XVIII Brazilian Conference on Soil Mechanics and Geotechnical Engineering
The Sustainable Future of Brazil goes through our Minas
COBRAMSEG 2016 – 19-22 October, Belo Horizonte, Minas Gerais, Brazil

Towards improved depth to bedrock modeling beneath hillslopes

Guilherme José Cunha Gomes
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil,
guilhermejcg@yahoo.com.br

Jasper Alexander Vrugt
University of California, Irvine, California, USA, jasper@uci.edu

Eurípedes do Amaral Vargas Jr.
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil,
vargas@puc-rio.br

SUMMARY: As hillslope interiors are very difficult and costly to illuminate and access, the topography of the bedrock surface is largely unknown. This paper is concerned with the prediction of spatial patterns in the depth to bedrock (DTB) using high-resolution topographic data, numerical modeling and Bayesian analysis. Our DTB model predicts a thick soil mantle at the ridge top and gets progressively thinner downslope. We reconcile the DTB model with field observations using Bayesian analysis. We illustrate our method using real-world data set of bedrock depth observations collected at the Papagaio river basin in Rio de Janeiro, Brazil. Our results demonstrate that the DTB model predicts accurately the observed bedrock depth data. We also derive the posterior prediction uncertainty of the DTB model.

KEYWORDS: depth to bedrock, Bayesian analysis, MCMC simulation, parameter estimation, field observations, prediction uncertainty.

1 INTRODUCTION

The depth to bedrock (DTB) controls a large array of geomorphologic, hydrologic, geochemical, ecologic and atmospheric processes, yet is large unknown as hillslope interiors are very difficult and costly to illuminate and access. The regolith thickness determines groundwater flow, infiltration and redistribution, subsurface saturation, runoff generation, storage capacity, the shape of the hydrograph, and variably saturated water flow. The bedrock topography is also of paramount importance in geotechnical engineering as it determines slope stability, pore pressure responses to infiltration, and landslide potential. An accurate characterization of the DTB is thus a prerequisite to describe adequately many different Earth-surface processes.

Several different modeling approaches have been proposed to estimate depth to bedrock. Examples of such attempts include interpolation methods from sparse DTB observations, empirical-statistical methods, and physically-based models. Interpolation methods can be classified in two main groups including (i) deterministic (e.g., triangulated irregular networks) and (ii) geostatistical (e.g., ordinary kriging) approaches. Empirical-statistical methods has led to the development of multivariate linear/nonlinear regression and other statistical methods (e.g., generalized additive models, artificial neural networks, support vector machines) that predict DTB from environmental covariates. Physically-based techniques are rooted on landscape evolution models that solve the soil mass-balance equation over geological time scales. A comprehensive review on DTB models is presented by Gomes et al. (2016a), and interested readers are referred to this publication for further details.

Whereas much progress has been made on the development and use of models for DTB prediction, surprisingly little attention has been given to inference of their parameters. Many of the parameters in these models cannot be measured directly in the field but can only be meaningfully inferred from field data. What is more, some parameters might be depth-dependent or vary spatially depending on hillslope position and lithology.

In this paper, we introduce a Bayesian framework for DTB model parameter estimation. We use Markov chain Monte Carlo (MCMC) simulation with the DiffeRential Evolution Adaptive Metropolis (DREAM) algorithm (Vrugt, 2016) to infer the parameters of the DTB model from spatially distributed regolith depth observations. We also investigate the benefits of using spatially distributed DTB parameter values for the prediction of bedrock depths. We validate our approach with real-world data collected at the Papagaio river basin (PRB) in Rio de Janeiro, Brazil.

2 MODEL DESCRIPTION

Our model builds on the bottom-up control on fresh-bedrock hypothesis of Rempe and Dietrich (2014) and calculates the thickness of the weathered zone from the difference between the measured surface topography and predicted groundwater profile derived from analytic solution of the one-dimensional steady-state Boussinesq equation. The regolith thickness, h [L] of a soil-mantled hillslope can be derived by calculating the difference between the elevation of the ground surface, Z_{s} [L] and the underlying topography, Z_{b} [L] of the fresh bedrock

$$h(x,y) = Z_{\mathrm{s}}(x,y) - Z_{\mathrm{b}}(x,y), \tag{1}$$

where the coordinates (x,y) are used to denote spatial location. Spatial maps of Z_{s} are readily available from digital elevation models (DEMs). The bedrock depth model of Rempe and Dietrich (2014) builds on the one-dimensional, steady-state form of the Boussinesq equation for groundwater flow (Bear, 2013)

$$\frac{1}{2}K\frac{\partial^2 Z_{\mathrm{b}}^2}{\partial x^2} + \phi C_o = 0, \tag{2}$$

where K [LT^{-1}] denotes the saturated hydraulic conductivity of the bedrock, x [L] is the horizontal distance from the ridge, ϕ [-] signifies the porosity, and C_o [LT^{-1}] represents the channel incision rate at the base of the hillslope.

