

Predicting Snow Velocity in Large Chute Flows Under Different Environmental Conditions

Jonathan Rougier*
Department of Mathematics
University of Bristol, UK

Martin Kern
Department of Snow and Avalanches
BFW Institute for Natural Hazards and Alpine Timberline
Innsbruck, Austria

February 8, 2010

Abstract

Observations, model evaluations, and expert judgements are combined to make predictions of snow velocity in large chute experiments. Different experimental variables, namely the environmental conditions snow density and snow-surface temperature, affect all aspects of this inference. We show how the effect of these two variables can be incorporated into our judgements regarding the uncertain parameters of the physical model, the discrepancy between the physical model and reality, and the observation error. We adopt a Bayes linear approach to avoid the necessity of fully-probabilistic belief specifications, and demonstrate visual tools for statistical validation. Our results represent an important first step in improving the specification of uncertainty in model-based avalanche hazard mapping.

KEYWORDS: Avalanche, Bayes linear, computer experiment, Herschel-Bulkley rheology, model discrepancy

1 Introduction

It is very important to understand the behaviour of avalanches: how, for example, the speed of an avalanche depends upon the inclination of the slope, on the snow density, and on the snow temperature. One source of data is from avalanches themselves but, for obvious reasons, these opportunistic

*Corresponding author: Department of Mathematics, University Walk, Bristol BS8 1TW, U.K. Email j.c.rougier@bristol.ac.uk.

observations are hard to come by. Experiments provide a second source of data. In this paper we focus on chute experiments carried out at the WSL Institute for Snow and Avalanche Research SLF, at Davos, Switzerland.

These experiments take place under different environmental conditions, partly by design and partly through circumstance. We would like to extrapolate from the experiments we have, to learn about snow velocities across a range of plausible environmental conditions. For this purpose we introduce a mathematical model which expresses the velocity profile as a function of these conditions. The general idea is to use the mathematical model to construct a joint probability distribution over the experiments we have and those environmental conditions for which we would like predictions, and then the observations are assimilated into the predictions by probabilistic conditioning.

The statistical field of Computer Experiments is concerned with combining model evaluations and observations. A particular challenge in this field is to account for the fact that some of the model parameters are imperfectly known, and that the model itself is imperfect (Kennedy and O’Hagan, 2001; Craig et al., 2001; Goldstein and Rougier, 2004, 2009). This challenge becomes more acute when the model-outputs and the system behaviour are multivariate. In our application, for example, the model output is functional, and the system is observed at a discrete set of abscissae (heights). But our application also introduces a further complication: variations in the environmental conditions. One methodological contribution of this paper is to show how environmental variables can be included in a computer experiment, taking account of the fact that they can affect all aspects of the statistical model that links the model parameters, the model evaluations and the system, and the system observations. Although our environmental variables are what are sometimes termed *concomitant* variables (Cox, 1958), the method we describe extends immediately experimental variables as well. A second statistical contribution is to demonstrate a detailed elicitation, including validation, for a complex physical process.

The outline of the paper is as follows. Section 2 describes the background to the experiments, and the mathematical model we adopt. Section 3 describes the statistical framework that we use to combine model evaluations and observations, over a range of different environmental conditions. Section 4 describes our choice of inferential treatment, the Bayes linear approach, and the simplifications that follow from it. Sections 5 and 6 describe our statistical modelling choices, and the results of our analysis, including diagnostic assessment. Section 7 concludes. An Appendix describes our approach for specifying random quadratic functions.

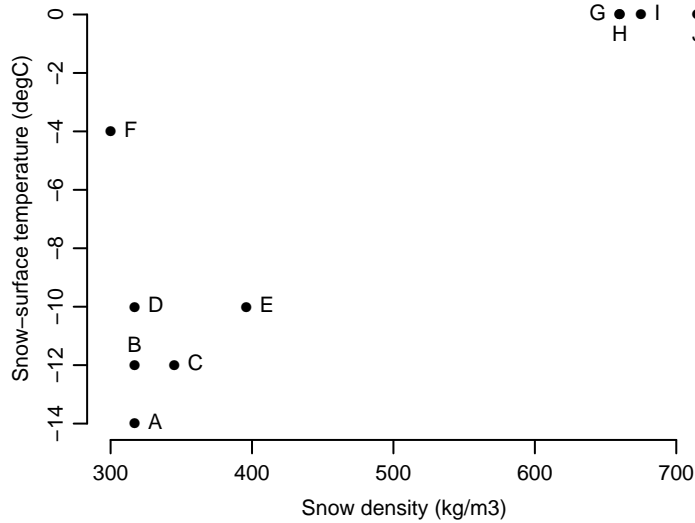


Figure 1: Environmental conditions in the ten experiments (labelled A, \dots, J), indexed by snow density (kg/m^3) and snow-surface temperature ($^{\circ}\text{C}$).

2 The experiments and the physical model

In this section we present a summary of the chute and the experiments, and describe the notational modifications we have made to the physical model. A more detailed account can be found in Kern et al. (2004), and the references below.

The chute and the experiments. Our experiments were performed on the SLF snow chute at Weissfluhjoch, which is 34 m long and 2.5 m wide. In each experiment, 8.4 m^3 of snow was released from a hopper at the top of the chute. The snow then accelerated along a 10 m section, at the end of which it was approximately in steady state. Downslope velocities of the flow were measured by an optical sensor array placed in a half-wedge at the centreline of the chute, half a metre beyond this section. This array comprised 40 sensors placed in eight rows of 5 sensors each, with a vertical spacing of 13 mm and a horizontal spacing of about 10 mm.

For our analysis we use ten experiments which were performed under different environmental conditions, indexed by the snow density ρ , the ambient air temperature T_a and the snow-surface temperature T_{ss} . For the snow and air temperatures, we use measurements recorded by an automatic weather station at the nearby Weissfluhjoch experimental site. The environmental conditions for the ten experiments are summarised in Figure 1, and

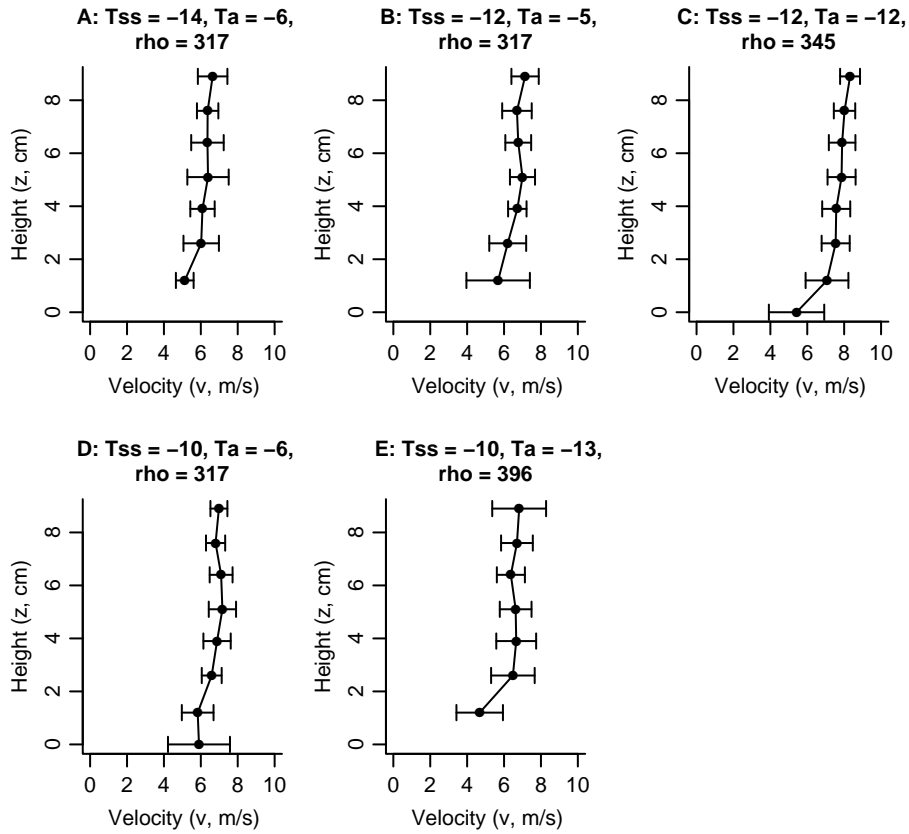


Figure 2: Those experiments with low snow-surface temperatures. T_{ss} : snow-surface temperature ($^{\circ}\text{C}$); T_a : atmospheric temperature ($^{\circ}\text{C}$); ρ : snow density (kg/m^3). The dots indicate the measurements and the error bars ± 2 standard deviations; the measurements have been interpolated to make the velocity profile easier to see. Note that the velocity plots are oriented using the common convention in snow science, in which velocity (the response) is plotted on the horizontal axis and height (= distance above the chute floor, in a direction normal to the chute floor) is plotted on the vertical axis.

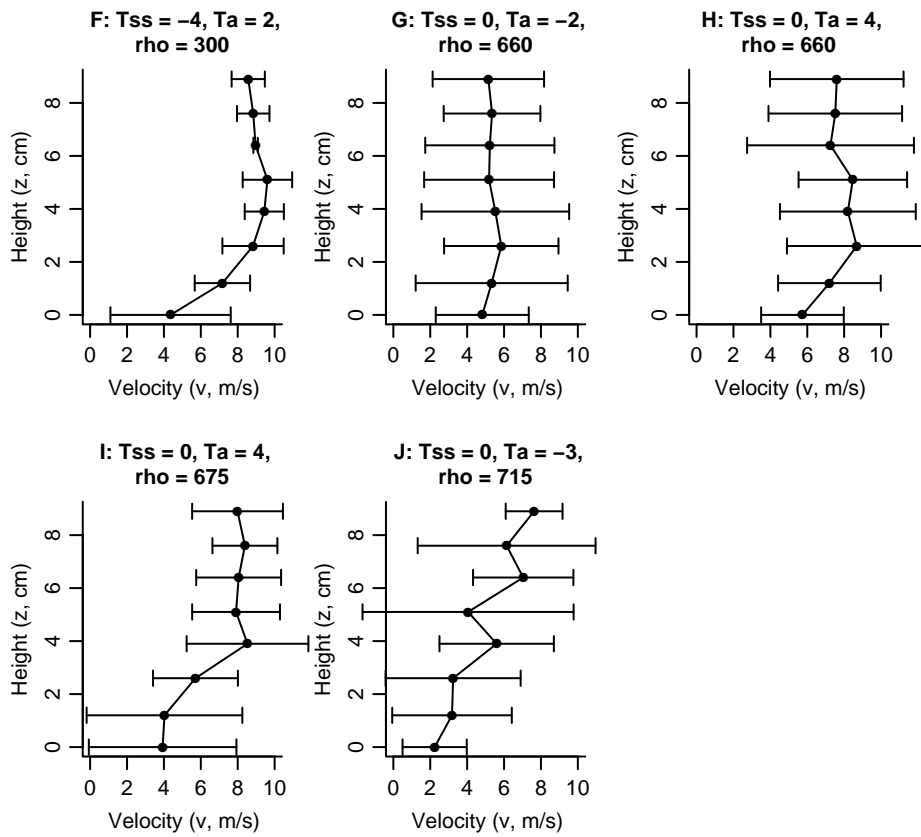


Figure 3: Those experiments with high snow-surface temperatures. See the caption to Figure 2 for details.



