STOCHASTIC METHODS IN OPERATIONAL AVALANCHE SIMULATION - FROM BACK CALCULATION TO PREDICTION

Valentin Hellweger[b], Jan-Thomas Fischer[a], Andreas Kofler[a], Andreas Huber[a], Wolfgang Fellin[b], Michael Oberguggenberger[b]

[a]*Austrian Research Centre for Forests (BFW), Innsbruck, Austria*
[b]*University of Innsbuck, Austria*

Avalanche simulations are an integral part of hazard assessment. Determining the potential hazard requires a multidisciplinary approach, including different scientific fields such as geography, meteorology, physics, civil engineering and mathematics. The application of probabilistic methods allows one to develop a complete, comprehensive applicational concept for snow avalanche simulations, ranging from back calculation to prediction. In this context optimal parameter sets or runout distances are represented by probability distributions.

Existing deterministic avalanche dynamics models contain several parameters (e.g. friction), some of them more conceptual than physical. Direct measurement of these parameters in the field is hardly possible. Hence, a parameter identification has to be undertaken, matching simulation results to field observations. This inverse problem can be solved by optimization or by a Bayesian approach (Markov chain Monte Carlo). An important task in snow avalanche simulation is to predict process intensities (runout, flow velocity and depth, ...). The identification process yields parameter distributions, that can be utilized for probabilistic reconstruction and prediction. Arising challenges include: the limited amount of observations, correlations appearing in model parameters or observed avalanche characteristics (e.g. velocity and runout) and the effective and accurate handling of ensemble simulations, always taking into account the related uncertainties.

Key words: avalanche dynamics, parameter estimation, stochastic simulation, Bayes' theorem, Metropolis-Hastings algorithm, posterior distribution, back calculation, prediction

1. INTRODUCTION

Over the last decade probabilistic methods have been successfully applied in snow avalanche simulation for simple flow models, compare e.g. Eckert et al. (2010) or Eckert et al. (2007). Here we present an approach for operational avalanche simulations, performed with a 2-dimensional operational simulation software (SamosAT Sampl and Granig, 2009; Zwinger et al., 2003) combined with a simulation results-post-processing technique (Fischer, 2013; Fischer et al., 2015b) in order to compare and rate simulation results with respect to one observed extreme avalanche event.

The first goal of this work is to back calculate the distribution of the model parameters fitting a specific avalanche. The parameter distributions $\pi(\theta)$ are derived, such that an avalanche simulation tool approximates measured field data $y_{\text{obs}} = f(x, \theta)$ for given input data x and a simulation model f with flow model parameters θ. To do so, we apply a stochastic approach explicitly treating the arising uncertainties in the measurements as well as in the simulation model. We employ Bayes' theorem and the theory of Markov chains to derive the previously unknown parameter distributions $\pi(\theta)$ through back calculation of a known event.

The second goal of this work is to perform a prediction of characteristics y_{pred} for an avalanche on a given path considering the parameter distributions $\pi(\theta)$ given through the back calculation. Here Monte Carlo simulations allow one to estimate the variability of the simulation results y_{pred}, explicitly taking into account the obtained parameter distribution.

2. Mathematical model

Mathematically we can realize the parameter estimation by solving an inverse problem. To model certain kinds of uncertainties all variables are considered as random variables. With this step the inverse problem turns into a so called statistical inverse problem. The objective of statistical inversion theory is to extract information and assess the uncertainty about the variables based on all available knowledge of the measurement process as well as information and models of the unknown parameters that are available prior to the measurement. The statistical inversion approach is based on the following principles:

1. all variables included in the model f are modeled as random variables,

2. the randomness describes our degree of information concerning their realizations,

3. the degree of information concerning these values is encoded in the probability distributions π,

4. the posterior probability distribution is the solution of the inverse problem.