By assuming strictly horizontal flow, topographic symmetry about the ridge, and a channel elevation at the bottom of the hillslope, the following closed-form equation can be derived for the elevation of the transition from fresh to weathered bedrock

$$Z_{\mathrm{b}}(x) = \sqrt{\frac{\phi C_o}{K}(L^2 - x^2)}, \tag{3}$$

where L [L] is the hillslope length, and the term $(L^2 - x^2)$ can be interpreted as distance to the drainage channel. A step-by-step derivation of Equation (3) is given in Rempe and Dietrich (2014). Equation (3) predicts that the depth of the weathered zone decreases from the hilltop to the valley floor with convexity and depth of the bedrock surface determined by the parameters ϕ, C_o, K.

Our DTB model uses as basic building block the analytic solution of Equation (3) but includes two important extensions that enhance applicability of the model to watersheds with convex and/or concave bedrock surfaces underneath the drainage valley and thin weathered zones and/or exposed rock on steep hillslopes subject to mass movement. This DTB model solves for the bedrock depth at two spatial coordinates, x and y and contains two new variables, Ψ and Λ whose values are derived from the slope angle and drainage distance, respectively, and three additional (quasi)-physical parameters. The basic formulation of the DTB model is given by the following closed-form equation

$$Z_{\mathrm{b}}(x,y) = \frac{\Psi}{\Lambda}\sqrt{\Phi L_{\mathrm{d}}^2(x,y)}, \tag{4}$$

where Ψ [-] measures the effect of mass movement on the bedrock surface, Λ [-] determines the shape and depth of the bedrock valley, and is hereafter also referred to as the bedrock-valley shape term, L_d [L] denotes the horizontal distance from the drainage, and $\Phi = \phi C_0/K$ [-] is a scalar that summarizes conveniently the combined effect of rock porosity, permeability, and the channel incision rate on the elevation of the bedrock surface, Z_b. The scalar variables Λ and Ψ are bounded between zero and one and determine the regolith thickness underneath valleys and steep slopes.

Stochastic mass movements are described as

$$\Psi = 1 - \min\left[1, (|\nabla Z_\mathrm{s}|/S_\mathrm{c})^2\right], \tag{5}$$

where ∇Z_s [-] denotes the slope gradient of the surface topography and S_c [-] signifies the critical slope angle beyond which mass movement is initiated. Equation (5) predicts regolith loss on hillslopes steeper than the threshold angle S_c. This movement of mass (due to landslides) gives rise to exposed rock.

The variable Λ in Equation (4) determines the hillslope-to-valley transition morphology and is computed as follows

$$\Lambda = \exp\left[-\lambda_1(1-\bar{L_\mathrm{d}})^{\lambda_2}\right], \tag{6}$$

where $\bar{L_\mathrm{d}}$ denotes the normalized drainage distance, and λ_1 and λ_2 are shape parameters that determine the bedrock shape (curvature) and depth in the valley at the base of the hillslope.

Figure 1 presents a sensitivity analysis of the parameters of the DTB model. The four different horizontal panels show the DTB model predicted regolith profiles underneath an artificial hillslope for different values of the parameters Φ (top), λ_1 (top-middle), λ_2 (bottom-middle), and S_c (bottom). The artificial topography (surface) is separately indicated in each plot with the black line.

In summary, Figure 1 shows that, regardless of the parameter values used in the DTB model, the weathered zone appears largest at the hilltop and then progressively thins downwards. This profile of the bedrock depth underneath the hillslope is in agreement with field observations of upland and lowland areas and mimics qualitatively the output of the Rempe and Dietrich (2014) model. The effect of the parameter S_c (regolith movement due to landslides) on the output of the DTB model is shown in Figure 1 (g) and (h) and reduces, as expected, the thickness of the weathered zone along the sideslope. The effect of the DTB variable Λ (hillslope-to-valley morphology) is visible in most of the displayed bedrock depth profiles with a shape and curvature of the bedrock surface in the valley (drainage) that deviates considerably from the concave drainage profiles simulated exclusively by Equation (3) of Rempe and Dietrich (2014).