Figure 4: Photograph of the SLF snow-chute at Weissfluhjoch, in Davos, Switzerland. The circle in the centre of the picture indicates the half-wedge containing the optical sensor array. Photograph: M. Schaefer, SLF.

the observations are given in Figures 2 and 3.

For a detailed description of the measurement setup, the working principle of the optical velocity sensors, and a discussion of the systematic measurement errors, see Tiefenbacher and Kern (2004), Kern et al. (2004), and McElwaine and Tiefenbacher (2005).

The Herschel-Bulkley model. Observations on flowing avalanches (Dent et al., 1997; Sovilla et al., 2008) and from chute experiments suggest that the flow of snow is characterised by a relatively thick layer where the shear rates (dv/dz) are low or even zero. This so-called ‘plug-layer’ travels on a comparatively thin shear-layer in which $dv/dz > 0$. Basal sliding ($v(0) > 0$) may occur in this shear-layer, depending on the environmental conditions. In avalanche dynamics, *rheology* is used to denote the closure scheme of the equations of motion, in terms of a relation between internal forces and deformation of the flowing snow. The simplest rheology which is able to reproduce the observed snow behaviour is the Bingham rheology (Bingham, 1922; Oldroyd, 1947). However, Kern et al. (2004) show that qualitatively better fits can be obtained using a generalised Bingham rheology, the so-called Herschel-Bulkley model (see, e.g., Barnes et al., 1989).

The equations of the Herschel-Bulkley (HB) model, adapted for our purposes, are presented in Figure 5, with a simple schematic of the main features in Figure 6. The HB model is described in detail in Kern et al. (2004); here we outline one particular feature that helps us to identify the set of model parameters about which we are uncertain, and for which we can specify a marginal distribution. The uncertain parameters in the original model are: v_0 , denoting basal slip velocity (m/s); τ_c , the yield stress (Pa) at zero shear rate; α , a unitless exponent (shape parameter); and K , which enters into the stress equation for a positive shear rate:

$$\text{stress} = \tau_c + K \left(\frac{dv}{dz} \right)^\alpha \quad \frac{dv}{dz} > 0, \quad (1)$$

where $v(z)$ is the velocity profile at height z . K is a complicated quantity which has units that depend on α (the units are Pa s^α). This makes it hard to elicit K : it would be simpler if we could decouple K from α , in such a way that we would be comfortable treating these two quantities as probabilistically independent. Therefore we reparameterise the stress relationship as

$$\text{stress} = \tau_c \left[1 + \left(t_c \frac{dv}{dz} \right)^\alpha \right] \quad \frac{dv}{dz} > 0, \quad (2)$$

where t_c has units of time (seconds). According to (2), $t_c = (K/\tau_c)^{1/\alpha}$, and t_c replaces K as our fourth uncertain model parameter. Neither K nor t_c are well-defined physically because HB is an a model that, while motivated by physical considerations, is nonetheless primarily empirical in nature. Our

The Herschel-Bulkley Model

The velocity profile has the form:

$$v(z) = \begin{cases} v_h + (v_0 - v_h) \left(1 - \frac{z}{h}\right)^{\frac{1+\alpha}{\alpha}} & 0 \leq z < h \\ v_h & z \geq h, \end{cases} \quad (\text{HB-a})$$

where

$$v_h = v_0 + \frac{h}{t_c} \frac{\alpha}{1 + \alpha} \left(\frac{h}{H - h}\right)^{1/\alpha} \quad (\text{HB-b})$$

and h solves

$$\tau_c = (H - h)g\rho \sin \theta \quad (\text{HB-c})$$

subject to $h \geq 0$, with $h = 0$ otherwise.

Notation

z	Height ordinate (m)	$v(z)$	Velocity (m/s)
h	Height to plug-layer (m)	v_h	plug-layer velocity

Treated as known

θ	Inclination (32°)	g	Acceleration (9.8 m/s^2)
H	Height to top of flow (0.4 m)		

Environmental variables

ρ	Snow density (kg/m^3)	T_{ss}	Snow-surface temperature ($^\circ\text{C}$)
--------	----------------------------------	-----------------	---

Uncertain model parameters

v_0	Basal velocity (m/s)	α	Stress coefficient
τ_c	Yield stress (Pa) at $dv/dz = 0$	t_c	Time constant (s)

(the variable T_{ss} does not appear explicitly in the model, but it influences our judgements about the best value of τ_c).

Figure 5: Velocity profile of a steady 2D Herschel-Bulkley flow of snow. Our alterations to the standard notation are described in Section 2. See Figure 6 for a schematic.

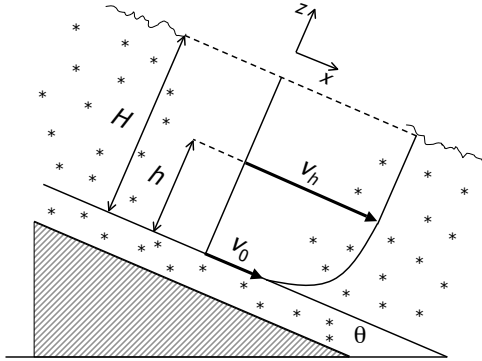


Figure 6: A simple schematic of a typical velocity profile of steady Herschel-Bulkley flow.

marginal distribution for t_c will therefore be largely defined by its predictive implications (see section 5).

Model-consistent approximate values for the model parameters and model results are given in Figure 7, along with the resulting velocity profile.

3 Outline of the statistical inference

3.1 General features

Our objective is to estimate the velocity profile under different experimental conditions. We write this velocity profile in the general form $v(\mathbf{m}, z)$, where z indicates height (normal to the surface) along the velocity profile, which we term the *index variable*, and \mathbf{m} denotes the value of relevant environmental variables. In our case these variables are

$$\mathbf{m} = (\rho, T_{ss}),$$

where ρ is the snow density and T_{ss} is the snow-surface temperature; we also have information on the atmospheric temperature but we have not used it in this analysis.

We have three sources of information. First, we have the outcome of experiments on the large chute. These observations include measurement error, and are distinguished from the true results by writing $v^{\text{obs}}(\mathbf{m}, z)$. We have ten experiments, and for simplicity we will suppose that we record velocities at the same eight heights for each experiment—in fact not all the heights were recorded for every experiment: we make this simplification purely to avoid an extra layer of subscripts. Therefore our observations

Variable	App. value
g	10 m/s ²
H	0.5 m
$\sin \theta$	1/2
ρ	500 kg/m ³
T_{ss}	-5 °C
α	2
v_0	5 m/s
v_h	10 m/s
h	0.1 m
t_c	$2/3 \times 10^{-2}$ s
τ_c	1000 Pa

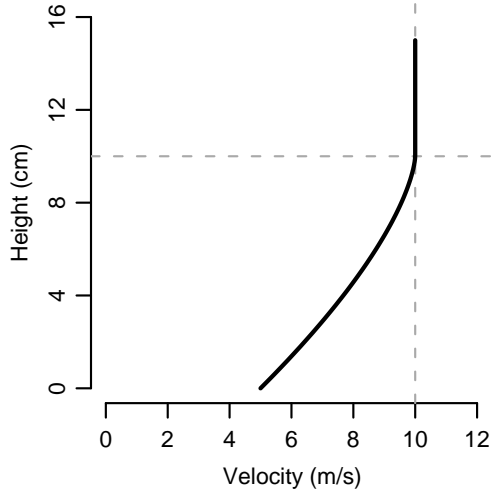


Figure 7: Model-consistent approximate values for the model parameters and results (lefthand table) and the resulting velocity profile (righthand figure). The T_{ss} value in the table is just for reference.

comprise the matrix

$$V^{\text{obs}} = \begin{pmatrix} v^{\text{obs}}(\mathbf{m}_1, z_1) & \dots & v^{\text{obs}}(\mathbf{m}_1, z_8) \\ \vdots & \ddots & \vdots \\ v^{\text{obs}}(\mathbf{m}_{10}, z_1) & \dots & v^{\text{obs}}(\mathbf{m}_{10}, z_8) \end{pmatrix}; \quad (3)$$

a matrix in which the rows correspond to (potentially) different values for the environmental variables and the columns to the different values for the index variable.

Our second source of information is evaluations of a physical model, namely the Herschel-Bulkley (HB) model described in Figure 5. This model can be seen as a function mapping (\mathbf{m}, z) into a scalar output. But there will be a further set of model-inputs, namely those model parameters about which we are uncertain. This uncertainty has two sources. First, the model contains semi-empirical relationships that stand in for physics we do not understand, or which we choose not to represent fully. These empirical relationships may have uncertain coefficients. Second, deficiencies in the model compromise the interpretation of the model's parameters, even those with well-defined physical meanings. For example, the value of shear stress in actual snow in the chute may not be the best value to use for shear stress in the model. Therefore, although we are guided by the physical interpretation of the parameters, we do not necessarily want to fix them at their physical values. These uncertain parameters are denoted \mathbf{x} ,

$$\mathbf{x} = (v_0, \tau_c, \alpha, t_c).$$

The model output is then denoted $g(\mathbf{x}, \mathbf{m}, z)$. Below we will treat the ‘best’ values of \mathbf{x} as unknown, and allow our judgements about these best values to be influenced by the environmental variables, \mathbf{m} .

