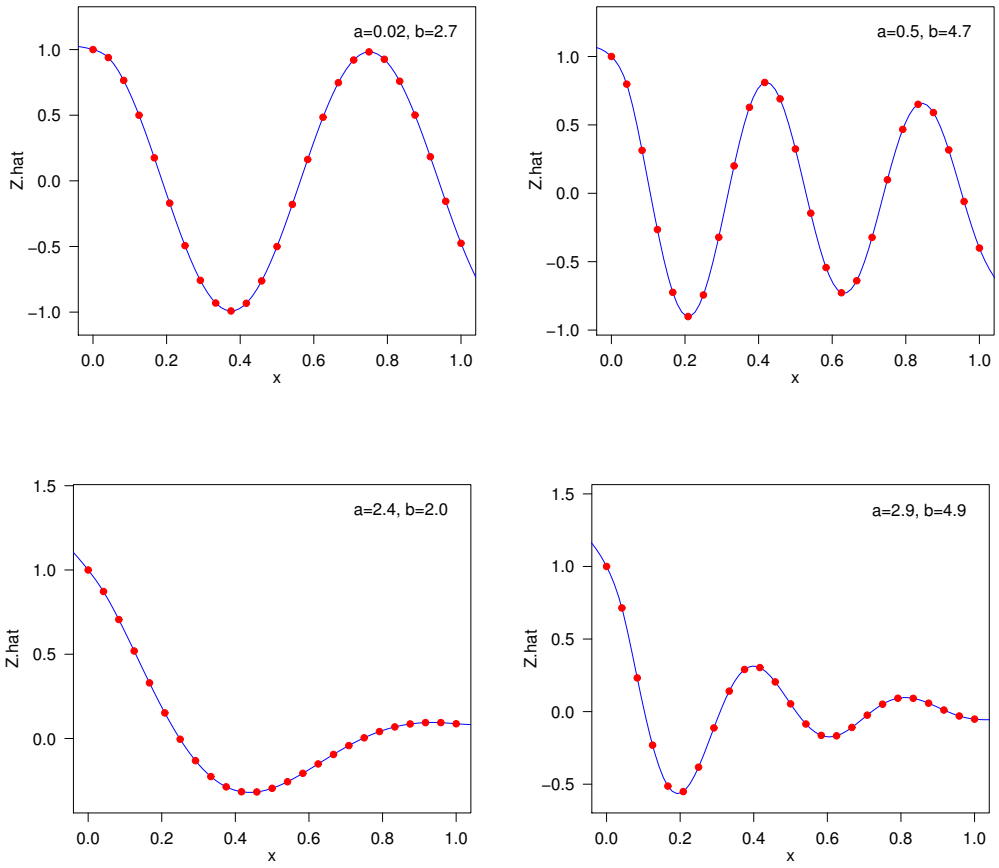

Figure 6.6: Kriging model fits corresponding to the response surfaces for the one-dimensional damped cosine function, with $\boldsymbol{\phi}^\top = [0.02, 2.7]$, $[0.5, 4.7]$, $[2.4, 2.0]$ and $[2.9, 4.9]$, respectively.

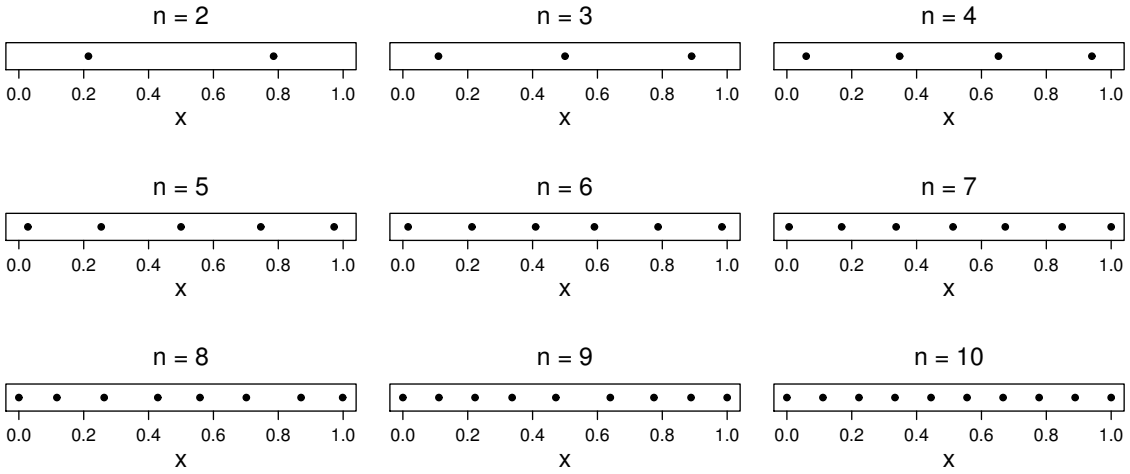

Figure 6.7: ave-IMSPE optimal designs for the one-dimensional damped cosine function.

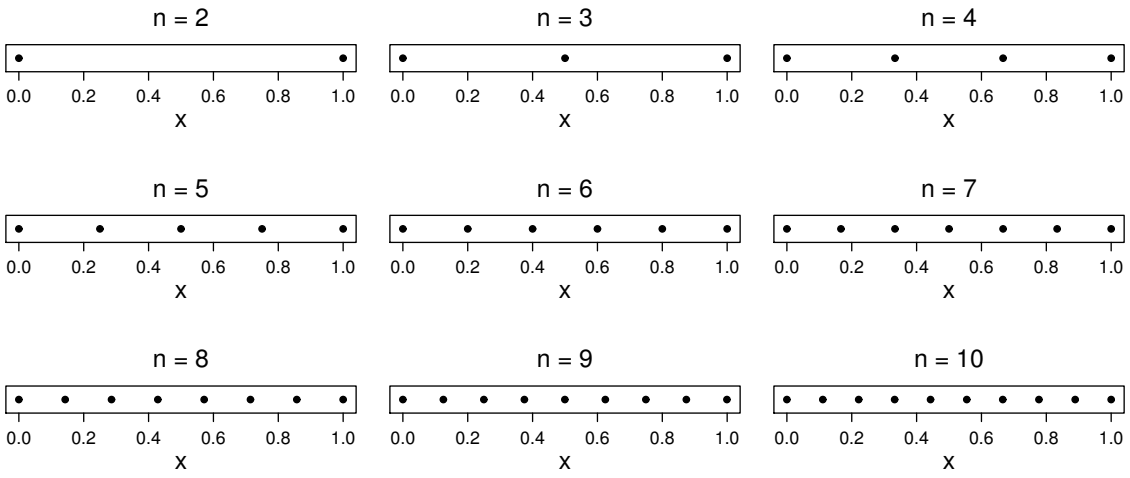

Figure 6.8: One-dimensional maximin designs.

of the two designs. Table 6.2 and Figure 6.9 show the resulting scaled ave-IMSPE values for the optimal designs and the maximin designs, for each n. As before, the ave-IMSPE optimal designs improve in performance as n increases. However, due to their similarity to the maximin designs, the performance is very similar.

Table 6.2: Scaled values of ave-IMSPE for the one-dimensional damped cosine function.

n	ave-IMSPE Optimal Design	Maximin Design
2	1.62×10^{-1}	3.07×10^{-1}
3	6.83×10^{-2}	9.93×10^{-2}
4	3.27×10^{-2}	3.92×10^{-2}
5	1.75×10^{-2}	1.88×10^{-2}
6	9.73×10^{-3}	1.02×10^{-2}
7	5.95×10^{-3}	5.99×10^{-3}
8	3.57×10^{-3}	3.83×10^{-3}
9	1.76×10^{-3}	2.58×10^{-3}
10	9.99×10^{-6}	1.25×10^{-5}