3 INVERSE MODELING

The DBT model contains several coefficients that are difficult to be measured directly in the field and thus have to be determined by calibration instead using some spatially distributed regolith depth observations. If we denote with $\mathcal{F}$ Equation (4) then we can write our DTB model as follows

$$\mathbf{H} \leftarrow \mathcal{F}(\boldsymbol{\theta}, \nabla Z_\mathrm{s}, \mathbf{L}_\mathrm{d}) + \mathbf{e}, \tag{7}$$

where $\mathbf{H} = \{h_1, \ldots, h_n\}$ is a n-vector of simulated bedrock depths at spatial coordinates, $(x_1, y_1)\ldots(x_n, y_n)$, $\boldsymbol{\theta} = \{\Phi, \lambda_1, \lambda_2, S_\mathrm{c}\}$ signifies the d-vector of model parameters, $\mathbf{L}_\mathrm{d} = \{L_\mathrm{d}(x_1, y_1), \ldots, L_\mathrm{d}(x_n, y_n)\}$ stores the n-values of the drainage distance of each measurement location, and $\mathbf{e} = \{e_1, \ldots, e_n\}$ represents the vector of observation errors. The vector $\mathbf{e}$ includes observation error as well as error due to the fact that the DTB model, $\mathcal{F}(\cdot)$ may be systematically different from reality, $\Im(\boldsymbol{\theta})$ for the parameters $\boldsymbol{\theta}$. The latter may arise from an improper model formulation (epistemic errors) and topographic uncertainty (due to DEM measurement errors and/or inadequate resolution).

If we adopt a Bayesian formalism then we can derive the posterior distribution of the parameters, $p(\boldsymbol{\theta}|\tilde{\mathbf{H}})$, by conditioning the spatial behavior of the model on the n-measured values of the bedrock depth, $\tilde{\mathbf{H}} = \{\tilde{h}_1, \ldots, \tilde{h}_n\}$ using

$$p(\boldsymbol{\theta}|\tilde{\mathbf{H}}) = \frac{p(\boldsymbol{\theta})p(\tilde{\mathbf{H}}|\boldsymbol{\theta})}{p(\tilde{\mathbf{H}})}, \tag{8}$$