Our third source of information is our judgements about the physical model, the actual behaviour of snow in the chute, the observations, and the relationships between them. Specifying these judgements occurs in two stages: first we construct a joint statistical model over all uncertain and observed quantities, describing our conditional independence choices. Then we quantify the marginal and conditional distributions that occur in this statistical model. Our choices will be informed by the physics of snow in large chutes, by the feasibility of the elicitation, and by the tractability of the resulting statistical inference. In this respect constructing a statistical model is no different from the process of constructing a physical model.

3.2 Statistical framework

For clarity, we start by describing a statistical framework for a single experiment and a single height, dropping \mathbf{m} and z from the notation, and writing $g(\mathbf{x})$ for the physical model, v for the actual value, and v^{obs} for the observation. The standard approach to constructing a joint statistical framework is to assert the existence of a ‘best’ value of the model parameters, denoted \mathbf{x}^* (Goldstein and Rougier, 2004, 2006, 2009). Then we link the evaluations of the physical model and the actual system behaviour through the model evaluated at this ‘best’ input. We can visualise this in the form of a Directed Acyclic Graph (or, DAG, see, e.g., Cowell et al., 1999), in which the vertices represent uncertain quantities, and the absence of an edge between vertices indicates an assertion of conditional independence. The DAG in this case is simply

$$\mathbf{x}^* \xrightarrow{g} v \longrightarrow v^{\text{obs}} \quad (4)$$

where the only simplification we have made in the joint structure is to choose $v^{\text{obs}} \perp\!\!\!\perp \mathbf{x}^* \mid v$, a completely standard and uncontroversial choice.

The superscript on the edge from \mathbf{x}^* to v indicates the ‘location’ of the HB model in the inference, but it does not imply that $v = g(\mathbf{x}^*)$, but simply that the distribution for v will be expressed conditionally on $g(\mathbf{x}^*)$. This edge represents our statistical model of the *discrepancy* between the model and reality: a typical form of this conditional distribution might be

$$\pi(v \mid \mathbf{x}^*) = \varphi(v; g(\mathbf{x}^*), \sigma^2(\mathbf{x}^*)) \quad (5)$$

where φ is the Gaussian density function with specified mean and variance, and we specify the variance $\sigma^2(\mathbf{x})$ as an explicit function of the model parameters. A common simplification in (5) is to make $\sigma^2(\mathbf{x})$ invariant to \mathbf{x} , i.e. set $\sigma^2(\mathbf{x}) = \sigma^2$, a scalar. A regrettable further simplification is to set $\sigma^2 = 0$, which asserts that the model has no structural error, and that it is

only uncertainty about \mathbf{x}^* that prevents us from performing a perfect evaluation. This, unfortunately, has been the dominant practice in much applied science, where practitioners have either not been aware that it is possible to incorporate structural error into their analysis, or have been reluctant to quantify it. This reluctance to quantify can also be seen in the choice of marginal distribution $\pi(\mathbf{x}^*)$, which is often taken to be rectangular with specified limits. This class of distribution is supposed, mistakenly, to be the ‘neutral’ choice.

The second edge represents our statistical model of the measurement processes: a typical form for this distribution might be

$$\pi(v^{\text{obs}} | v) = \varphi(v^{\text{obs}}; v, \tau^2) \quad (6)$$

where for simplicity we treat measurement error as invariant to flow speed, and specify the size of the error in terms of a standard deviation τ .

Multiple heights. Generalising from one value of the index variable height to a collection of heights presents no conceptual problems. We write (4) as

$$\mathbf{x}^* \xrightarrow{g} \mathbf{v} \longrightarrow \mathbf{v}^{\text{obs}} \quad (7)$$

where $\mathbf{v} = (v(z_1), \dots, v(z_8))$ and $\mathbf{v}^{\text{obs}} = (v^{\text{obs}}(z_1), \dots, v^{\text{obs}}(z_8))$.

The scalar variance function $\sigma^2(\mathbf{x})$ in (5) is replaced by a more general vectorised relationship such as

$$\text{Cov}(v(z_j), v(z_{j'}) | \mathbf{x}^*) = \sigma(\mathbf{x}^*, z_j) \times \sigma(\mathbf{x}^*, z_{j'}) \times \kappa_z(z_j, z_{j'}), \quad (8)$$

where $\kappa_z(\cdot)$ is a correlation function, and for simplicity we treat the correlation structure as invariant to \mathbf{x}^* . Our judgements of how the physical model’s performance varies according to both \mathbf{x} and z are represented in $\sigma(\mathbf{x}, z)$. For example, the HB velocity profile comprises two parts, one for heights below the shear layer height h , and a much simpler (vertical) part for heights equal to or above h , where h is a known function of \mathbf{x} (see Figure 7). If we judge the simpler model more likely to be in error, then our $\sigma(\mathbf{x}, z)$ will be a function of both \mathbf{x} and z .

Our judgements about systematic model bias are represented in the correlation function $\kappa_z(\cdot)$. If, for example, we believe that the model will tend to systematically over- or under-predict velocity across the heights, then we would specify a positive correlation for $\kappa_z(z_j, z_{j'})$, indicating that $v_j - g(\mathbf{x}^*, z_j)$ and $v_{j'} - g(\mathbf{x}^*, z_{j'})$ are likely to have the same sign, particularly if z_j and $z_{j'}$ are proximate. We will use this representation of systematic model bias in section 5.3.

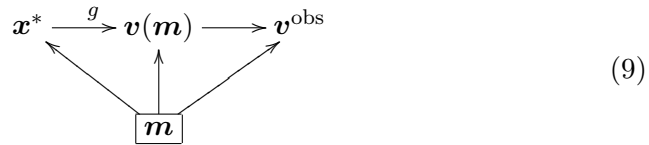
In a similar way, the scalar measurement error variance τ^2 in (6) is replaced by the matrix T . We need to include common sources of variation

in the off-diagonal elements of T . The dominant source of these is measurements made by the same sensor, which might have a bias. In our large chute, though, there are different sensors at each height, so this is not an issue. However, the sensors are identical, and so there might be a common source of variation from a fault that is particular to sensors of this general type. We have not incorporated this possibility in our analysis below, however, and the matrix T is treated as diagonal.

3.3 Including environmental variables

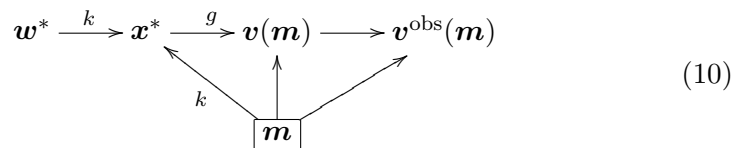
The purpose of including environmental variables in our framework is to allow us to predict the behaviour of snow in a chute under environmental conditions we have not observed. For this to be possible, we must believe that there is some relationship between $\mathbf{v}(\mathbf{m})$ and $\mathbf{v}(\mathbf{m}')$ where \mathbf{m}' is not dissimilar to \mathbf{m} , so that an experiment at \mathbf{m} is informative about what happens at \mathbf{m}' .

Explicit single experiment. We start by including the environmental variables explicitly in our statistical model for a single treatment:



where the box around \mathbf{m} denotes that it is specified, not uncertain. This indicates that our judgements about \mathbf{x}^* , about our physical model's discrepancy, and about the observation error might all depend on the values of the environmental variables that prevail at the time.

When we introduce multiple experiments, though, we are faced with a problem. We would like to abstract the best value of the model parameters from a particular experiment, so that this best value becomes a generic quantity that is applicable across a range of experiments. However, our judgements about the best value of the model parameters are influenced by the environmental variables. The solution is to propose a higher level of model-parameters, that are invariant to the value of the environmental variables. Formally, we suppose that we can specify an uncertain vector \mathbf{w}^* with marginal density $\pi(\mathbf{w}^*)$, and a vector-valued deterministic function $\mathbf{x} = \mathbf{k}(\mathbf{w}, \mathbf{m})$, such that $\mathbf{k}(\mathbf{w}^*, \mathbf{m})$ has the distribution $\pi(\mathbf{x}^*; \mathbf{m})$, for all \mathbf{m} . This gives



Here, \mathbf{w}^* is not influenced by \mathbf{m} , and \mathbf{x}^* for any particular set of environmental variables is induced from \mathbf{w}^* and \mathbf{m} . Therefore learning about the ‘generic’ model parameters is learning about \mathbf{w}^* . We will illustrate judgements of this form in section 5.1, for the two model parameters v_0 and τ_c , which will depend on ρ and T_{ss} , respectively.

Multiple experiments. In order to handle multiple experiments we concatenate the actual values of the environmental variables with the values at which we would like predictions. Denote these latter values as $\mathbf{m}_a, \mathbf{m}_b, \dots$. We write the collection as

$$V = \begin{pmatrix} v(\mathbf{m}_1, z_1) & \dots & v(\mathbf{m}_1, z_8) \\ \vdots & \ddots & \vdots \\ v(\mathbf{m}_{10}, z_1) & \dots & v(\mathbf{m}_{10}, z_8) \\ \hline v(\mathbf{m}_a, z_1) & \dots & v(\mathbf{m}_a, z_8) \\ v(\mathbf{m}_b, z_1) & \dots & v(\mathbf{m}_b, z_8) \\ \vdots & \ddots & \vdots \end{pmatrix}. \quad (11)$$

In this format, the statistical model over multiple experiments and predictions can be written

$$\mathbf{w}^* \xrightarrow{g,k} V \longrightarrow V^{\text{obs}} \quad (12)$$

where the different values of \mathbf{m} are now internalised in V and V^{obs} . Looking back to (4), we seem to have come full-circle, with the important difference that we have had to redefine the model-parameters to ensure that they can be expressed independently of the environmental variables.

With this modification, the statistical modelling of the two conditional distributions can be extended quite naturally from the single experiment. For the discrepancy we might take the collection $\{v(\mathbf{m}_i, z_j) \mid \mathbf{w}^*\}$ to be jointly Gaussian with mean $g(\mathbf{k}(\mathbf{w}^*, \mathbf{m}_i), \mathbf{m}_i, z_j)$ and variance function

$$\begin{aligned} \text{Cov}(v(\mathbf{m}_i, z_j), v(\mathbf{m}_{i'}, z_{j'}) \mid \mathbf{w}^*) = \\ \sigma(\mathbf{k}(\mathbf{w}^*, \mathbf{m}_i), \mathbf{m}_i, z_j) \times \sigma(\mathbf{k}(\mathbf{w}^*, \mathbf{m}_{i'}), \mathbf{m}_{i'}, z_{j'}) \times \\ \kappa_m(\mathbf{m}_i, \mathbf{m}_{i'}) \times \kappa_z(z_j, z_{j'}) \end{aligned} \quad (13)$$

where $\kappa_m(\cdot)$ and $\kappa_z(\cdot)$ are correlation functions, and for simplicity we take the joint correlation structure as invariant to \mathbf{x}^* and separable in the environmental variables and the index variable.