Point number 4 is the crucial difference between the statistical approach and classical regularization methods such as maximum likelihood or residual squares. (see e.g. Kofler et al. (2016); Fischer et al. (2015b) for application in avalanche simulations).
We denote random variables by capital letters and their realizations by lower case letters. So let Θ, Y be random variables. In Bayesian theory it is assumed that we have some information about the model parameters Θ. Thus we can encode this information into a probability density $\theta \mapsto \pi_{\text{pr}}(\theta)$, called the *prior density*. The conditional probability density of Y, given a value of the unknown $\Theta = \theta$, is

$$\pi(y|\theta) = \frac{\pi(\theta, y)}{\pi_{\text{pr}}(\theta)}, \text{ if } \pi_{\text{pr}}(\theta) \neq 0. \tag{1}$$

This term is called the *likelihood function*, because it expresses how likely different measurement outcomes y_{obs} with given $\Theta = \theta$ are. Given measured data $Y = y_{\text{obs}}$, the conditional probability distribution

$$\pi(\theta|y_{\text{obs}}) = \frac{\pi(\theta, y_{\text{obs}})}{\pi(y_{\text{obs}})}, \text{ if } \pi(y_{\text{obs}}) \neq 0 \tag{2}$$

is called the *posterior distribution* regarding the input. It gives the probability for the unknown parameters θ, assuming given measured data y_{obs}.
Bayes theorem summarizes the previous notations and reads as

$$\pi_{\text{post}}(\theta) = \pi(\theta|y_{\text{obs}}) = \frac{\pi_{\text{pr}}(\theta)\pi(y_{\text{obs}}|\theta)}{\pi(y_{\text{obs}})}. \tag{3}$$

In summary we can say that solving an inverse problem can be divided into three subtasks:

1. finding a probability density π_{pr} that represents the prior information Θ

2. finding the likelihood function $\pi(y_{\text{obs}}|\theta)$ that describes the interrelation between the observation and the unknown parameters θ

3. developing methods to explore the posterior probability density.

2.1. Back calculation

To explore the posterior distribution we use Markov chain Monte Carlo (MCMC) methods, i.e. we utilize the Metropolis-Hastings algorithm to generate a chain of parameter combinations, which follows the posterior distribution. It is shown that under certain conditions a Markov chain converges to a unique stationary distribution. For more details we refer to Brooks et al. (2011), Nummelin (1984) and Kaipio and Somersalo (2005).

2.1.1. The algorithm

Metropolis et al. (1953) published a scheme to generate a Markov chain having a desired stationary distribution and Hastings generalized it in 1970. Simulations following his scheme are said to use the *Metropolis-Hastings* (M-H) algorithm.
The idea of the algorithm is as follows: One chooses a starting point in the state space. Then for a certain number of iterations (length of the chain) in every iteration one draws a candidate following the so called proposal distribution and computes the ratio r of the likelihood function times the prior of the current state of the chain to the likelihood function times the prior of the candidate. Then one adds the candidate to the chain if $r > 1$. If the ratio is smaller than 1 one adds the candidate with probability r and the actual state of the chain with probability $1 - r$. To add candidates with a smaller likelihood function compared to the actual state of the chain with a certain probability ensures that the chain does not get caught in a part of the state space with a high density.

2.1.2. Monitoring

A common method to assess the convergence behaviour of the M-H algorithm is to start various chains at different starting points. As mentioned earlier we use a 2-dimensional operational avalanche simulation software, which is computationally very expensive and we cannot apply most of the MCMC diagnostics proposed in the literature. Generating a chain containing 3000 elements (performing 3000 simulations runs) needs about one week. Therefore, we decided to do "graphical diagnostics" and to check three different characteristics of the chain suggesting a high convergence speed.
These properties are:

- The acceptance rate: it is the fraction of proposed moves which are accepted. An acceptance rate of about 25% is generally considered appropriate for the multivariate case.

- The mixing: a chain is said to have good mixing if it moves fast from one part of the state space into an other. How fast the chain moves can be read off the trace plots.

- The autocorrelation: it gives the correlation of the elements in the chain. A fast decreasing autocorrelation function indicates independent elements and a higher efficiency in exploring the target distribution.

Note that good mixing and small autocorrelations do not indicate that the chain has already reached its equilibrium distribution.