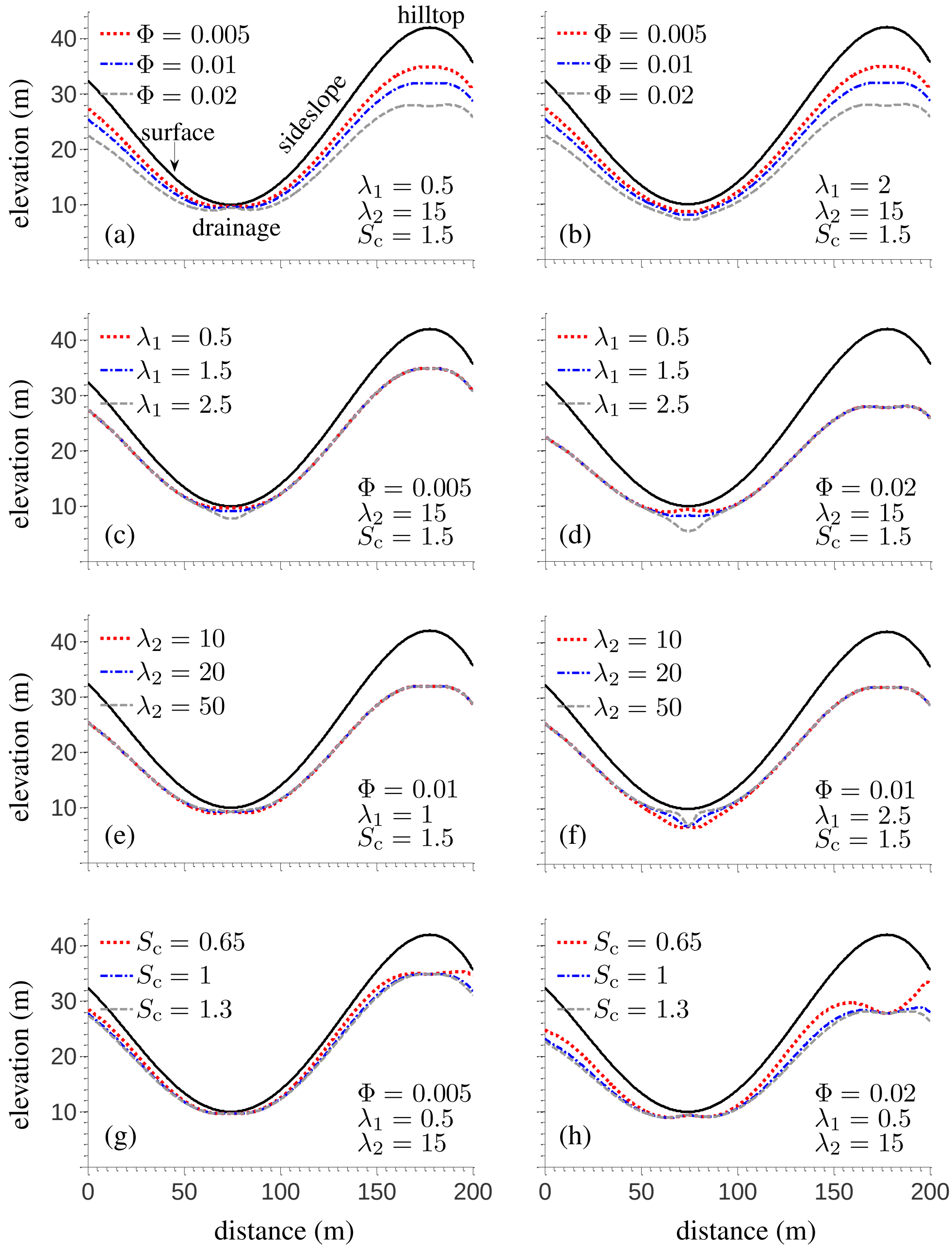

Figure 1: Sensitivity of the bedrock depth profile predicted by the DTB model to the values of the parameters Φ (a and b), λ_1 (c and d), λ_2 (e and f), and S_c (g and h). The dotted lines in red, blue, and gray display the simulated bedrock profiles.

where $p(\boldsymbol{\theta})$ is the prior parameter distribution, $L(\boldsymbol{\theta}|\tilde{\mathbf{H}}) \equiv p(\tilde{\mathbf{H}}|\boldsymbol{\theta})$ denotes the likelihood function, and $p(\tilde{\mathbf{H}})$ signifies the evidence. This latter variable is a constant that is independent of the parameter values and acts as a normalization constant (scalar) so that the posterior distribution integrates to unity over the parameter space, $\boldsymbol{\theta} \in \boldsymbol{\Theta} \in \mathbb{R}^d$. In practice, $p(\tilde{\mathbf{H}})$ is not required for posterior estimation as all statistical inferences about $p(\boldsymbol{\theta}|\tilde{\mathbf{H}})$ can be made from the unnormalized density

$$p(\boldsymbol{\theta}|\tilde{\mathbf{H}}) \propto p(\boldsymbol{\theta})L(\boldsymbol{\theta}|\tilde{\mathbf{H}}). \tag{9}$$

We conveniently assume that the prior distribution, $p(\boldsymbol{\theta})$ is uniform, $p(\boldsymbol{\theta}) \propto c$, where c is a constant. This means that we a-priori do not favor any values of the model parameters, and instead use uniform prior ranges. The main culprit now resides in the definition of the likelihood function, $L(\boldsymbol{\theta}|\tilde{\mathbf{H}})$, used to summarize the distance between the model simulations, $\mathbf{H}(\boldsymbol{\theta})$, and corresponding observations, $\tilde{\mathbf{H}}$. If we assume the error residuals of the observed and simulated bedrock depths to be normally distributed and uncorrelated, with homoscedastic measurement errors, then the likelihood function can be written as

$$L(\boldsymbol{\theta}|\tilde{\mathbf{H}}) \propto \sum_{i=1}^{n} |\tilde{h}_i - h_i(\boldsymbol{\theta})|^{-n}. \tag{10}$$