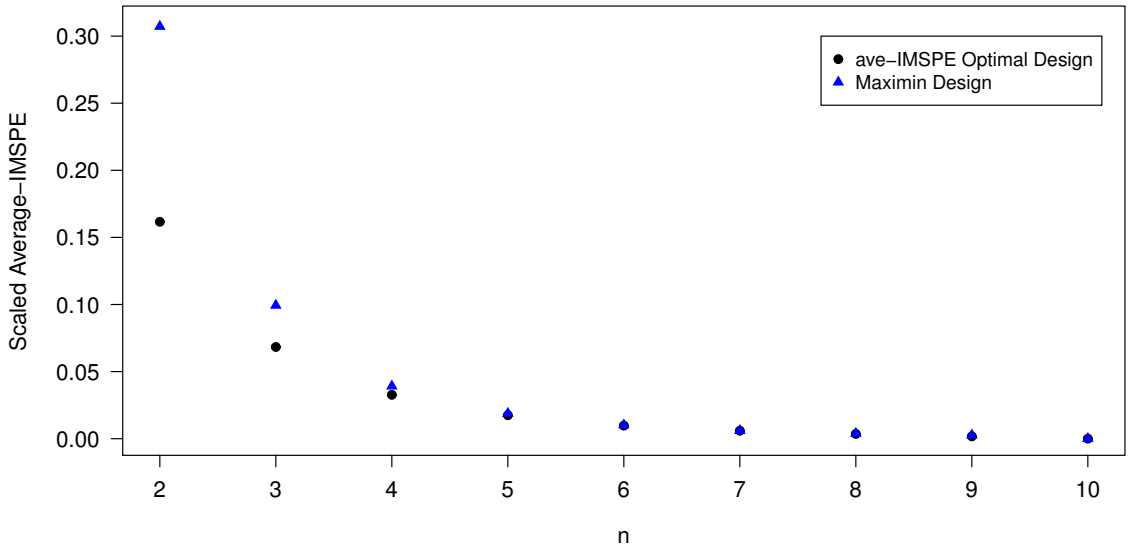

Figure 6.9: Scaled ave-IMSPE values at the optimal designs and maximin designs, for the one-dimensional damped cosine function.

6.3 Ackley Function

The final example that we consider is the 2-dimensional Ackley function,

$$Z(\boldsymbol{x}) = -a \exp\left(-b\sqrt{\frac{1}{d}\sum_{j=1}^{d} x_j^2}\right) - \exp\left(\frac{1}{d}\sum_{j=1}^{d}\cos(cx_j)\right) + a + \exp(1), \tag{6.5}$$

where $d = 2$, and we use the parameter values $a = 20$ and $b = 0.2$ (obtained from the Virtual Library of Simulation Experiments, Surjanovic and Bingham, 2013). This function is commonly used for testing numerical optimization methods, because it contains many local minima. We scale it so that its domain, initially $[-32.768, 32.768]^d$, becomes $\mathcal{A} = [0, 1]^d$.

The unknown parameter in the model is c, whose distribution is defined to be Gaussian with a mean of 2π and a standard deviation of $\pi/2$. As before, $N_{\text{params}} = 20$ values of $\phi = c$ are sampled from this distribution, responses are obtained for each sample at 25 equally-spaced locations in $\mathcal{A}$, and a kriging model is fit to each dataset (with $\nu^2 = 10^{-3}$). Two of the kriging model surfaces are shown in Figure 6.10. The number of local minima differs greatly for the two surfaces. However, both surfaces appear to be nearly symmetrical with respect to x_1 and x_2. Thus, it can be expected that the IMSPE-optimal designs for the two fits will be approximately evenly distributed in the two dimensions, and will thus be similar to each other.

The ave-IMSPE optimal designs are found for each of $n \in \{2, 3, \ldots, 10\}$, using the same PSO algorithm. They are shown in Figure 6.11. They appear to fill the space $\mathcal{A} = [0, 1]^2$

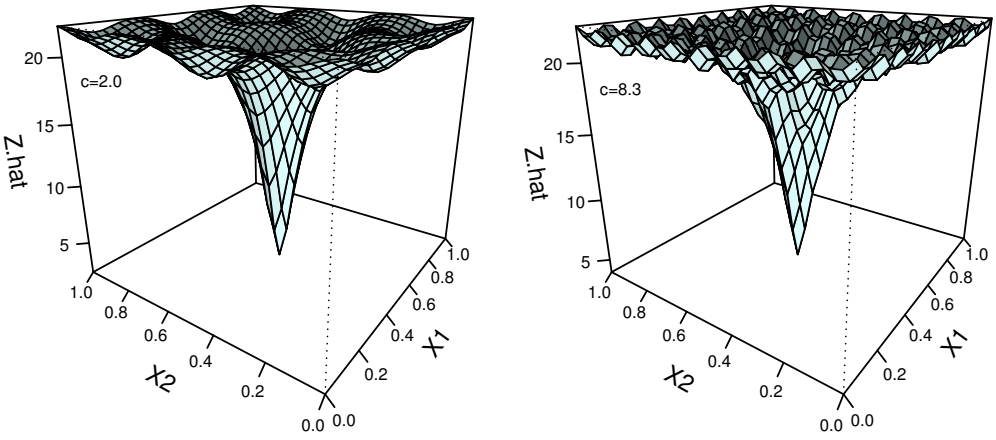

Figure 6.10: Kriging model fits corresponding to the response surfaces for the two-dimensional Ackley function, with $\phi = 2.0$ and 8.3, respectively.

very well, and are all nearly symmetrical with respect to x_1 and x_2, due the response surface having equal variability in each dimension.

The maximin designs are the same as in § 6.1. Table 6.3 and Figure 6.12 show the resulting scaled ave-IMSPE values for the optimal designs and the maximin designs, for each n. As before, the ave-IMSPE optimal designs perform better than the maximin designs, and their performance improves with n.

Table 6.3: Scaled values of ave-IMSPE for the two-dimensional Ackley function.

n	ave-IMSPE Optimal Design	Maximin Design
2	0.666	0.847
3	0.547	0.696
4	0.453	0.670
5	0.382	0.535
6	0.328	0.413
7	0.285	0.368
8	0.249	0.299
9	0.220	0.318
10	0.193	0.242

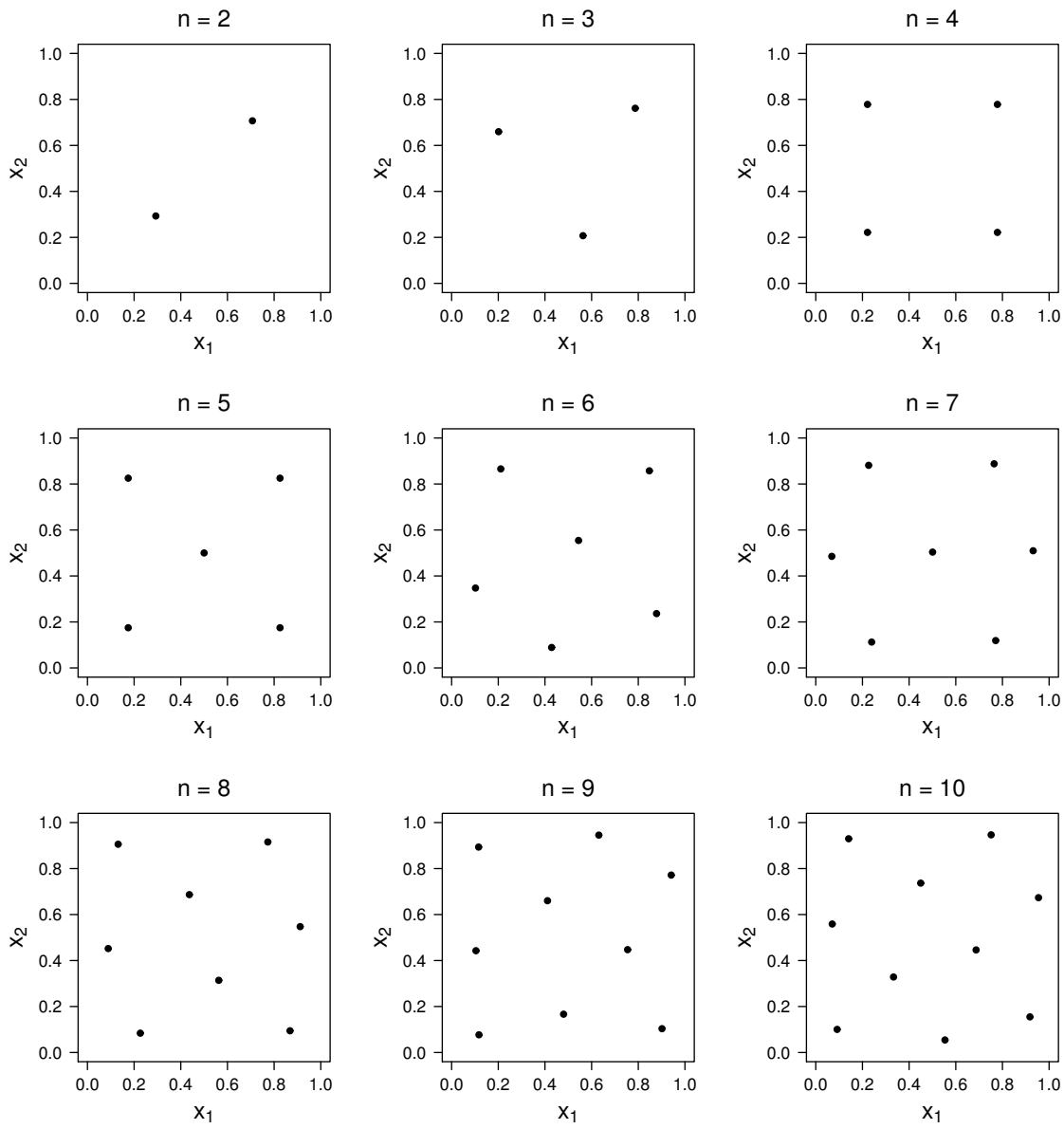

Figure 6.11: ave-IMSPE optimal designs for the two-dimensional Ackley function.