2.2. Prediction

The next step is to use the posterior distribution $\pi_{\text{post}}(\theta)$ to perform predictive simulations. The simulation results give probabilities for certain avalanche characteristics (runout, velocities) and explicit information about their variabilty and are directly linked to the back calculated posterior distribution. To realize this we use a Monte Carlo approach: we generate a sample of the model parameters following the posterior distribution, perform simulations with these parameter combinations and evaluate the simulation results statistically.
Thus we should generate a sample following the posterior distribution, i.e. it should be sampled from the same distribution as obtained from the Markov chain. One common method to realize this is to apply the inversion method componentwise. However, we cannot use this method in a straightforward way because the components of the chain are correlated. But it is possible to approximate the joint distribution by means of a copula, preserving the marginals and the linear correlation coefficients (estimated through the covariance matrix C_θ) of the original multivariate distribution. We shall apply this method of generating multivariate samples, using a Gaussian copula. For more details we refer to Nelson (2006).

3. Application in operational avalanche simulation

In the following section we will present the application of the theory discussed above and show how to derive the posterior distribution of θ and how this distribution can be used for avalanche prediction.

3.1. Model

To perform the avalanche simulations we utilize the snow avalanche simulation software SamosAT (**S**now **A**valanche **MO**delling and **S**imulation - **A**dvanced **T**echnology, Zwinger et al., 2003; Sampl and Zwinger,2004). The simulation results are transfered in a path dependent coordinate system (Fischer, 2013) in order to be able to compare simulation results from paths with different topographies.

The model parameters
As bottom friction and entrainment relation we use the well known Voellmy friction relation for the basal shear stress $\tau^{(b)}$, which combines a Coulomb bottom friction with a velocity dependent drag term

$$\tau^{(b)} = \sigma^{(b)}\mu + \frac{g}{\xi}\bar{\mathbf{u}}^2,$$

with the dimensionless Coulomb friction parameter μ and the turbulent friction coefficient ξ, combined with a simple assumption for the entrainment rate

$$\dot{q} = \frac{\tau^{(b)}}{e_b}\|\bar{\mathbf{u}}\|, \tag{4}$$

that includes the erosion energy parameter e_b, (Fischer et al., 2015a). The choice of the friction and entrainment model defines a set of flow model parameters $\theta = \{\mu, \xi, e_b\}$. The appropriate parameter ranges of each model parameter are constrained by physically relevant ranges, results of scaling analysis, experimental work or prior model optimization through back calculations. We choose the intervals of the single components to be $\mu \in [0.1, 0.5]$, $\xi \in [0, 25000]$ and $e_b \in [1000, 30000]$.

Simulation results
The input for each simulation is the documented release volume and the path topography. Additionally an initial snow distribution (Mountain Snow Cover - MSC) is taken into account, which is determined with respect to the topography of the path and extreme snow heights (Leichtfried, 2010; Fischer et al., 2015b). After each simulation run (which computes the spatio temporal evolution of flow depths and velocities), a post processing step uses the peak pressure values to obtain the following main simulation result variables:

- runout r
- velocity $u_{\max}$
- normed true positive tp
- normed true negative tn
- deposit volume V_{dep}

3.1.1. Documented data

One avalanche path, namely Alpenlahner, is given with a documentation which can be considered extreme and can serve as a design event in terms of hazard zoning. Observations of the avalanche event include a release volume of $218517\,\text{m}^3$, more observational variables are summarized in table 1.