Once the prior distribution and likelihood function have been defined, what is left in Bayesian analysis is to summarize the posterior distribution. For models such as Equation (4) which is nonlinear in its parameters, the posterior distribution $p(\boldsymbol{\theta}|\tilde{\mathbf{H}})$ cannot be obtained by analytic means nor by analytic approximation. We therefore resort to iterative methods that approximate the posterior probability density function by generating a large sample from this distribution. The most powerful of such sampling methods is Markov chain Monte Carlo (MCMC) simulation using the Metropolis algorithm. The basis of MCMC simulation is a Markov chain that generates a random walk through the search space and successively visits solutions with stable frequencies stemming from a stationary distribution. To explore the target distribution a MCMC algorithm alternates between three basic steps. First, a proposal $\boldsymbol{\theta}_{\mathrm{p}}$ is generated from the current state of the Markov chain, $\boldsymbol{\theta}_t$. Next, this proposal is accepted with Metropolis probability

$$P_{\mathrm{acc}}(\boldsymbol{\theta}_t \rightarrow \boldsymbol{\theta}_{\mathrm{p}}) = \min\left[1, \frac{p(\boldsymbol{\theta}_{\mathrm{p}})}{p(\boldsymbol{\theta}_t)}\right]. \tag{11}$$

Finally, if the proposal is accepted, the chain moves to $\boldsymbol{\theta}_{\mathrm{p}}$, and thus $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_{\mathrm{p}}$, otherwise the current position is retained, $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t$. Repeated application of these three steps results in a Markov chain which, under certain regularity conditions, has a unique stationary distribution with posterior probability density function. This selection rule has become the basic building block of the random walk Metropolis (RWM) algorithm, the earliest MCMC method.

In this paper, MCMC simulation of the DTB model has been performed using the DREAM algorithm (Vrugt, 2016). This multi-chain MCMC simulation algorithm automatically tunes the scale and orientation of the proposal distribution en route to the target distribution. We evaluate the DTB model using lumped and spatially distributed parameter values. These values are stored in the d-vector $\boldsymbol{\theta}$.

4 BAYESIAN INFERENCE

Figure 2 summarizes the setup of the DTB model for invariant (lumped) and variant (distributed) parametrization. The top panel displays the surface of an idealized DEM consisting of P cells (pixels). To simplify notation we use a single variable, $i = \{1, \ldots, P\}$ to denote the xy coordinates of a DEM cell. The input data of the DTB model in Equation (4) differs per grid cell and is stored in the vector, $\mathbf{U}_i = \{\nabla Z_{\mathrm{b}i}, L_{\mathrm{d}i}\}$. If a lumped parameterization of the DTB model is used then it suffices to use the same parameter values, $\boldsymbol{\theta} = \{\Phi, \lambda_1, \lambda_2, S_{\mathrm{c}}\}$ for each cell of the xy grid. A distributed DTB parameterization uses different parameter values for each region of the DEM. The spatially distributed framework requires the user to define a spatial pattern for each of the model parameters. For example, in Figure 2 we assume a simple block pattern of $r = 16$ equal-sized squares for each of the parameters. After the parameterization of the DTB model has been defined, the DREAM algorithm proceeds with statistical inference of the model parameters using distributed observations of the bedrock depth. Both implementations use the same source code of the DTB model but differ in their assignment of the parameter values.

The method described is now applied to the borehole observations at the experimental site detailed in Gomes et al. (2016b). The bedrock data set is split randomly into two parts, designated for DTB-model calibration (75%) and evaluation (25%). We now estimate the posterior distribution of the DTB model parameters, $\boldsymbol{\theta} = \{\Phi, \lambda_1, \lambda_2, S_{\mathrm{c}}\}$ using Bayesian inference