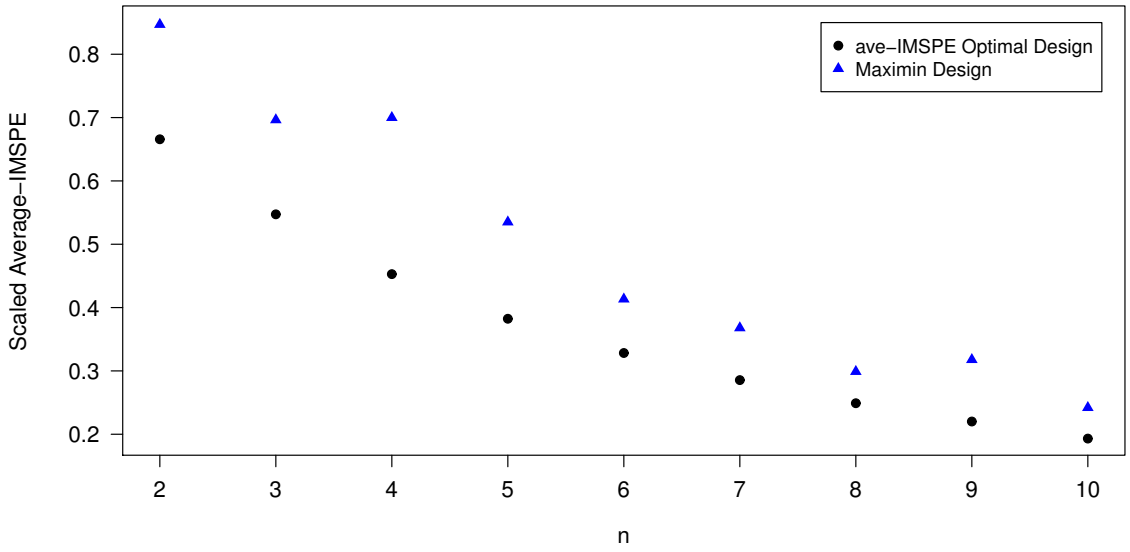

Figure 6.12: Scaled ave-IMSPE values at the optimal designs and maximin designs, for the two-dimensional Ackley function.

Chapter 7

Return to the Glacier

Using the same approach as in the previous chapter, we obtain results for the glaciology application discussed in Chapter 2. The results are compared to those of a maximin design and a design commonly applied in the glaciology literature.

7.1 Proposed Distribution of ϕ

The computer model parameters for the melt model in (2.2) are: MF (temperature melt factor, measured in mm w.e. day^{-1} °C^{-1}) and $a_{\text{snow/ice}}$ (radiation coefficients for snow and ice surfaces, given in 10^{-3} mm w.e. h^{-1} °C^{-1} m^2 W^{-1}). As with the computer model parameters in the simulation examples, these are inputs to the computer model whose values are not known in advance. Since the distribution of $\boldsymbol{\phi} = [MF, a_{\text{snow}}, a_{\text{ice}}]^\top$ is not precisely known, a distribution is specified based on values in the literature.

A literature review for uses of the Hock (1999) melt model was done. From this, twenty separate vectors $\boldsymbol{\phi} = [MF, a_{\text{snow}}, a_{\text{ice}}]^\top$ were found, corresponding to studies on various glaciers around the world. The values are given in Table 7.1. The top of Figure 7.1 shows the scatterplot matrix of the parameter values. It appears that, in the data collected, there is a strong positive linear relationship between a_{snow} and a_{ice}, and a very weak relationship between the other two pairs of parameters. The proposed joint distribution $\pi(\boldsymbol{\phi})$ on the computer model parameters is a multivariate Gaussian distribution with means, variances and correlations equal to the empirical values obtained from the literature, which are given in Table 7.2. The distribution is truncated to require that $MF, a_{\text{snow}}, a_{\text{ice}} > 0$.

From this joint distribution, $N_{\text{params}} = 50$ values of $\boldsymbol{\phi}$ are sampled, and the resulting scatterplot matrix is given in the bottom of Figure 7.1. The results closely resemble the data from the literature.

Table 7.1: Values found in the literature for the Hock (1999) computer model parameters: MF (mm w.e. day^{-1} °C^{-1}) and $a_{\text{snow/ice}}$ (10^{-3} mm w.e. h^{-1} °C^{-1} m^2 W^{-1}).

Source	Glacier	Location	MF	a_{snow}	a_{ice}
Hock (1999)	Storglaciären	Sweden	1.8	0.6	0.8
Schuler et al. (2002)	Unteraargletscher	Switzerland	0.45	0.5	0.7
De Woul et al. (2006)	Hofsjökull	Iceland	2.0	0.9	1.2
Huss et al. (2007)	Gornergletscher	Switzerland	0.744	0.35	0.61
Trüssel et al. (2015)	Yakutat Glacier	Alaska	4.3	0.5	0.792
Schuler et al. (2005a)	Engabreen	Norway	4.2	0.23	0.43
Schuler et al. (2005b)	Gruvefonna	Svalbard	3.32	1.21	1.43
Zhang et al. (2008)	Tuotuo River Basin Glaciers	China	2.1	0.254	0.6
Kronenberg et al. (2016)	Glacier No. 354	Kyrgyzstan	1.41	0.471	0.942
Huss et al. (2008)	Valley of Zinal Glaciers	Switzerland	0.622	0.528	0.622
Klok et al. (2001)	Rhone Basin Glaciers	Switzerland	1.85	0.4	0.7
Shi et al. (2016)	Xiao Dongkemadi Glacier	Tibetan Plateau	4.0	0.167	0.625
Pellicciotti et al. (2005)	Haut Glacier d'Arolla	Switzerland	1.968	0.52	1.06
Gabbi et al. (2014)	Rhonegletscher	Switzerland	2.64	0.3	0.5
Wheler et al. (2014)	"South Glacier"	Yukon	4.5	0.15	0.8
Hock et al. (2007)	Storglaciären	Sweden	1.4	0.5	0.7
Schuler et al. (2007)	Austfonna	Svalbard	1.8	1.125	1.667
MacDougall et al. (2011)	"South Glacier"	Yukon	1.5	0.70	1.3
MacDougall et al. (2011)	"North Glacier"	Yukon	0.8	0.80	0.95
Flowers and Clarke (2000)	Trapridge Glacier	Yukon	1.201	0.400	0.529