For the observation error, we might take the collection $\{v^{\text{obs}}(\mathbf{m}_i, z_j) \mid V\}$ to be jointly Gaussian, and simplify by treating $v^{\text{obs}}(\mathbf{m}_i)$ as conditionally independent of $v^{\text{obs}}(\mathbf{m}_{i'})$ given $v^{\text{obs}}(\mathbf{m}_i)$, with mean $v^{\text{obs}}(\mathbf{m}_i)$, and variance function

$$\text{Cov}(v^{\text{obs}}(\mathbf{m}_i, z_j), v^{\text{obs}}(\mathbf{m}_{i'}, z_{j'}) \mid V) = \begin{cases} T_{jj'} & i = i' \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

3.4 Summary

We summarise the statistical requirements described in this section, under the simplifications we have made. Note that the simplifications are not critical: we have made them to clarify the presentation, and also because they seem reasonable in our application and, perhaps, more widely. More importantly, the Gaussian form of the conditional distributions is not critical either. In fact, in the following sections we will dispense with it, and proceed using a Bayes Linear approach.

1. We specify the deterministic function $\mathbf{k}(\cdot)$ and the marginal distribution \mathbf{w}^* so that the inferred marginal distribution of $\mathbf{x}^* = \mathbf{k}(\mathbf{w}^*, \mathbf{m})$ is consistent with our judgements about \mathbf{x}^* with environmental variables \mathbf{m} .
2. We quantify our judgement about the discrepancy in our physical model in terms of the standard deviation function $\sigma(\mathbf{x}, \mathbf{m}, z)$, and the two correlation functions $\kappa_m(\mathbf{m}, \mathbf{m}')$ and $\kappa_z(z, z')$; see eq. (13).
3. We quantify our judgements about the measurement errors in our observations in terms of the variance matrix T ; see eq. (14).

4 Bayes linear inference

Our intention is to predict velocities at values for the environmental variables that we have not observed. To make this more concrete, in this paper we will predict the velocity for a range of densities and snow-surface temperatures. If we want to predict on, say, a 21×21 grid in these two variables then V will have $10 + 21^2$ rows and eight columns. i.e. comprise about 3600 components. Therefore this is quite a large inference. At the same time, though, we are aware that our data, although the best of their kind, are noisy, and our physical model is rather simple. Therefore our judgements will play a large part in our predictions. This prioritises diagnostic information. It also makes us cautious about putting more structure in our judgements than we would willingly specify. For these reasons we strongly favour a Bayes linear analysis. The Bayes linear approach is outlined in Goldstein (1999) and described in detail in Goldstein and Wooff (2007); it has proved very powerful in large computer experiments (Craig et al., 1997, 2001; Goldstein and Rougier, 2006). It also underpins standard techniques such as Dynamic Linear Models (West and Harrison, 1997), also known as the Kalman filter.

In the Bayes linear approach, expectation is taken as primitive and judgements are specified in terms of the mean and variance of a collection of quantities. Therefore we are not required to make higher-order specifications, in contrast to the fully probabilistic approach. The resulting updating equations have a simple form that allows rapid computation, including of diag-

nostic information. Our predictions take the form of a mean vector and a variance matrix over the product of our specified values for the experiments $\mathbf{m}_a, \mathbf{m}_b, \dots$ and the abscissae of the velocity profile.

To implement the Bayes linear approach for prediction we must specify a mean and variance over the collection $\{V, V^{\text{obs}}\}$. If we want to do model calibration we must also include \mathbf{w}^* in that collection. Calibration is more complicated in a Bayes linear framework, because of the strong non-linearities that can exist between \mathbf{w}^* and V , induced mainly by non-linearities in the physical model $g(\cdot)$. The Bayes linear approach to calibration is described in Goldstein and Rougier (2006). In this paper we will focus on prediction, using the approach described in Craig et al. (2001).

We can write our statistical model in the general form

$$V \equiv G^* + D^* \tag{15a}$$

$$V^{\text{obs}} \equiv V + E \tag{15b}$$

where G^* is the collection of model evaluations with typical component $g(\mathbf{x}^*, \mathbf{m}_i, z_j)$, D^* is the collection of discrepancies with typical component

$$d(\mathbf{x}^*, \mathbf{m}_i, z_j) = v(\mathbf{m}_i, z_j) - g(\mathbf{x}^*, \mathbf{m}_i, z_j), \tag{15c}$$

and E is the collection of measurement errors with typical component

$$e(\mathbf{m}_i, z_j) = v^{\text{obs}}(\mathbf{m}_i, z_j) - v(\mathbf{m}_i, z_j). \tag{15d}$$

We induce a mean and variance on the collection $\{V, V^{\text{obs}}\}$ in terms of our choices for $\{G^*, D^*, E\}$. According to our choices in section 3.1, $E \perp\!\!\!\perp \{G^*, D^*\}$, has mean zero, and variance given by (14). D^* and G^* , however, covary, because they share a common source of uncertainty, namely \mathbf{x}^* . Including this covariance is a challenge for the Bayes linear approach. We judge that our purpose is better served by removing the dependence of D^* on \mathbf{x}^* . In fact, this is the standard approach in computer experiments where, as far as we are aware, no analysis has yet included the influence of the model parameters on the model’s discrepancy, not even in a fully probabilistic approach; see, e.g., the standard set-up in Kennedy and O’Hagan (2001) and the discussion in Rougier (2007). In our statistical model this means removing the effect of \mathbf{x} from the standard deviation function $\sigma(\mathbf{x}, \mathbf{m}, z)$, used in (13); we will show in section 5.3 how the lack of \mathbf{x} can be partially mitigated through the creative use of \mathbf{m} . Now we have $E \perp\!\!\!\perp G^* \perp\!\!\!\perp D$, where we have dropped the ‘*’ on D .

An alternative way of viewing the probabilistic independence between G^* and D is as an identification constraint, which helps us to generate judgements about V in two parts: judgements about the best input \mathbf{x}^* , and judgements about the discrepancy D . This separation makes judgements more transferable between inferences, since it is then possible to focus on

one or other of \boldsymbol{x}^* and D . Having said that, there are some concerns about whether it is coherent to treat \boldsymbol{x}^* and D as independent, discussed in Goldstein and Rougier (2009): at the moment, though, these concerns have not impacted on current practice (see the ensuing discussion and rejoinder in Goldstein and Rougier, 2009).

With this simplification we can write the joint mean and variance as

$$E\begin{pmatrix} V \\ V^{\text{obs}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}^* \\ H\boldsymbol{\mu}^* \end{pmatrix} \quad (16a)$$

$$\text{Var}\begin{pmatrix} V \\ V^{\text{obs}} \end{pmatrix} = \begin{pmatrix} \Xi & \Xi H^T \\ H\Xi & H\Xi H^T + \text{Var}(E) \end{pmatrix} \quad (16b)$$

where H is the incidence matrix, which picks out the observations, $\Xi = \text{Var}(V) = \Sigma^* + \text{Var}(D)$, and $\boldsymbol{\mu}^*$ and Σ^* are the mean and variance of G^* . This still leaves us to determine the mean and variance of G^* . For this purpose we find it helpful to specify a probability distribution for \boldsymbol{w}^* , and then to infer $\boldsymbol{\mu}^*$ and Σ^* using samples drawn from G^* . This is a cheap calculation in our application because the physical model is quick to evaluate; where the physical model is expensive we would use an emulator (see, e.g., O’Hagan, 2006; Rougier, 2008). This approach is not simply a case of estimating $\boldsymbol{\mu}^*$ and Σ^* , though. Our initial choice for the distribution $\pi(\boldsymbol{w}^*)$ will be informed by our judgements, but we are likely to tune this choice in the light of the implied mean and variance for G^* and, possibly, to modify the resulting mean and variance in the light of other judgements. This is particularly the case for t_c which, as we observed in section 2, is a model-parameter without a natural analogue. Thus the distribution $\pi(\boldsymbol{w}^*)$ is a contrivance that helps us to specify $\boldsymbol{\mu}^*$ and Σ^* , rather than a core part of our inference.

The scheme for adjusting our mean and variance for V on the basis of observed value for V^{obs} is described in Goldstein and Wooff (2007), ch. 3. The updating equations will be familiar because they are also the conditioning relations of a multivariate Gaussian distribution. However, as already explained, we have adopted the Bayes linear approach partly because of our reluctance to provide fully-probabilistic descriptions of our uncertainty, and therefore we do not judge (V, V^{obs}) to be Gaussian, and nor will we be using the extra structure that this would imply in our predictions. For example, our predictions are explicitly in terms of means, variances, and covariances: we cannot report quantiles without further restrictions.

5 Statistical modelling and results

5.1 Model parameters

In this subsection we specify \mathbf{w}^* , the function $\mathbf{k}(\cdot)$, and the marginal distribution $\pi(\mathbf{w}^*)$, as described in section 3.3. Recall from section 3.1 that $\mathbf{x} = (v_0, \tau_c, \alpha, t_c)$ and $\mathbf{m} = (\rho, T_{ss})$. The outline in this subsection is of necessity a very partial description of the way in which we arrived at our quantification. We deliberated for many hours, knowing that our choices would be influential in the predictions. We started with simple choices for each of the uncertain quantities, and then gradually iterated towards a joint choice. Throughout we were guided by the diagnostic plots given in section 5.4.

First we consider the dependence of v_0^* on ρ , which appears to be a feature of our observations. For snow in the coexistence regime of snow and water at 0°C , snow density and snow water content are typically positively related. High water content causes additional basal friction by lubrication-adhesion effects on the ground, especially if the flow is water-dominated, as in slush flows (Jaedicke et al., 2008). Compact ($\rho > 300 \text{ kg/m}^3$), dry snow exhibits slightly less basal friction on dry ground. On the other hand, for low snow densities ($\rho \leq 300 \text{ kg/m}^3$), a lower ratio between gravitational forces and basal friction may contribute to slower basal slip velocities: low snow densities are frequently associated with a fine-grained snow structure which may result in a higher effective basal friction coefficient. We judge that these two effects will probably combine in a concave relationship between ρ and v_0^* over our range of densities.