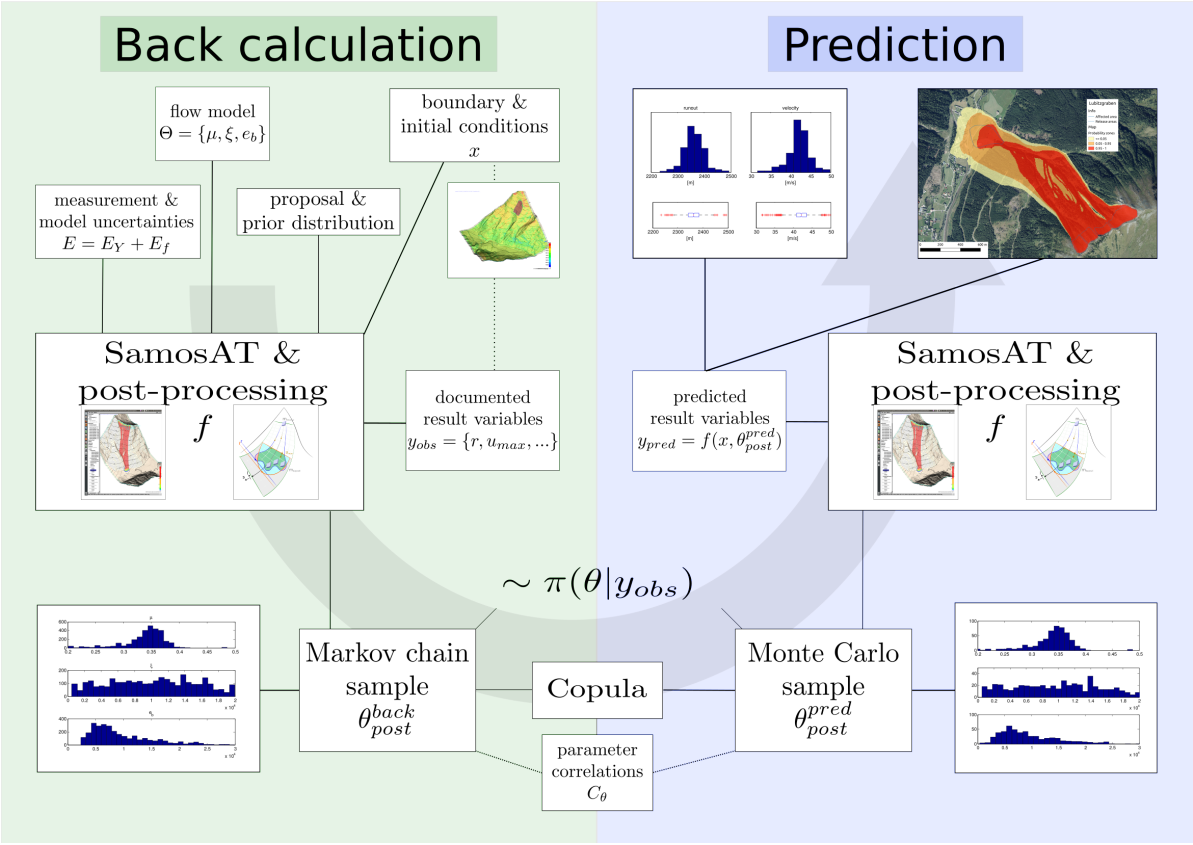

Figure 1: Flowchart of the back calculation and prediction concept

r [m]	$u_{\max}$ [m/s]	tp [m^2]	tn [m^2]	V_{dep} [m^3]
2384	61.7	85089	377687	388225

Table 1: Documented data for the path Alpenlahner

3.2. Back calculation

In this subsection we specify the components we need in order to apply the Metropolis-Hastings algorithm such as likelihood function, prior and proposal distribution.
As mentioned earlier the goal is to solve the statistical inverse problem

$$f(X, \Theta, E) = Y, \tag{5}$$

for the unknown parameter vector Θ. In the following we consider a statistical inverse problem with additive noise subsumed in the random variable E with expected value zero; the probability density π_{noise} of the noise is assumed to be known (see below). Here X corresponds to the input data such as release volume and topography and Y to the documented data. The error term E should model the arising uncertainties

- in the measured data (runout, velocity, affected area and deposit volume)
- in the model and its numerical implementation respectively.

Mathematically these assumptions lead to

$$f(X, \Theta, E) = Y \iff f(X, \Theta) + E_f = Y + E_Y. \tag{6}$$

where E_f denotes the error that we get from the implementation of the model and E_Y denotes the noise that we get from the uncertainty in the measured data in table 1.
The function f describes the simulation procedure and the postprocessing of the result. Note that we do not take into account any error resulting from uncertainties in the input data X.
These assumption on the error and Bayes' theorem lead to

$$\pi(\theta|y_{\text{obs}}) \propto \pi_{\text{pr}}(\theta) \cdot \pi_{\text{noise}}(f(x, \theta) - y_{\text{obs}}). \tag{7}$$

The notation indicates that the posterior distribution of the unknown parameter equals the prior distribution times the likelihood function up to a normalizing constant.

To apply algorithm 2.1.1 one has to define a proposal distribution, a prior distribution and an appropriate likelihood function.