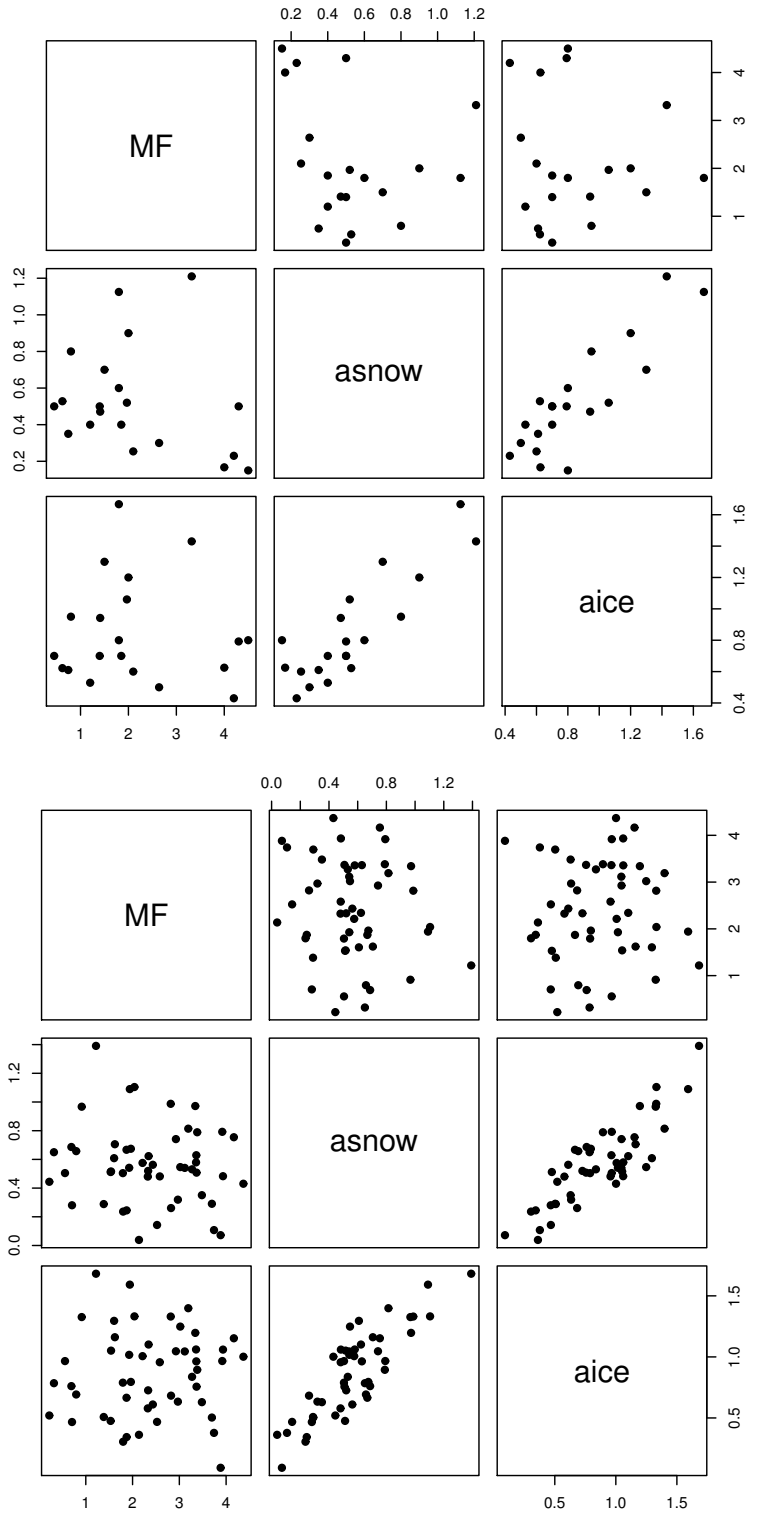

Figure 7.1: Hock (1999) model parameter values found in the literature (top) and obtained from the proposed joint distribution (bottom).

Table 7.2: Means and variances of the Hock (1999) model parameter values found in the literature (left), and the corresponding correlation matrix (right).

	MF	a_{snow}	a_{ice}
Mean	2.130	0.530	0.848
Variance	1.643	0.086	0.110

	MF	a_{snow}	a_{ice}
MF	1	-0.249	-0.056
a_{snow}	-0.249	1	0.864
a_{ice}	-0.056	0.864	1

7.2 Design Optimization Results

As in the simulation examples of Chapter 6, the preliminary modelling step consists of obtaining responses from the computer model for each of the $N_{\text{params}} = 50$ sampled values of $\boldsymbol{\phi}$ (on the 30 m $\times$ 30 m DEM grid), and fitting a kriging model to each dataset, with a small value of $\nu^2 = 10^{-3}$ added to the diagonal of the correlation matrix for computational reasons.

Looking at the sampled values of $\boldsymbol{\phi}$, it is difficult to determine which ones correspond to complex surfaces. However, in each of the $d = 2$ dimensions, the value of the correlation parameter θ_j represents the degree of spatial correlation in that dimension. A larger value of θ_j generally implies a smoother surface in that dimension. Four kriging model surfaces (for the rectangular region of interest) are shown in Figure 7.2, with correlation parameters $\boldsymbol{\theta}^\top = [0.32, 0.14]$, $[0.38, 0.07]$, $[0.15, 0.09]$ and $[0.14, 0.04]$. These kriging fits correspond to the responses obtained from the computer model using $\boldsymbol{\phi}^\top = [3.88, 0.07, 0.09]$, $[2.14, 0.04, 0.36]$, $[1.53, 0.51, 0.48]$ and $[0.56, 0.50, 0.97]$. The dimensions x_1 and x_2 correspond to standardized values of Easting and Northing, respectively. As expected, the surfaces differ with respect to the amount of variability across the surface. However, they share the common result that the melt tends to vary most greatly in x_2. In particular, the melt decreases as x_2 increases. This makes intuitive sense, because locations at a higher value of x_2 are generally at a higher elevation (see Figure 2.2).

The ave-IMSPE optimal designs are found for each of $n \in \{2, 3, \ldots, 10, 12, 14, 16, 18, 20, 25, 30\}$, using the PSO algorithm with $P = 80$ candidate particles moving simultaneously at each iteration, $T = 150$ iterations, and ten random restarts. Due to the complexity of the surfaces, each random restart results in a slightly different configuration of the optimal design. However, the best designs all share the property of being somewhat space-filling in x_2 (since this is the dimension in which the response varies most greatly), and being positioned close to the middle of the x_1-values. This is an interesting result, since it is qualitatively similar to the suggestions discussed in the glaciology literature (see Chapter 2). The ten random restarts for each value of n result in very similar values of ave-IMSPE. The overall best designs are shown in Figures 7.3 and 7.4.

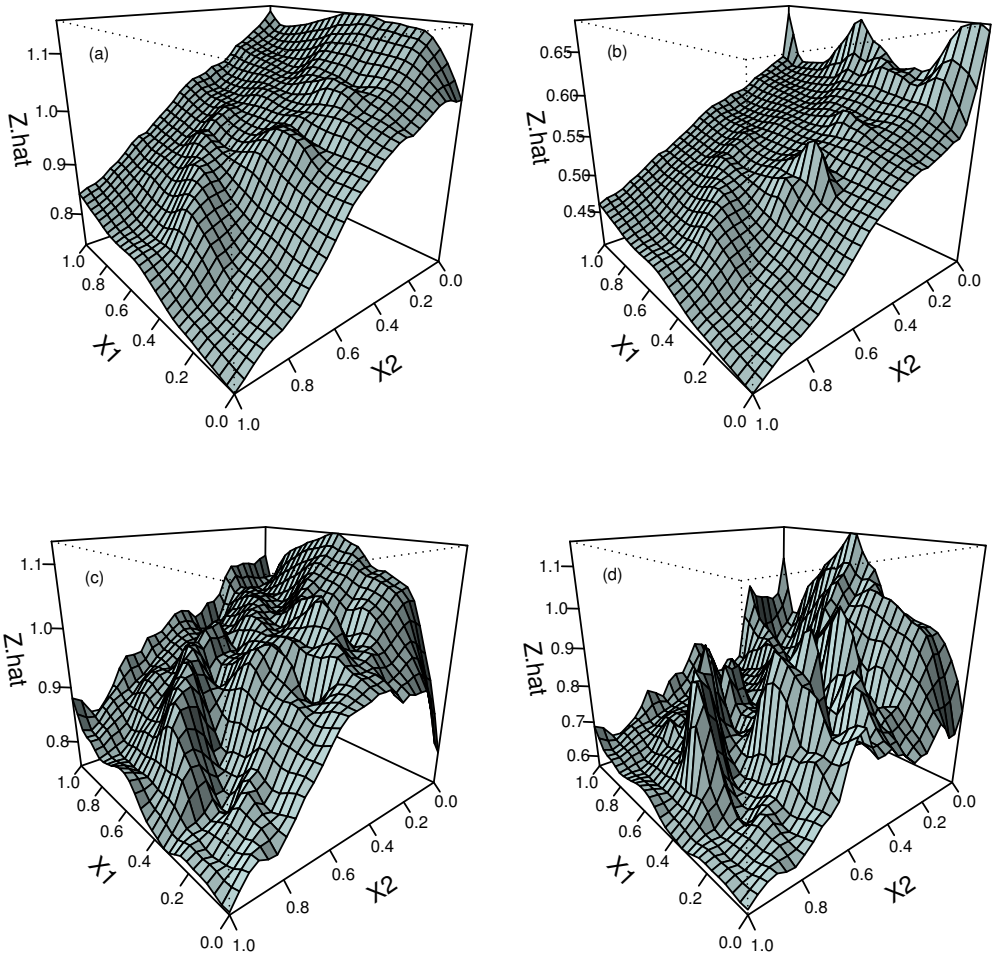

Figure 7.2: Kriging model fits corresponding to the response surfaces for the glacier application, with $\boldsymbol{\phi}^\top = [3.88, 0.07, 0.09]$, $[2.14, 0.04, 0.36]$, $[1.53, 0.51, 0.48]$ and $[0.56, 0.50, 0.97]$, respectively.