Second, we consider the dependence of τ_c on T_{ss} . Foehn et al. (1998) and Schweizer (1998) observed that τ_c depends both on the snow temperature and on the applied shear rate. For our Herschel-Bulkley flow, τ_c is the threshold stress for failure of the plug at $z = h$ where the shear rate dv/dz vanishes (note that, under our experimental conditions, this transition from shear flow to plug flow might not be perfectly smooth). We assign a value to τ_c approximately matching this model definition by using the experimental results for τ_c under very low shear rates, of the order $dv/dz \sim 10^{-3} \text{ s}^{-1}$. Under these conditions, τ_c is decreasing with increasing temperature.

The Appendix outlines a simple approach for quantifying both of these dependencies in terms of some basic elicitations. In summary, we restrict the form of the dependence to a quadratic with uncertain coefficients, and then we constrain the mean and variance of the coefficients. In the case of the basal slip velocity v_0^* (m/s), our constraints are (1) $E(v_0^*; \rho = 250 \text{ kg/m}^3) = 4.6 \text{ m/s}$ and $E(v_0^*; \rho = 800 \text{ kg/m}^3) = 2.5 \text{ m/s}$; (2) extremum at $\rho = 400 \text{ kg/m}^3$; (3) $\text{Pr}(\text{concave}) = 0.95$. In the case of the yield stress τ_c^* (Pa), our constraints are (1) $E(\tau_c^*; T_{ss} = -15^\circ\text{C}) = 1150 \text{ Pa}$ and $E(\tau_c^*; T_{ss} = -4^\circ\text{C}) = 950 \text{ Pa}$; (2) extremum at $T_{ss} = 1^\circ\text{C}$; (3) $\text{Pr}(\text{concave}) = 0.05$. In each case

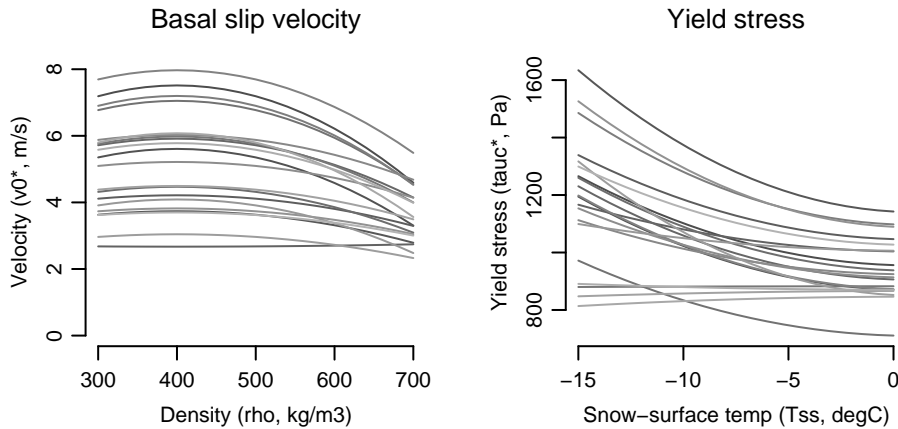


Figure 8: Realisations of the stochastic relationship between the environmental variables and the ‘best’ value of the model parameters.

we have one free statistical parameter with which to tune our choices, and we choose $\sigma_0 = 0.7$ m/s for v_0^* and $\sigma_0 = 100$ Pa for τ_c^* (σ_0 is defined in the Appendix). The results are shown in Figure 8, as realisations of the random functions. In fact, these realisations were our primary tool in setting the constraints and the free parameter to achieve a distribution of random functions that reflected our judgements.

Finally, we consider α^* and t_c^* . We found these the most difficult to quantify. For α^* , inspection of (HB-b) indicates that all values in the range $[0, \infty)$ are acceptable. Therefore we choose an Exponential distribution with mean 2. For t_c^* , however, values too close to zero give plug-layer velocities (v_h) that are too high. Therefore we selected a lower limit for t_c^* on the basis of an upper limit for v_h . A high value for v_h is associated with a thick shear layer, and the thickest shear layer we expect is 0.15 m, with $\rho = 700$ kg/m³ and $\tau_c^* = 900$ Pa. If we set the upper limit for v_h at 15 m/s, then this translates into a lower limit for t_c^* of 0.01 s. Therefore we choose an Exponential distribution for $t_c^* - 0.01$, with a mean of 0.01.

It is interesting to note that this lower limit for t_c^* is actually higher than the value we used in Figure 7, demonstrating that introducing uncertainty into our reasoning about model parameters can lead to conclusions that are inconsistent with results from a simpler ‘plug-in’ approach.

To summarise, our \mathbf{w} has the form

$$\mathbf{w} = (\beta_0, \beta_1, \beta_2, \gamma_0, \gamma_1, \gamma_2, \alpha, t_c) \quad (17)$$

and our function \mathbf{k} satisfying $\mathbf{x} = \mathbf{k}(\mathbf{w}, \mathbf{m})$ is

$$\mathbf{k}(\mathbf{w}, \mathbf{m}) = \begin{cases} \beta_0 + \beta_1 \rho + \beta_2 \rho^2 & (= v_0) \\ \gamma_0 + \gamma_1 T_{\text{ss}} + \gamma_2 (T_{\text{ss}})^2 & (= \tau_c) \\ \alpha \\ t_c \end{cases} \quad (18)$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ and $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2)$ are random coefficients used to represent the functional relationships described above and in the Appendix. We treat $\boldsymbol{\beta}^*, \boldsymbol{\gamma}^*, \alpha^*$ and t_c^* as probabilistically independent.

5.2 Observation error variance matrix

We treat the observation errors as uncorrelated, and use pooled estimates for the variances, based on whether $T_{\text{ss}} \leq -2^\circ\text{C}$. There are simple physical reasons for thinking that the accuracy of the measurements will depend on temperature. At a snow temperature of 0°C , snow is in the phase transition regime. That is, both snow and water are present in the flow. The measurement principle of the velocity sensors is based on a correlation analysis of signals obtained from infra-red (IR) reflectivity sensors. The quality of the signal is a function of the water content of the snow: the ambient water blurs the signal peaks related to passing snow particles. Moreover, the transition spectrum of fluid water has a gap in the IR band used for the reflectivity measurements which causes the water film between snow particles and sensor to dampen the signals. For a detailed discussion of this technical problem, see Guenther (2006). Strictly speaking, fluid water is present only for $T_{\text{ss}} = 0^\circ\text{C}$. But since T_{ss} was not measured in the snow chute but in a nearby study-plot, we use the slightly lower value of -2°C . The estimated standard deviations are 0.536 m/s for the low temperatures, and 1.670 m/s for the high temperatures. Note that this cut-off assigns experiment F to the low-temperature group, which is consistent with the measurement uncertainties shown in Figures 2 and 3.

5.3 Discrepancy variance matrix

Standard deviation function. We would like to express the standard deviation of $d(\mathbf{x}^*, \mathbf{m}_i, z_j)$, denoted $\sigma(\mathbf{x}^*, \mathbf{m}_i, z_j)$ in general, as a two-level function of h and z :

$$\sigma(\mathbf{x}^*, \mathbf{m}, z) = \begin{cases} \sigma_\ell & z < h \\ \sigma_u & z \geq h, \end{cases} \quad (19)$$

as explained in section 3.2. From (HB-c),

$$h = H - \frac{\tau_c}{g\rho \sin \theta} \quad (20)$$

and so h depends on both $\tau_c^* \in \{\mathbf{x}^*\}$ and $\rho \in \{\mathbf{m}\}$. However, as explained in section 4, for tractability we choose to exclude \mathbf{x}^* from $\sigma(\cdot)$. But since h is linear in τ_c and the expectation of τ_c^* depends on \mathbf{m} , we can express $E(h; \mathbf{m})$ as a function of \mathbf{m} . Thus we replace h in (19) with its expectation, to give

$$\sigma(\mathbf{m}, z) = \begin{cases} \sigma_\ell & z < E(h; \mathbf{m}) \\ \sigma_u & z \geq E(h; \mathbf{m}). \end{cases} \quad (21)$$

We choose the values $\sigma_\ell = 1.0$ m/s and $\sigma_u = 1.5$ m/s.

Correlation functions. We choose to treat $\kappa_m(\cdot)$ as separable in ρ and T_{ss} , so that

$$\kappa_m(\mathbf{m}, \mathbf{m}') = \kappa_\rho(\rho, \rho') \times \kappa_t(T_{\text{ss}}, T_{\text{ss}}'). \quad (22)$$

We treat both correlation functions as stationary, and use the Matérn family. In each case we use the same roughness parameter, $\nu = 5/2$, reflecting our judgement that none of the relationships is particularly smooth. In this case the Matérn correlation function has the form

$$\kappa(d; \ell) = \left(1 + \frac{\sqrt{5}d}{\ell} + \frac{5d^2}{3\ell^2} \right) \exp\left(\frac{-\sqrt{5}d}{\ell} \right) \quad (23)$$

where d is Euclidean distance, and ℓ is a scale parameter that sets the correlation length (see, e.g., Rasmussen and Williams, 2006, ch. 4). The expected number of up-crossings of zero in the unit interval for this correlation function is $u = \sqrt{5/3} (1/2\pi\ell)$, which allows us to choose ℓ in an intuitive manner, by thinking about the typical size of regions of the input space where the model consistently over- or under-predicts, which can be translated into the rate of upcrossings on a per-input basis. We choose the scale parameters to give one upcrossing, on average, over the intervals $T_{\text{ss}} \in [-15, 0]^\circ\text{C}$ and $\rho \in [300, 700]$ Pa. One up-crossing gives sample paths that are ‘quadratic’, indicating a typical region size of half the total width.

We must also specify the correlation function $\kappa_z(\cdot)$, which represents our judgements about systematic effects in the model bias, as explained in section 3.2. If we were to choose $\kappa_z(z_j, z_{j'}) = 0$ for $j \neq j'$, then we would be asserting that there was no systematic bias. We believe that there is a systematic bias, and so we use a Matérn correlation function for $\kappa_z(\cdot)$, with a scale parameter chosen to give half an up-crossing, on average, over the interval $z \in [0, 15]$ cm. This difference in the expected number of upcrossings compared to $\kappa_\rho(\cdot)$ and $\kappa_t(\cdot)$ reflects our judgement that the discrepancy is more systematic in z than in the two environmental variables. Quantifying these correlation lengths is not easy, but in our case the variance in V is dominated by variance in G^* , and so these choices are not critical.