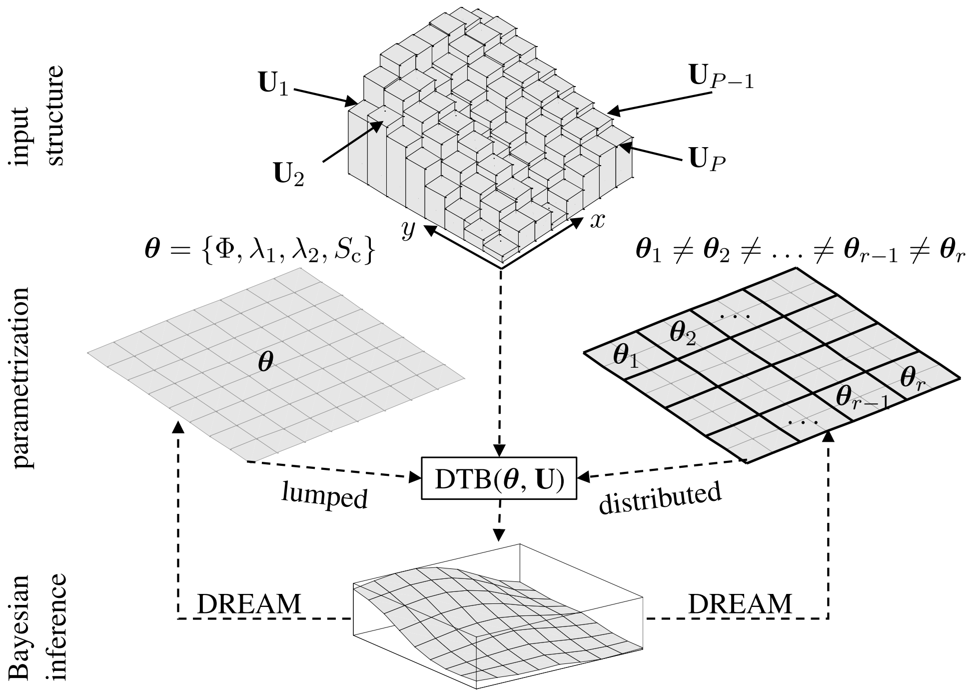

Figure 2: Schematic overview of the DTB modeling framework for a lumped (left) and distributed (right) parameterization of the watershed.

with DREAM. The prior ranges for the parameters are listed in Table 1.

Table 1: Prior uncertainty ranges of the DTB parameters for the observed bedrock depth data.

Range	Φ [-]	λ_1 [-]	λ_2 [-]	S_c [-]
minimum	10^{-4}	0.1	1.0	0.8
maximum	10^{-1}	3.0	20.0	1.5

Figure 3 presents a scatter plot matrix of the posterior samples derived from DREAM. The graphs on the main diagonal present marginal distributions of each of the parameters, whereas the off-diagonal elements display bivariate scatter plots of the posterior samples. The posterior distribution of the parameter Φ follows closely a normal distribution with median posterior solution that is in excellent agreement with the MAP value (cross symbol in (a)). The marginal distribution extends only a small portion of the uniform prior distribution (see Table 1), which demonstrates that this parameter is well defined by calibration against the real-world bedrock depth data. The marginal distributions of the other three parameters occupy almost their entire prior distribution, which suggest that these parameters are poorly defined by calibration against our bedrock depth observations. The relatively low values of the parameter λ_1 (including the MAP value) suggest that the bedrock is close to the surface in the channel zone with a thin soil mantle overlying a weathered bedrock zone (Figure 3 (f)). The high MAP value for λ_2 (Figure 3 (k)) signifies that the bedrock valley topography approximates a smooth concave shape (see Figure 1). The marginal distribution of parameter S_c attains high values, demonstrating the presence of a much thicker regolith zone underneath steep slopes. The bivariate scatter plots (off-diagonal) highlight the presence of some negligible correlation between the DTB parameters, Φ and λ_2 and Φ and S_c.

The posterior mean parameter values derived from the calibration are now used to determine the performance of the DTB model on the independent evaluation data set for different parametrizations. The performance of the DTB model is summarized in Table 2. The listed value of the RMSE of 1.80 m and the ρ-statistic of 0.83 (lumped case) can be considered acceptable for