These results are compared to the two-dimensional maximin designs (which, for $n \leq 10$, are the same as the ones shown in Figure 6.4), as well as *expert-knowledge designs* constructed using the common approach in the glaciology literature. As described in Chapter 2, the recommended approach is to arrange the stakes along the centre line following the longitudinal axis of the glacier. Due to the way in which the rectangular region is defined, this corresponds approximately to a centre line down the middle of the region, in the direction of x_2. As suggested by Fountain and Vecchia (1999) and Kaser et al. (2003), the stakes should be arranged so as to evenly sample as much of the elevation as possible. Using this approach, the expert-knowledge designs (for $n \leq 10$) are shown in Figure 7.5.

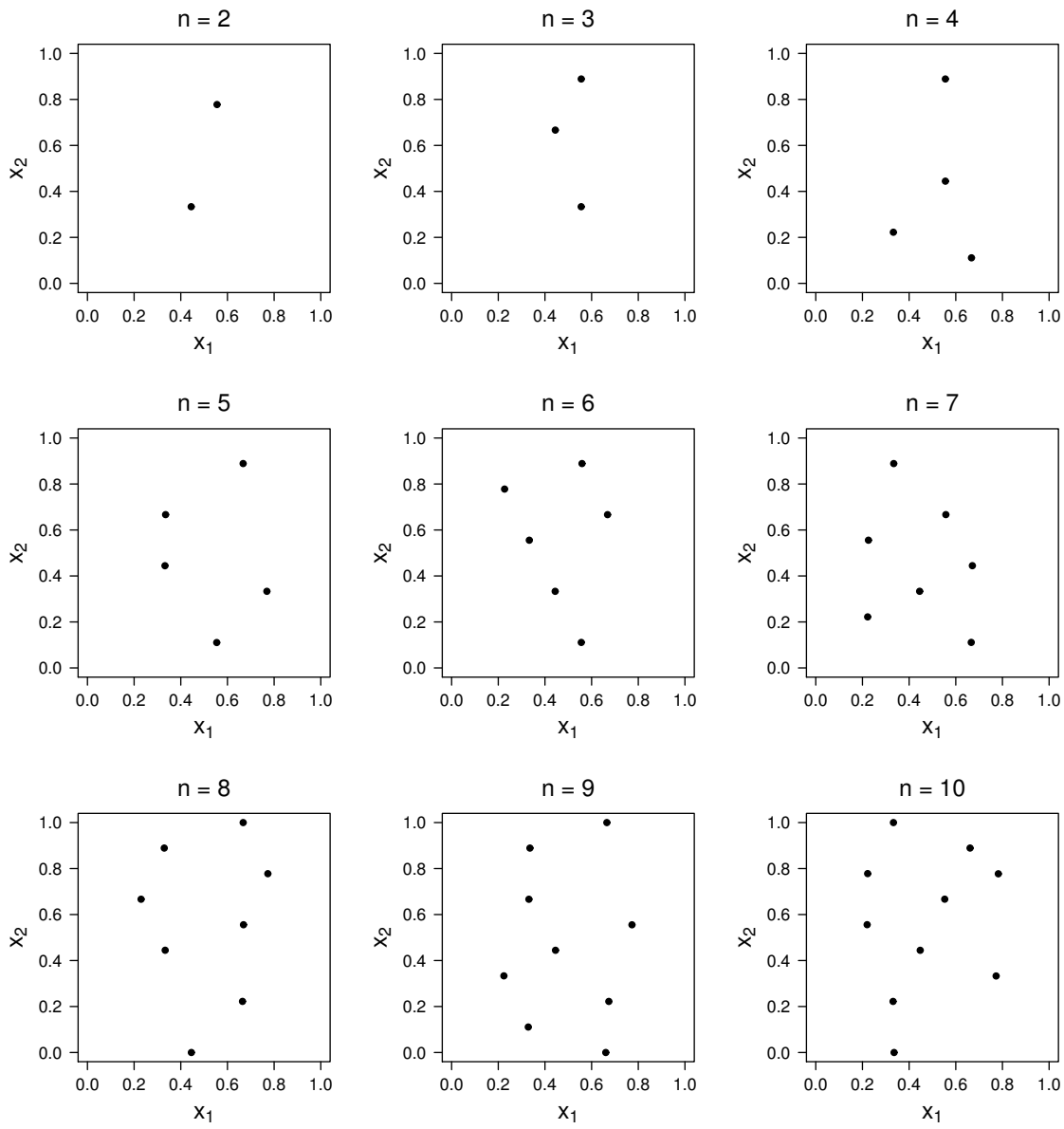

Figure 7.3: ave-IMSPE optimal designs for the glacier application, for $n \leq 10$.