Conditioning at zero height. We make one further modification to the variance matrix $\text{Var}(D)$. One of the model inputs is v_0 , the basal slip velocity. This is also one of the outputs, corresponding to $z = 0$. We identify this input and output by conditioning $d(\mathbf{m}, z)$ on $d(\mathbf{m}, 0) = 0$ for all (\mathbf{m}, z) ; in practice we condition on $d(\mathbf{m}_g, 0) = 0$ for all \mathbf{m}_g in a dense regular grid. After this modification, the only source of uncertainty about $v(\mathbf{m}, 0)$ is $(v_0^*; \mathbf{m})$.

5.4 Diagnostics and re-modelling

Prior predictive mean and standard deviation. It is difficult for us to make judgements about the marginal distribution $\pi(\mathbf{w}^*; \mathbf{m})$ directly, because the semi-empirical nature of our physical model means that the quantities \mathbf{w}^* and \mathbf{x}^* are not operationally defined. Therefore, we make these judgements partly indirectly, by examining their implications for the velocity profile at different values of the environmental variables, which *is* operationally defined. The velocity profile at any particular \mathbf{m} will be an uncertain quantity that synthesises the HB model and our statistical choices. We treat the HB model as inviolate, so that if the velocity profile does not accord with our judgements, then we need to change the marginal distribution of \mathbf{w}^* , the mapping function $\mathbf{k}(\cdot)$, or the variance function of the discrepancy. Once the velocity profile looks about right (a more formal criterion is not possible here), we know at least that our judgements on the two domains, \mathbf{w}^* and $v(\mathbf{m}, z)$, are consistent.

Figure 9 shows our main prior predictive summary diagnostic: the mean and standard deviation for the velocity profile in a 3×3 layout of snow densities and snow-surface temperatures. Each panel shows both the total uncertainty, and, inside that, the uncertainty attributable to uncertainty in the model parameters: the difference is uncertainty attributable to the discrepancy. The first source of uncertainty dominates, more so at higher densities. To us these figures ‘look right’, both in terms of the qualitative structure (e.g. stronger shearing at higher snow densities, especially for higher temperatures), and also in terms of the means and standard deviations. The prior standard deviations are large, but that is to be expected when we are predicting such a complicated system based on such a simple model. We used this diagnostic mainly in setting the marginal distribution for \mathbf{w}^* , as described in 5.1.

Joint structure. Figure 9 summarises the marginal structure of our judgements, i.e. taken pointwise at different values for the environmental variables. Our judgements about the joint structure are much less well-formed: certainly not well-enough formed that we might use them as the basis for further adjustments to our choices for the marginal distribution of \mathbf{w}^* . However, it is still interesting to see what our choices and the HB model imply

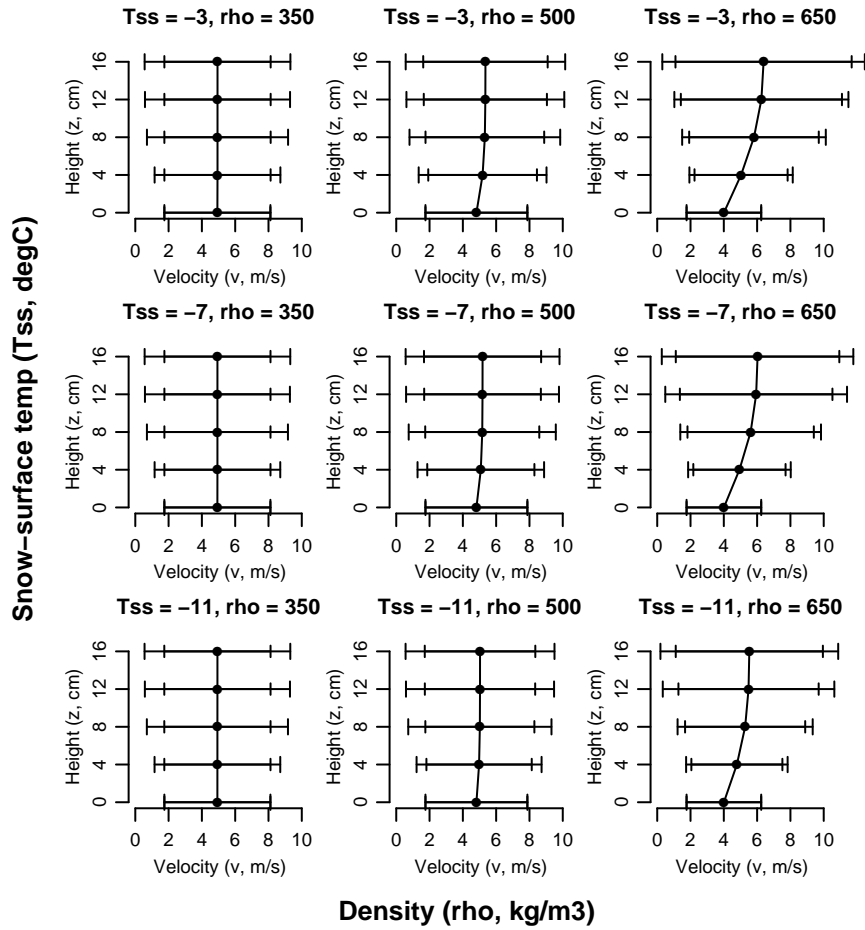


Figure 9: Prior predicted velocities on a 3×3 grid in snow density (horizontal) and snow-surface temperature (vertical). The share attributable to uncertainty in the model parameters is shown by the ticks inside each error bar, with the rest being due to the discrepancy between the model and actual snow behaviour.

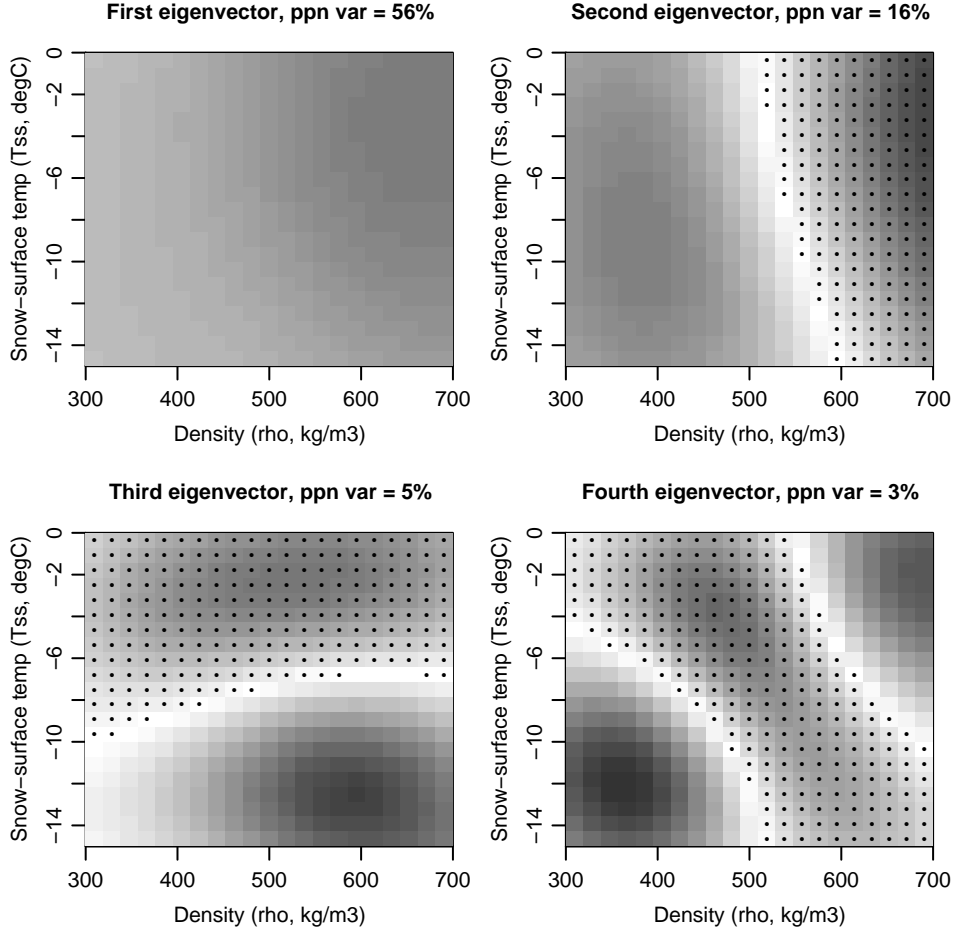


Figure 10: First four eigenvectors of the prior predicted plug-layer velocity variance matrix. The grey-scale indicates absolute size, while the dots indicate positive values.

for the covariance of velocities over the treatment space.

For simplicity, we restrict attention to a single height, $z = 0.4$ m. The velocity $v(\mathbf{m}, 0.4)$ is the velocity of the plug-layer, v_h . From (HB-b), this is a linear function of v_0 and a non-linear function of τ_c (through h , see (HB-c)), α , and t_c . Therefore our uncertainty about v_h will be affected by both of the environmental variables, since ρ affects h directly and v_0^* indirectly, and T_{ss} affects τ_c^* indirectly. To visualise the joint structure of this uncertainty, we compute the spectral decomposition of the variance matrix of v_h over a high-resolution grid in the two environmental variables, and plot the first four eigenvectors: see Figure 10.

The first eigenvector accounts for about 50% of total variation, and describes uncertainty about the general height of the function, without much

differentiation according to the values of the two environmental variables. The second eigenvector (about 20%) describes the tilt of the function along a fulcrum roughly parallel to the temperature axis, and the third eigenvector (about 5%) picks out the high-density low-temperature corner (this is the least likely corner in practice). All of the remaining eigenvalues are quite small.

From the point of view of experimental design, we conclude that experiments with extreme values of snow density would be more valuable at reducing uncertainty over the range of the environmental variables than those with extreme values of snow-surface temperature. This accords with our intuition that density plays a larger role than temperature. But note the *caveat* that a quarter of the total variation is made up of many small contributions. Therefore many experiments will be required to reduce uncertainty substantially everywhere: certainly more than the ten experiments we currently have.