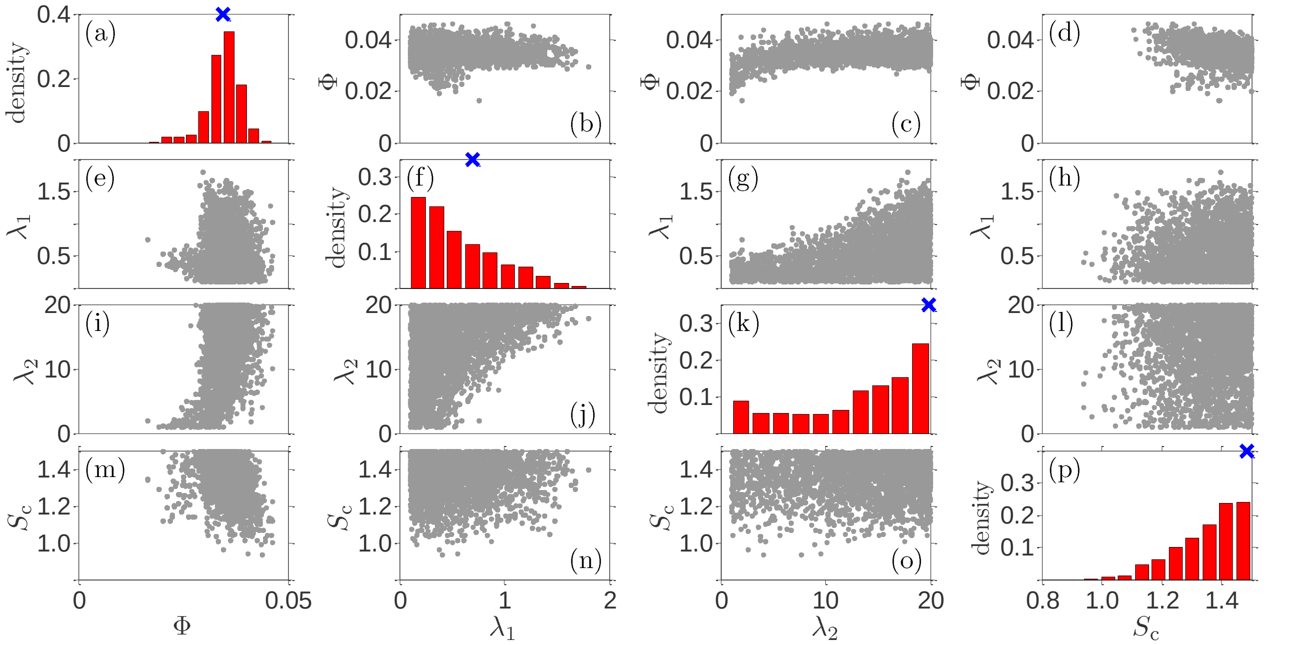

Figure 3: Scatter-plot matrix of the posterior samples generated with the DREAM algorithm. The main diagonal plots histograms of the marginal posterior distribution of the DTB model parameters, Φ, λ_1, λ_2 and S_c respectively, and the off-diagonal graphs present scatter plots of the posterior samples of the different parameter pairs.

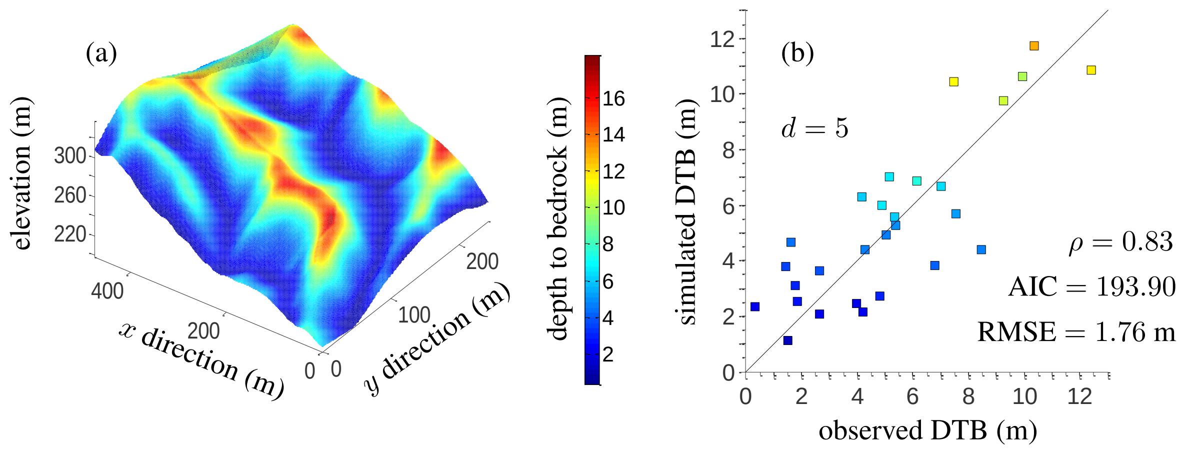

Figure 4: (a) Spatial distribution of the regolith thickness predicted at the PRB experimental watershed using a distributed ($d = 5$) parameterization of Φ. (b) The bivariate scatter plot compares the observed and simulated bedrock depths of the evaluation data set.

the PRB experimental watershed.

Table 2: Performance statistics of the calibrated DTB model after Bayesian inversion with DREAM.

Case	d	RMSE	ρ	AIC
Lumped	4	1.80 m	0.83	194.00
Distributed	5	1.76 m	0.83	191.10
Distributed	8	1.76 m	0.82	193.90

We now plot in Figure 4 (a) the DTB simulated regolith thicknesses using the mean posterior solution of the distributed ($d = 5$) parameterization (Gomes et al., 2016a). The scatter plot (b) compare the observed and simulated regolith depth values at the different measurement locations. The color coding in the regression plot matches the color bar used in (a). The DTB model predicts a smooth topography from the hilltop (thick regolith) to the drainage channel (thin or even exposed rock). These results are in agreement with theory (Rempe and Dietrich, 2014) and our field expertise, and provides support for the claim that the DTB model gives an adequate description of the bedrock surface at the PRB field site.