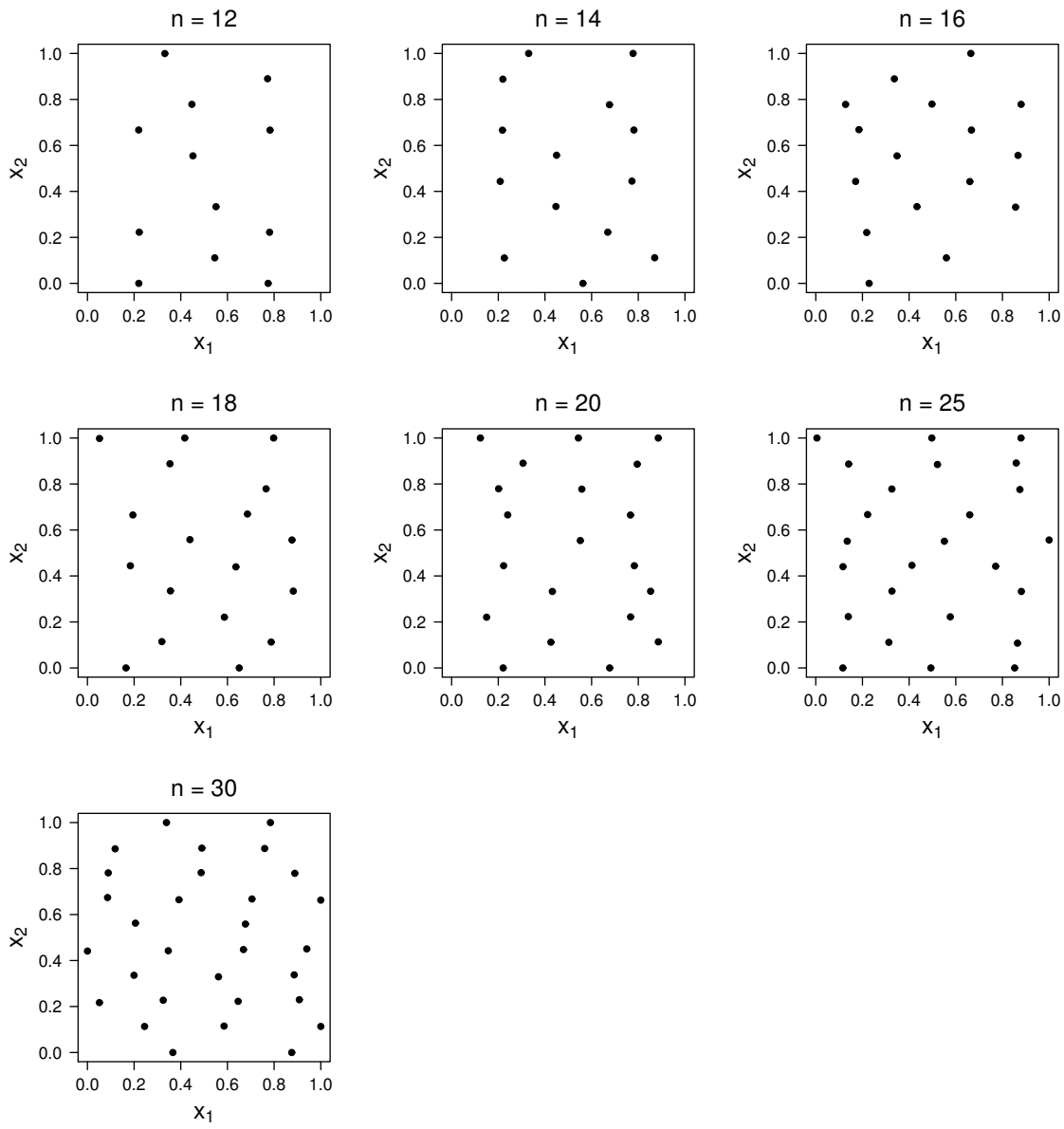

Figure 7.4: ave-IMSPE optimal designs for the glacier application, for $10 < n \leq 30$.

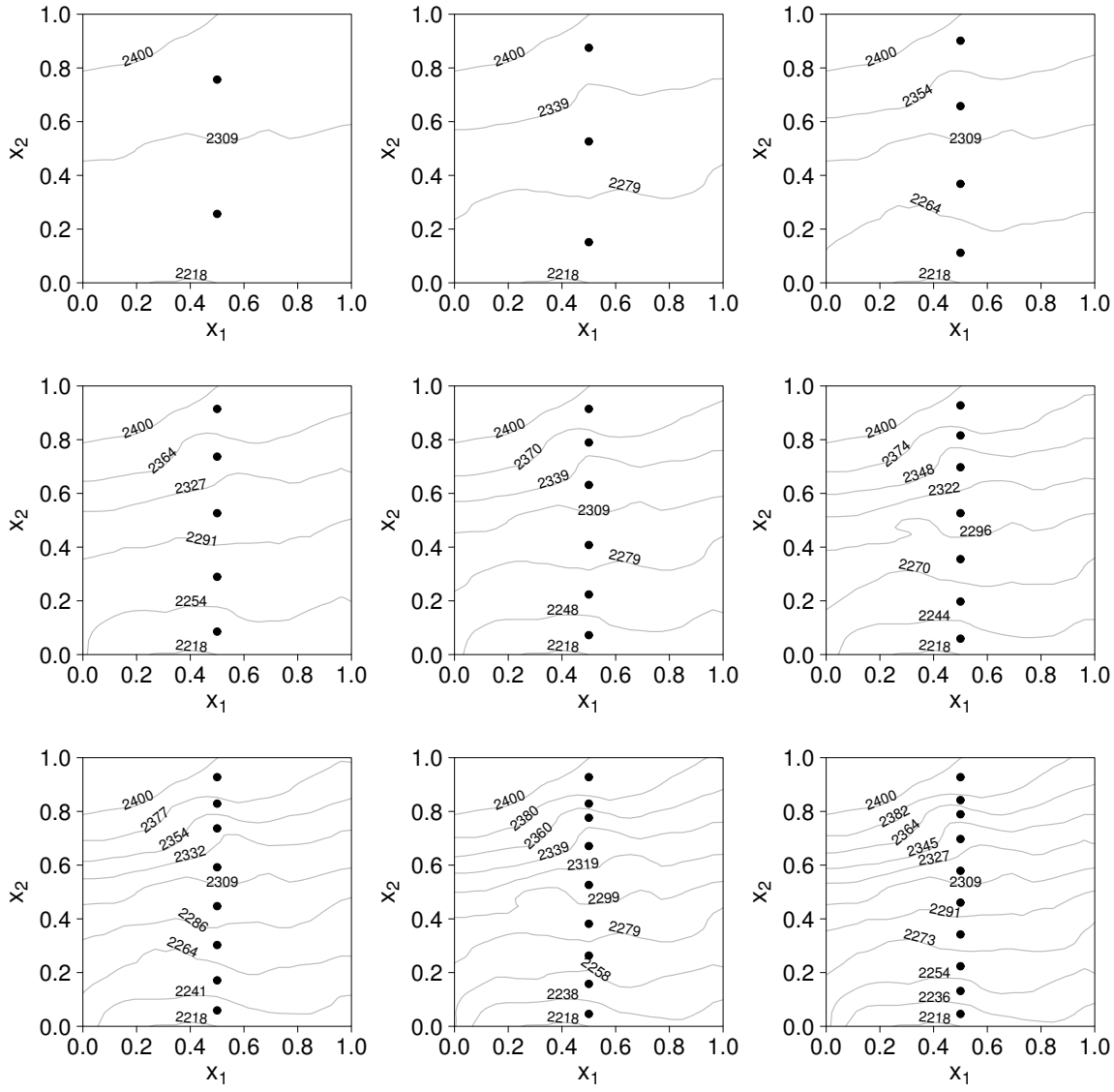

Figure 7.5: Expert-knowledge designs for the glacier application, along with elevation contours (m above sea level), for $n \leq 10$.

Table 7.3 and Figure 7.6 show the resulting scaled ave-IMSPE values for the ave-IMSPE optimal designs, maximin designs and expert-knowledge designs, for each n. As expected, the ave-IMSPE optimal designs perform better than either of the two alternatives, and the value of ave-IMSPE decreases with n. Also, as the value of n increases past $n = 20$, the ave-IMSPE for the optimal designs begins to display the same convex shape as in the simulation results in Chapter 6. Note also that the difference in ave-IMSPE between the optimal designs, maximin designs and expert-knowledge designs tends to increase with n.

Table 7.3: Scaled values of ave-IMSPE for the glacier application.

n	ave-IMSPE Optimal Design	Maximin Design	Expert-Knowledge Design
2	0.951	0.966	0.966
3	0.927	0.951	0.949
4	0.903	0.932	0.917
5	0.879	0.920	0.923
6	0.855	0.884	0.897
7	0.832	0.877	0.897
8	0.809	0.873	0.894
9	0.785	0.858	0.879
10	0.763	0.830	0.842
12	0.717	0.752	0.842
14	0.673	0.763	0.824
16	0.636	0.731	0.820
18	0.595	0.712	0.815
20	0.550	0.693	0.806
25	0.466	0.625	0.797
30	0.397	0.534	0.795

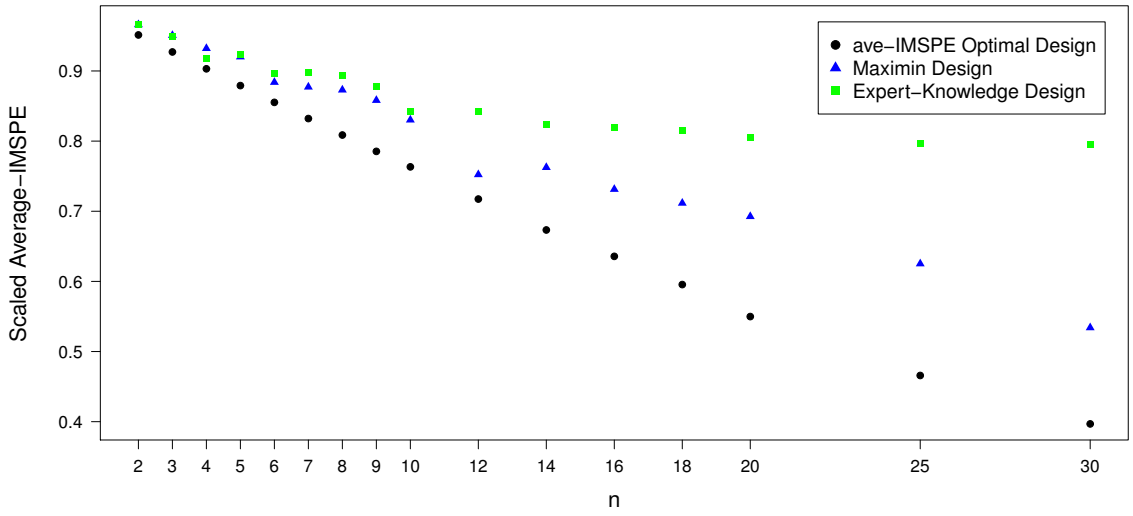

Figure 7.6: Scaled ave-IMSPE values at the optimal designs, maximin designs and expert-knowledge designs, for the glaciology application.