Leave-one-out diagnostic. For our observation-based diagnostic we use leave-one-out plots. For each experiment in turn, we predict the observations using the outcome of the other nine experiments, and compare the prediction with the actual observations. Each prediction takes the form of a mean vector and a variance matrix, therefore we transform the joint prediction errors so that they ought to have mean zero and standard deviation one, and be uncorrelated, using the pivoted Cholesky approach described by Bastos and O’Hagan (2009). In our case these standardised joint prediction errors look quite similar to the standardised marginal prediction errors (not shown), because the variance matrix is roughly diagonal; experiment F has the largest difference, which is consistent with its unique position in Figure 1. The resulting diagnostics for each experiment are shown in Figure 11.

This is not the first such plot we made, as we have permitted ourselves a small amount of tuning of this diagnostic, mainly in the choices of the discrepancy standard deviations σ_ℓ and σ_u (eq. 21), and the scale parameters in the correlation functions (eq. 23). While the location and spread of the standardised prediction errors is broadly satisfactory, there is clearly a systematic effect across the experiments, and also across height within several of the experiments, which would not be present if the physical and statistical modelling were entirely consistent with the observations. However, we judge these deviations to be tolerable for our application. Where more care is required, we could revisit our distributional choices, or we could generalise the statistical framework from the ‘best input’ to the ‘reified modelling’ approach (Goldstein and Rougier, 2009).

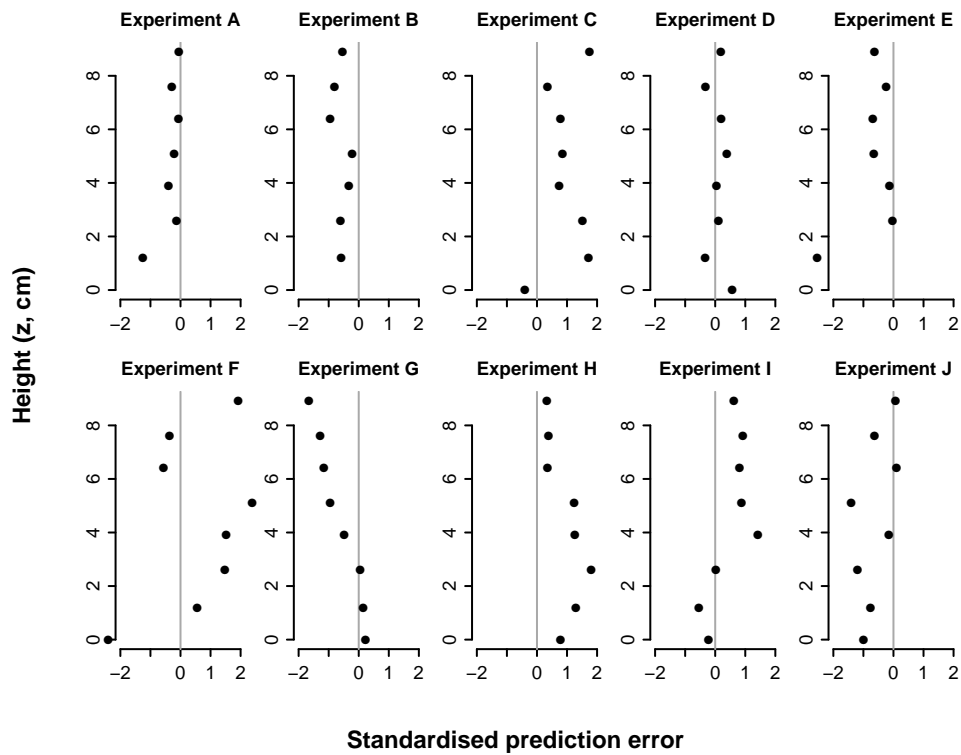


Figure 11: Leave-one-out diagnostic. The observations for each experiment in turn are predicted using observations from the other nine. The prediction errors from each experiment are standardised so that they ought to have mean zero and standard deviation one, and be uncorrelated, using the pivoted Cholesky approach of Bastos and O’Hagan (2009). Each panel shows the standardised prediction error on the horizontal axis and the height on the vertical axis.

6 Results

Our main interest is in the velocity profile across a range of values for the two environmental variables. For simplicity, we restrict attention to the velocity of the plug-layer (i.e. $z = 0.4$ m). We present both our prior assessment, and our assessment after adjusting by the observations from the ten experiments. The results are given in Figures 12 and 13. Each Figure shows the mean function, represented in terms of grey-scale and contours, and also an indication of the pointwise standard deviation, in terms of the size of the grey boxes: details are given in the caption to Figure 12.

In Figure 12, the prior mean field shows that snow density is more important than snow-surface temperature, but that there is an interaction between the two, so that temperature is much more influential when density is high. The mean and standard deviation functions are flat for densities below about 500 kg/m^3 . The prior standard deviations range from 2.2 m/s (low density, low temperature) to 3.6 m/s (high density, high temperature).

In Figure 13, the most noticeable effect of adjusting by the observations from our ten experiments is that the mean field has increased by about 1 m/s everywhere and the standard deviation has decreased everywhere; the interaction between high density and high temperature is also slightly more pronounced. This is consistent with the spectral analysis of the prior variance, which indicated that the overall level of the function was the dominant source of uncertainty. The spectral analysis also explains why uncertainty is reduced in the top-left and bottom-right corners, despite us having no experiments in these regions: the first two eigenvectors indicate that the experiments we have are informative for these regions. The adjusted standard deviations have been reduced to the range 1.5 m/s to 2.2 m/s . The reduction in uncertainty is most pronounced for high density and high temperature, where the standard deviation has fallen by 1.4 m/s .

7 Conclusion

In this study we have extended the ‘standard’ approach to model-based inference for complex physical systems to multiple experiments. This involves introducing an explicit role for the environmental variables in the assessment of the model’s discrepancy with reality, and also in the observation process. In our analysis of ten chute experiments using the Herschel-Bulkley (HB) model the environmental variables described environmental conditions, namely snow density and snow-surface temperature. We incorporated the effect of these into the variance function of the model discrepancy, where density affected the height of the shear layer, which interacted with the output index-variable, height. We also incorporated them into the observation error, to account for the larger observation errors that arise when the snow

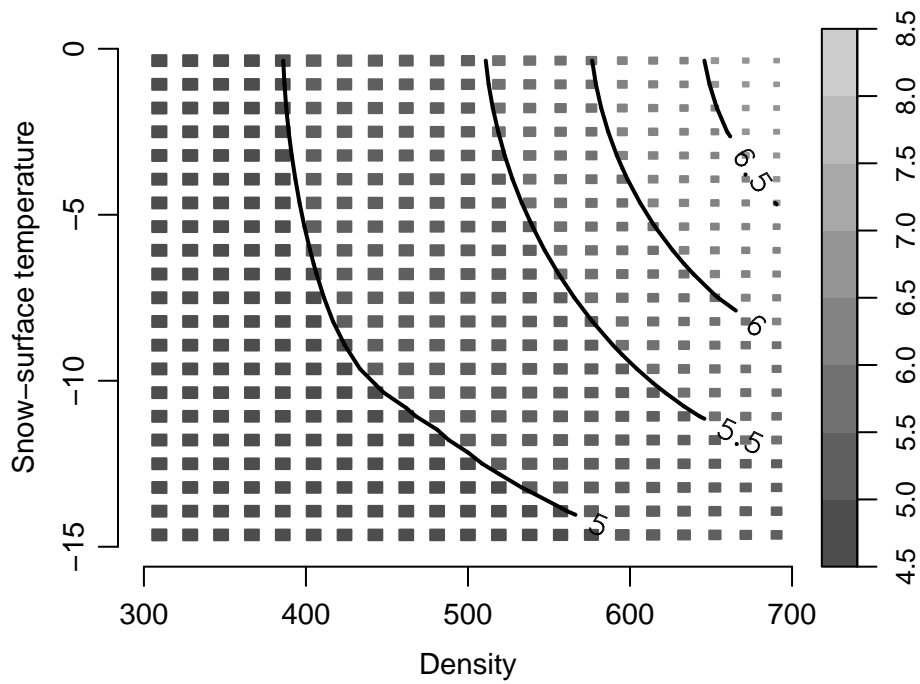


Figure 12: Prior predictive mean and standard deviation of velocity at a height of 0.4 m, by snow density and snow-surface temperature. The grey-scale and contours show the mean velocity. The width of each grey tile is inversely proportional to the standard deviation (i.e. the area is proportional to the precision), so that regions with more white indicate more uncertainty; the standard deviations run from 0 m/s (full width) to ≥ 4 m/s (no width, all white).

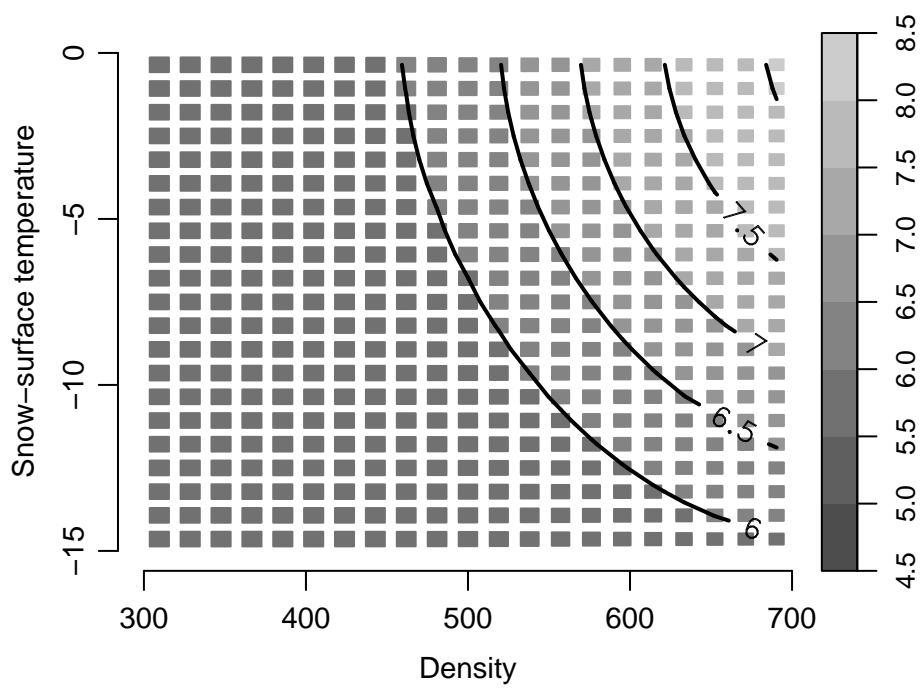


Figure 13: Adjusted predictive mean and standard deviation of velocity at a height of 0.40 m, *cf.* Figure 12.

is close to melting.