We conclude the paper with Figure 5, which plots the DTB simulated bedrock profile of the mean posterior solution for the **BB**' transect (see location in Gomes et al. (2016b)) using the distributed ($d = 5$) parameterization. The topographic surface is indicated with the black line and the observed bedrock depth data are indicated separately with a blue dot. The dark gray region represents the 95% confidence intervals of the output prediction due to parameter uncertainty, whereas the light gray region denotes the corresponding total prediction uncertainty. The simulated posterior mean bedrock surface (solid red line) appears rather smooth and fits nicely the observed bedrock depth observations. The 95% parameter uncertainty bounds appear relatively small and track closely the observed regolith depth data. The total (model + parameter) 95% prediction uncertainty intervals are rather large and encompass the observations.

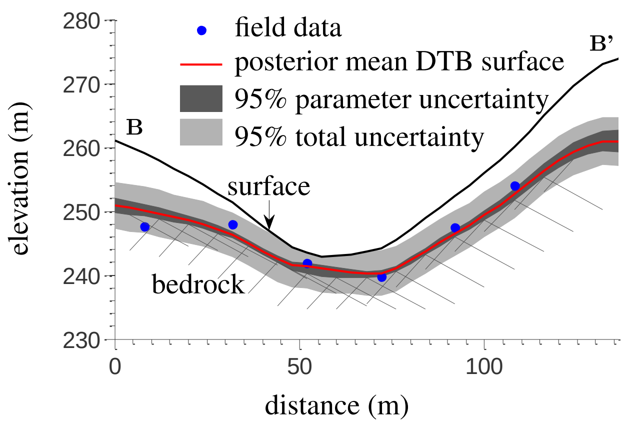

Figure 5: Transect **BB**' of Gomes et al. (2016b) summarizing the posterior mean (red line) and prediction uncertainty bounds (gray regions) of the DTB model.

5 CONCLUSIONS

We have introduced a model to predict depth to bedrock in hillslopes from a high-resolution topography. Our model builds on the bottom-up control on fresh-bedrock hypothesis of Rempe and Dietrich (2014) and calculates the depth to bedrock from the difference between the measured surface topography and predicted groundwater profile derived from analytic solution of the one-dimensional steady-state Boussinesq equation. Two additional terms are used in our DTB model to characterize adequately the effect of mass movement on steep hillslopes, and the shape and depth of the bedrock surface in the drainage valley.

Bayesian analysis with the DREAM algorithm was used to reconcile the DTB-model simulations with field observations. Our results demonstrate that the proposed DTB model with lumped parameters mimics reasonably well the observed regolith depth data with root mean square error (RMSE) of the posterior mean simulation of 1.80 m for evaluation data set, respectively. The performance of the DTB model can be enhanced if a distributed parameterization of Φ is used with RMSE reduced to 1.76 m. The use of a distributed parameterization provides a means to account implicitly for geologic/geomorphic heterogeneities that are difficult, or impossible, to characterize adequately in the field.

REFERENCES

Bear, J. (2013). *Dynamics of Fluids in Porous Media*, Dover publications, New York, NY, USA, 784 p.

Gomes, G.J.C.; Vrugt, J.A. and Vargas Jr., E.A. (2016a). Toward Improved Prediction of the Bedrock Depth Underneath Hillslopes: Bayesian Inference of the Bottom-up Control Hypothesis using High-Resolution Topographic Data, *Water Resources Research*, Vol. 52.

Gomes, G.J.C.; Araújo, J.P.; Vargas Jr., E.A.; Fernandes, N.F. (2016b). In-situ regolith depth measurements by DPL tests, *XVIII Braz. Conf. on Soil Mech. and Geotech. Eng.*, COBRAMSEG, Belo Horizonte, Vol. 1.

Rempe, D.M. and Dietrich, W.E. (2014). A Bottom-up Control on Fresh-bedrock Topography under Landscapes, *Proceedings of the National Academy of Sciences*, Vol. 111, p. 6576–6581.

Vrugt, J.A. (2016). Markov Chain Monte Carlo Simulation Using the DREAM Software Package: Theory, Concepts, and MATLAB Implementation, *Environmental Modelling & Software*, Vol. 75, p. 273–316.