Chapter 8

Discussion and Further Work

In this project, a methodology has been presented for using a computer model to aid in the design of a physical experiment. The designs minimize the uncertainty of prediction, and are robust to changes in the computer model parameters, whose true values are not known. Thus, the resulting designs often display space-filling properties, with more points placed along the dimension that contains more variability in the response values. They are, in a sense, "averages" of the IMSPE-designs that would be obtained for any single set of computer model parameter values.

For the glacier application, the ave-IMSPE optimal designs all lie near the middle of the Easting dimension (x_1), and are spread out along the length of the Northing dimension (x_2). This result resembles the description of the expert-knowledge designs commonly used in the glaciology literature (e.g. Fountain and Vecchia, 1999; Kaser et al., 2003; Østrem and Brugman, 1991). However, since the ave-IMSPE optimal designs seek to minimize the uncertainty of prediction across the entire region of interest, they also contain some variability in the x_1-dimension. As a result, they provide a compromise between the centre line designs (expert-knowledge designs) and space-filling designs that seek to place points throughout the entire space (e.g. maximin designs). As shown in Chapter 7, they outperform both of these alternative approaches.

Several points must be considered for future work on this problem. Most importantly, an assumption that has been made is that the region of interest is a hyperrectangle in d dimensions. On the glacier, this corresponds to the rectangle defined in § 2.3. In the future, the methodology will be expanded, using the methods in Pratola et al. (2015), to be applicable to a non-rectangular (and non-convex) surface such as South Glacier.

Future work will also consider the relationship between designs that are optimal for predicting the spatial distribution of the response across the surface (i.e. designs that minimize ave-IMSPE, as in this project), and designs that are optimal for estimating the average or total response value across the surface. Much interest in glaciology lies in estimating the total melt across the glacier, and so this result would be beneficial to researchers.

Furthermore, uncertainty in the physical conditions represented by the computer model in § 2.2 can also be represented by a further parameter to the model: the year in which the field data was obtained. Data from South Glacier exist for several years, and considering them would allow the design criterion, ave-IMSPE, to more accurately incorporate the uncertainty in the parameters governing the system. Thus, the resulting optimal designs could more effectively be used to predict the spatial distribution of melt for any given year.

A final potential problem to be considered stems directly from an application in glaciological field work. During the melt season, the snow/ice level at certain points on the glacier may decrease to a level lower than the bottom of a stake. As a result, the measurement of snow/ice level at this stake cannot accurately be measured, as only a lower bound is known. This amounts to a type of censoring problem, and could be the focus of further studies.

Bibliography

Adler, D., Murdoch, D., Nenadic, O., Urbanek, S., Chen, M., Gebhardt, A., Bolker, B., Csardi, G., Strzelecki, A., and Senger, A. (2016). *rgl: 3D Visualization Using OpenGL*. R package version 0.95.1441.

Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., Dasgupta, P., Dubash, N. K., Edenhofer, O., Elgizouli, I., Field, C. B., Forster, P., Friedlingstein, P., Fuglestvedt, J., Gomez-Echeverri, L., Hallegatte, S., Hegerl, G., Howden, M., Jiang, K., Cisneros, B. J., Kattsov, V., Lee, H., Mach, K. J., Marotzke, J., Mastrandrea, M. D., Meyer, L., Minx, J., Mulugetta, Y., O'Brien, K., Oppenheimer, M., Pachauri, R. K., Pereira, J. J., Pichs-Madruga, R., Plattner, G.-K., Pörtner, H.-O., Power, S. B., Preston, B., Ravindranath, N. H., Reisinger, A., Riahi, K., Rusticucci, M., Scholes, R., Seyboth, K., Sokona, Y., Stavins, R., Stocker, T. F., Tschakert, P., van Vuuren, D., and van Ypersele, J.-P. (2014). *Climate Change 2014: Synthesis Report (Longer Report)*. Intergovernmental Panel on Climate Change.

Bales, R. C., McConnell, J. R., Mosley-Thompson, E., and Csatho, B. (2001). Accumulation over the Greenland ice sheet from historical and recent records. *Journal of Geophysical Research*, 106(D24):33813–33825.

Bansal, J. C., Singh, P., Saraswat, M., Verma, A., Jadon, S. S., and Abraham, A. (2011). Inertia weight strategies in particle swarm optimization. In *2011 Third World Congress on Nature and Biologically Inspired Computing (NaBIC)*, pages 640–647. IEEE.

Calaway, R., Revolution Analytics, Weston, S., and Tenenbaum, D. (2015). *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*. R package version 1.0.10.

Chen, R.-B., Chang, S.-P., Wang, W., Tung, H.-C., and Wong, W. K. (2015). Minimax optimal designs via particle swarm optimization methods. *Statistics and Computing*, 25(5):975–988.

Chen, R.-B., Hsieh, D.-N., Hung, Y., and Wang, W. (2013). Optimizing Latin hypercube designs by particle swarm. *Statistics and Computing*, 23(5):663–676.

Cogley, J. G. (1999). Effective sample size for glacier mass balance. *Geografiska Annaler: Series A, Physical Geography*, 81(4):497–507.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley-Interscience.

De Paoli, L. and Flowers, G. E. (2009). Dynamics of a small surge-type glacier using one-dimensional geophysical inversion. *Journal of Glaciology*, 55(194):1101–1112.

De Woul, M., Hock, R., Braun, M., Thorsteinsson, T., Jóhannesson, T., and Halldórsdóttir, S. (2006). Firn layer impact on glacial runoff: a case study at Hofsjökull, Iceland. *Hydrological Processes*, 20(10):2171–2185.

Diggle, P. J. and Ribeiro Jr., P. J. (2007). *Model-Based Geostatistics*. Springer-Verlag.

Flowers, G. E. and Clarke, G. K. C. (2000). An integrated modelling approach to understanding subglacial hydraulic release events. *Annals of Glaciology*, 31(1):222–228.

Flowers, G. E., Roux, N., Pimentel, S., and Schoof, C. G. (2011). Present dynamics and future prognosis of a slowly surging glacier. *The Cryosphere*, 5(1):299–313, doi:10.5194/tc–5–299–2011.

Fountain, A. G. and Vecchia, A. (1999). How many stakes are required to measure the mass balance of a glacier? *Geografiska Annaler: Series A, Physical Geography*, 81(4):563–573.

Fu, P. and Rich, P. M. (2000). *The Solar Analyst 1.0 Manual*. Helios Environmental Modeling Institute.

Gabbi, J., Carenzo, M., Pellicciotti, F., Bauder, A., and Funk, M. (2014). A comparison of empirical and physically based glacier surface melt models for long-term simulations of glacier response. *Journal of Glaciology*, 60(224):1140–1154.

Hock, R. (1999). A distributed temperature-index ice- and snowmelt model including potential direct solar radiation. *Journal of Glaciology*, 45(149):101–111.

Hock, R. and Jensen, H. (1999). Application of kriging interpolation for glacier mass balance computations. *Geografiska Annaler: Series A, Physical Geography*, 81(4):611–619.

Hock, R., Radić, V., and De Woul, M. (2007). Climate sensitivity of Storglaciären, Sweden: an intercomparison of mass-balance models using ERA-40 re-analysis and regional climate model data. *Annals of Glaciology*, 46(1):342–348.

Holmlund, P., Jansson, P., and Pettersson, R. (2005). A re-analysis of the 58 year mass-balance record of Storglaciären, Sweden. *Annals of Glaciology*, 42(1):389–394.

Huss, M., Bauder, A., Werder, M., Funk, M., and Hock, R. (2007). Glacier-dammed lake outburst events of Gornersee, Switzerland. *Journal of Glaciology*, 53(181):189–200.

Huss, M., Farinotti, D., Bauder, A., and Funk, M. (2008). Modelling runoff from highly glacierized alpine drainage basins in a changing climate. *Hydrological Processes*, 22(19):3888–3902.