In general the value of the environmental variables can also affect our judgement about the ‘best’ value of the model parameters. This requires a transformation of the model parameter space; put simply, we have to construct ‘hyper-model-parameters’ that are uncertain and independent of the environmental variables. These hyperparameters are the quantities in which information from experiments under different values for the environmental variables is combined. For the HB model we required two sets of such hyperparameters: to account for the effect of density on basal slip velocity, and temperature on yield stress. In both cases they took the form of uncertain coefficients in quadratic equations. Hyperparameters are commonly used in hierarchical statistical modelling, but this is, to our knowledge, the first time they have been used in this way in a computer experiment with a physical model.

Throughout our analysis we have made judgements about the HB model and the statistical framework linking model parameters, model evaluations, actual snow behaviour, and observations. These types of judgements are never easy, and can be contentious. But one of the great advantages of a Bayesian analysis is that it forces us to quantify our judgements in a transparent manner, to facilitate discussion, and—in due course—revision. We have used a number of different methods for quantifying our judgements, including a novel but effective way of specifying simple uncertain functions. We have validated these judgements, collectively, in an assessment of prior predictive means and variances, and in a leave-one-out diagnostic analysis. They remain, however, *our* judgements, and we do not expect that they will meet with the approval of all statisticians and snow experts.

From a practical point of view, our results show how we can generalise the HB model to allow for environmental conditions, and thereby extend the prediction range of the model to environmental conditions that have not been observed. In this study we have focused on chute experiments, but our longterm goal is to construct better avalanche hazard maps, which are based on numerical modelling of avalanche velocity and run-out length. At the heart of these models are empirical rheology models similar to the HB model considered here. However, in current practice the discrepancy between these models and actual avalanche behaviour is never formally accounted for, and information on the snow cover properties (e.g. density and temperature) is underused; both of these create uncertainty in the resulting hazard maps. The approach described here allows us to quantify these uncertainties and should lead to a more time- and site-specific approach to avalanche hazard estimation, by considering the local temporal and spatial properties of the snow-cover in the avalanche release zones. Many more variables would have to be taken into account (topography properties, erosional properties, spreading in 3D avalanche flow). But we are convinced that this is a necessary and feasible task for future development of avalanche hazard

management tools.

Acknowledgements

These experiments were performed on the SLF snow chute at Weissfluhjoch, while the second-named author was a researcher at that Institute. We would like to thank Jürg Schweizer and Martin Schneebeli for guidance on the experimental data for τ_c and their validity for avalanche dynamics modelling, and Ana García Suárez and Cari Kaufman for suggestions about visualising the mean and standard deviation fields, and the eigenvectors. We would also like to thank the Associate Editor and two referees for their very perceptive and helpful comments on earlier drafts of this paper.

Appendix: Imposing constraints on random quadratic functions

Suppose we have a random function

$$y = \alpha_0 + \alpha_1 x + \alpha_2 x^2 \quad (\text{A1})$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$ is an uncertain vector subject to the following constraints:

- (1) $E(y; x = x_1) = y_1$, and $E(y; x = x_2) = y_2$, where $x_1 \neq x_2$;
- (2) y has an extremum at x_e with probability 1, where $x_e \neq (x_1 + x_2)/2$;
- (3) y is concave with probability p ;
- (4) α_0 and α_1 are independent;
- (5) $\boldsymbol{\alpha}$ is Multivariate Normal.

These conditions may or may not be consistent. Suppose that they are. Then the mean vector and variance matrix of $\boldsymbol{\alpha}$ are completely determined by one additional value, or free parameter, the standard deviation of α_0 .

First, note that (5) implies we need only consider the mean and variance of $\boldsymbol{\alpha}$. Then (2) implies that we must have $\alpha_1 + 2\alpha_2 x_e = 0$, which constrains both the mean and variance of (α_1, α_2) . Thus the three components of the mean vector must satisfy three linear constraints, given by (1) and (2):

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 0 & 1 & 2x_e \end{pmatrix} \begin{pmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix}. \quad (\text{A2})$$

The conditions $x_1 \neq x_2$ and $x_e \neq (x_1 + x_2)/2$ ensure that the square matrix is non-singular.

Consider the case $x_e \neq 0$. Then $\alpha_2 = -\alpha_1/2x_e$, so if we specify the variance matrix of $\{\alpha_0, \alpha_1\}$ then we induce a variance matrix for $\boldsymbol{\alpha}$. By (4), α_0 and α_1 are independent, and so we only need to specify the two standard deviations. One of these is the free parameter $\sigma_0 = \text{Sd}(\alpha_0)$. The other, σ_1 , controls the probability that y is concave, or, equivalently, the probability that α_2 is negative:

$$\Pr(\alpha_2 < 0) = \Pr(-\alpha_1/2x_e < 0) = \begin{cases} \Pr(\alpha_1 < 0) & x_e < 0, \\ \Pr(\alpha_1 > 0) & x_e > 0. \end{cases} \quad (\text{A3})$$

In the first case we have $\sigma_1 = -\mu_1/\Phi^{-1}(p)$, where $\Phi^{-1}(\cdot)$ is the quantile function for the Standard Normal, using (3), and (5) again, and in the second case $\sigma_1 = -\mu_1/\Phi^{-1}(1-p)$. The constraints are consistent if $\sigma_1 > 0$.

In the case where $x_e = 0$, we must have $\mu_1 = \sigma_1 = 0$. In this case $\Pr(\alpha_2 < 0) = \Phi(-\mu_2/\sigma_2)$, and $\sigma_2 = -\mu_2/\Phi^{-1}(p)$. The constraints are consistent if $\sigma_2 > 0$.

References

- Barnes, H. A., J. F. Hutton, and K. Walters (1989). *An Introduction to Rheology*. Amsterdam: Elsevier.
- Bastos, L. and A. O'Hagan (2009). Diagnostics for Gaussian Process emulators. *Technometrics* 51(4), 425–438.
- Bingham, E. C. (1922). *Fluidity and Plasticity*. New York: McGraw-Hill.
- Cowell, R., A. David, S. Lauritzen, and D. Spiegelhalter (1999). *Probabilistic Networks and Expert Systems*. New York: Springer.
- Cox, D. R. (1958). *Planning of Experiments*. New York: John Wiley & Sons, Inc.
- Craig, P., M. Goldstein, J. Rougier, and A. Seheult (2001). Bayesian forecasting for complex systems using computer simulators. *Journal of the American Statistical Association* 96, 717–729.
- Craig, P., M. Goldstein, A. Seheult, and J. Smith (1997). Pressure matching for hydrocarbon reservoirs: A case study in the use of Bayes Linear strategies for large computer experiments. In C. Gatsonis, J. Hodges, R. Kass, R. McCulloch, P. Rossi, and N. Singpurwalla (Eds.), *Case Studies in Bayesian Statistics III*, pp. 37–87. New York: Springer-Verlag. With discussion.
- Dent, J., K. Burrell, D. Schmidt, M. Louge, E. Adams, and T. Jazbutis (1997). Density, velocity and friction measurements in a dry snow avalanche. *Annals of Glaciology* 26, 247–252.
- Foehn, P., C. Camponovo, and G. Kruesi (1998). Mechanical and structural properties of weak snow layers measured in situ. *Annals of Glaciology* 26, 1–6.
- Goldstein, M. (1999). Bayes linear analysis. In S. Kotz (Ed.), *Encyclopaedia of Statistical Sciences, update vol. 3*, pp. 29–34. London: John Wiley & Sons.
- Goldstein, M. and J. Rougier (2004). Probabilistic formulations for transferring inferences from mathematical models to physical systems. *SIAM Journal on Scientific Computing* 26(2), 467–487.
- Goldstein, M. and J. Rougier (2006). Bayes linear calibrated prediction for complex systems. *Journal of the American Statistical Association* 101, 1132–1143.

- Goldstein, M. and J. Rougier (2009). Reified Bayesian modelling and inference for physical systems. *Journal of Statistical Planning and Inference* 139, 1221–1239. With discussion.
- Goldstein, M. and D. Wooff (2007). *Bayes Linear Statistics: Theory & Methods*. Chichester, England: John Wiley & Sons.
- Guenther, G. (2006, March). Optimierung optischer Geschwindigkeitssensoren für slushflows. Technical report, Swiss Federal Institute for Snow and Avalanche Research.
- Jaedicke, C., M. Kern, P. Gauer, M. Baillifard, and K. Platzler (2008). Chute experiments on slushflow dynamics. *Cold Regions Science and Technology* 51, 156–156.
- Kennedy, M. and A. O’Hagan (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society, Series B* 63, 425–450. With discussion, pp. 450–464.
- Kern, M., F. Tiefenbacher, and J. McElwaine (2004). The rheology of snow in large chute flows. *Cold Regions Science and Technology* 39, 181–192.
- McElwaine, J. and F. Tiefenbacher (2005). Calculating internal average velocities from correlation with error analysis. *Surveys in Geophysics* 24, 499–524.
- O’Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety* 91, 1290–1300.
- Oldroyd, J. G. (1947). A rational formulation of the equation of plastic flow for a Bingham solid. *Proceedings of the Cambridge Philosophical Society* 43, 100–105.
- Rasmussen, C. and C. Williams (2006). *Gaussian Processes for Machine Learning*. MIT Press. Available online at <http://www.GaussianProcess.org/gpml/>.
- Rougier, J. (2007). Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change* 81, 247–264.
- Rougier, J. (2008). Efficient emulators for multivariate deterministic functions. *Journal of Computational and Graphical Statistics* 17(4), 827–843.
- Schweizer, J. (1998). Laboratory experiments on the shear failure of snow. *Annals of Glaciology* 26, 97–102.
- Sovilla, B., M. Schaer, M. Kern, and P. Bartelt (2008). Impact pressures and flow regimes in dense snow avalanches observed at the Vallée de la Sionne test site. *Journal of Geophysical Research* 113, F01010.

Tiefenbacher, F. and M. Kern (2004). Experimental devices to determine snow avalanche basal friction and velocity profiles. *Cold Reg. Sci. Tech.* 38, 17–30.

West, M. and J. Harrison (1997). *Bayesian Forecasting and Dynamic Models* (2nd ed.). New York: Springer-Verlag.