Jansson, P. (1999). Mass balance of Storglaciären 1998/99. In Klingbjer, P., editor, *Tarfala Research Station Annual Report 1998–99*, pages 7–10. Department of Physical Geography at Stockholm University.

Jansson, P. and Pettersson, R. (2007). Spatial and temporal characteristics of a long mass balance record, Storglaciären, Sweden. *Arctic, Antarctic, and Alpine Research*, 39(3):432–437.

Johnson, M. E., Moore, L., and Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, 26(2):131–148.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492.

Kaser, G., Fountain, A., and Jansson, P. (2003). *A manual for monitoring the mass balance of mountain glaciers*. Unesco Paris.

Kaser, G., Großhauser, M., and Marzeion, B. (2010). Contribution potential of glaciers to water availability in different climate regimes. *Proceedings of the National Academy of Sciences of the United States of America*, 107(47):20223–20227.

Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948.

Klok, E. J., Jasper, K., Roelofsma, K. P., Gurtz, J., and Badoux, A. (2001). Distributed hydrological modelling of a heavily glaciated Alpine river basin. *Hydrological Sciences*, 46(4):553–570.

Kronenberg, M., Barandun, M., Hoelzle, M., Huss, M., Farinotti, D., Azisov, E., Usubaliev, R., Gafurov, A., Petrakov, D., and Kääb, A. (2016). Mass-balance reconstruction for Glacier No. 354, Tien Shan, from 2003 to 2014. *Annals of Glaciology*, 57(71):92–102.

Kuhn, M., Dreiseitl, E., Hofinger, S., Markl, G., Span, N., and Kaser, G. (1999). Measurements and models of the mass balance of Hintereisferner. *Geografiska Annaler: Series A, Physical Geography*, 81(4):659–670.

Leatherman, E. R., Dean, A. M., and Santner, T. J. (2014). Computer experiment designs via particle swarm optimization. In *Topics in Statistical Simulation*, pages 309–317. Springer.

Lemke, P. J., Ren, J., Alley, R. B., Allison, I., Carrasco, J., Flato, G., Fujii, Y., Kaser, G., Mote, P., Thomas, R. H., and Zhang, T. (2007). Observations: changes in snow, ice and frozen ground. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 337–383. Cambridge University Press.

MacDougall, A. H., Wheler, B. A., and Flowers, G. E. (2011). A preliminary assessment of glacier melt-model parameter sensitivity and transferability in a dry subarctic environment. *The Cryosphere*, 5(4):1011–1028.

Mandal, A., Wong, W. K., and Yu, Y. (2015). Algorithmic searches for optimal designs. In Dean, A., Morris, M., Stufken, J., and Bingham, D., editors, *Handbook of Design and Analysis of Experiments*, pages 755–783. CRC Press.

Minder, J. R., Mote, P. W., and Lundquist, J. D. (2010). Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. *Journal of Geophysical Research*, 115(D14122).

Moore, R. D., Fleming, S. W., Menounos, B., Wheate, R., Fountain, A., Stahl, K., Holm, K., and Jakob, M. (2009). Glacier change in western North America: influences on hydrology, geomorphic hazards and water quality. *Hydrological Processes*, 23(1):42–61.

Østrem, G. and Brugman, M. (1991). *Glacier mass-balance measurements: a manual for field and office work*. National Hydrology Research Institute.

Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M., and Corripio, J. (2005). An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. *Journal of Glaciology*, 51(175):573–587.

Peng, C.-Y. and Wu, C. F. J. (2014). On the choice of nugget in kriging modeling for deterministic computer experiments. *Journal of Computational and Graphical Statistics*, 23(1):151–168.

Pratola, M. T., Harari, O., Bingham, D., and Flowers, G. E. (2015). Design and analysis of experiments on non-convex regions. *Technometrics*.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radić, V., Bliss, A., Beedlow, A. C., Hock, R., Miles, E., and Cogley, J. G. (2013). Regional and global projections of twenty-first century glacier mass changes in response to climate scenarios from global climate models. *Climate Dynamics*, 42:37–58.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

Roustant, O., Ginsbourger, D., and Deville, Y. (2012). DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, 51(1):1–55.

Sacks, J., Schiller, S. B., and Welch, W. J. (1989a). Designs for computer experiments. *Technometrics*, 31(1):41–47.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989b). Design and analysis of computer experiments. *Statistical Science*, 4(4):409–423.

Santner, T. J., Williams, B. J., and Notz, W. I. (2003). *The Design and Analysis of Computer Experiments*. Springer Series in Statistics.

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York. ISBN 978-0-387-75968-5.

Schoof, C., Rada, C. A., Wilson, N. J., Flowers, G. E., and Haseloff, M. (2014). Oscillatory subglacial drainage in the absence of surface melt. *The Cryosphere*, 8(3):959–976.

Schuler, T., Fischer, U. H., Sterr, R., Hock, R., and Gudmundsson, G. H. (2002). Comparison of modeled water input and measured discharge prior to a release event: Unteraargletscher, Bernese Alps, Switzerland. *Nordic Hydrology*, 33(1):27–46.

Schuler, T. V., Hock, R., Jackson, M., Elvehøy, H., Braun, M., Brown, I., and Hagen, J.-O. (2005a). Distributed mass-balance and climate sensitivity modelling of Engabreen, Norway. *Annals of Glaciology*, 42(1):395–401.

Schuler, T. V., Loe, E., Taurisano, A., Eiken, T., Hagen, J. O., and Kohler, J. (2007). Calibrating a surface mass-balance model for Austfonna ice cap, Svalbard. *Annals of Glaciology*, 46(1):241–248.

Schuler, T. V., Melvold, K., Hagen, J. O., and Hock, R. (2005b). Assessing the future evolution of meltwater intrusions into a mine below Gruvefonna, Svalbard. *Annals of Glaciology*, 42(1):262–268.

Shi, P., Duan, K., Liu, H., Yang, J., Zhang, X., and Sun, J. (2016). Response of Xiao Dongkemadi Glacier in the central Tibetan Plateau to the current climate change and future scenarios by 2050. *Journal of Mountain Science*, 13(1):13–28.

Shi, Y. and Eberhart, R. (1998). A modified particle swarm optimizer. In *The 1998 IEEE International Conference on Evolutionary Computation Proceedings, IEEE World Congress on Computational Intelligence*, pages 69–73. IEEE.

Surjanovic, S. and Bingham, D. (2013). Virtual library of simulation experiments: Test functions and datasets. Retrieved June 1, 2016, from `http://www.sfu.ca/~ssurjano`.

Thibert, E., Blanc, R., Vincent, C., and Eckert, N. (2008). Glaciological and volumetric mass-balance measurements: error analysis over 51 years for Glacier de Sarennes, French Alps. *Journal of Glaciology*, 54(186):522–532.

Trüssel, B. L., Truffer, M., Hock, R., Motyka, R. J., Huss, M., and Zhang, J. (2015). Runaway thinning of the low-elevation Yakutat Glacier, Alaska, and its sensitivity to climate change. *Journal of Glaciology*, 61(225):65–75.

U.S. Department of the Interior (2013). Glossary of glacier terminology – text version. Retrieved June 30, 2016, from `http://pubs.usgs.gov/of/2004/1216/text.html`.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.

Wasserman, L. (2013). *All of Statistics: A Concise Course in Statistical Inference*. Springer Texts in Statistics.

Wheler, B. A., MacDougall, A. H., Flowers, G. E., Petersen, E. I., Whitfield, P. H., and Kohfeld, K. E. (2014). Effects of temperature forcing provenance and extrapolation on the performance of an empirical glacier-melt model. *Arctic, Antarctic, and Alpine Research*, 46(2):379–393.

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1):1–29.

Wilson, N. J., Flowers, G. E., and Mingo, L. (2013). Comparison of thermal structure and evolution between neighboring subarctic glaciers. *Journal of Geophysical Research: Earth Surface*, 118(3):1443–1459.

Zhang, Y., Liu, S., Xu, J., and Shangguan, D. (2008). Glacier change and glacier runoff variation in the Tuotuo River basin, the source region of Yangtze River in western China. *Environmental Geology*, 56(1):59–68.

Zhu, Z. and Evangelou, E. (2015). Optimal design for spatial models. In Dean, A., Morris, M., Stufken, J., and Bingham, D., editors, *Handbook of Design and Analysis of Experiments*, pages 549–573. CRC Press.