The proposal distribution
We decided to take a Gaussian proposal distribution, where the mean should be the current state of the Markov chain. Trial and error suggests the covariance matrix $C_p = \mathrm{diag}(0.02^2, 8000^2, 8000^2)$ in order to let the Markov chain have the desired properties discussed in section 2.1.2.

The prior distribution
A priori we do not know anything about the model parameters apart from physically relevant intervals, see Fischer et al. (2015a). This information can be expressed by a probability distribution, namely in a continuous uniform distribution which formally reads as

$$\pi_{\mathrm{pr}}(\theta) = \chi_{[0.1,0.5]}(\theta_1) \times \chi_{[0,25000]}(\theta_2) \times \chi_{[0,30000]}(\theta_3) \tag{8}$$

for $\theta_1 = \mu, \theta_2 = \xi$ and $\theta_3 = e_b$.

The likelihood function
As commonly accepted we decided to model the error E as Gaussian random variable with zero mean. Then the likelihood function reads

$$\pi(y|\theta) = \pi_{\mathrm{noise}}(f(x,\theta) - y) = \\ \mathrm{const} \cdot \exp\left(-\frac{1}{2}(f(x,\theta) - y)^{\intercal}\Sigma_E^{-1}(f(x,\theta) - y)\right), \tag{9}$$

with $\mathrm{const} = 1/\sqrt{(2\pi)^5 \cdot \det(\Sigma_E)}$.

The likelihood covariance matrix
Again we assume E_f and E_Y to be Gaussian.
Further we can use the fact that the sum of two independent Gaussian random vectors with covariance matrices Σ_{E_Y} and Σ_{E_f} is again Gaussian with covariance matrix $\Sigma_E = \Sigma_{E_Y} + \Sigma_{E_f}$.

3.2.1. Generation of the Markov chain $\theta_{\mathrm{post}}^{\mathrm{back}}$

In this section we present the results of the Metropolis-Hastings algorithm. Due to the fact that we cannot exactly determine when the chain has reached its equilibrium distribution we decide to search for an appropriate proposal distribution and stop the algorithm after 3000 iterations, such that the monitoring properties are sufficiently met.
We assume the resulting Markov chain to follow the posterior distribution.
For a presentation of the results we use the path Alpenlahner. Trace plot and autocorrelation function of the generated Markov chain indicate good mixing and the acceptance rate equals 0.42.

3.3. Prediction

To produce a sample $\theta_{\mathrm{post}}^{\mathrm{pred}}$ from the posterior distribution we use its marginals and a Gaussian copula, as described in subsection 2.2. To get an upper bound for the sample size in order to get meaningful statistical results we use Bikelis' theorem, see Graham and Talay (2013). It turns out that 500 simulations are enough to estimate the expectation value of the runout up to $3\,\mathrm{m}$ with a confidence level of 95%.
To check if the sample follows the posterior distribution we consider the marginals and the linear correlations between the components. A componentwise two-side Kolmogorov-Smirnov test does not reject the hypothesis that the components of the Markov chain and the components of the sample are sampled from the same distributions for μ, ξ and e_b. Also the linear correlation matrices are very close.

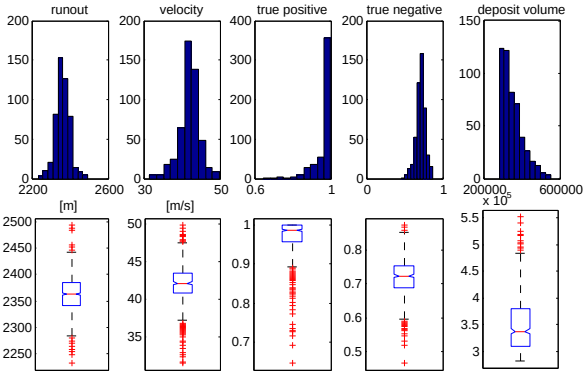

Figure 2: Histograms and boxplots of the simulation results generated with parameter combinations from a sample $\theta_{\mathrm{post}}^{\mathrm{pred}} \sim \pi(\theta|y)$.

Figure 2 shows the histograms and the boxplots of the simulation results generated with the above mentioned sample $\theta_{\mathrm{post}}^{\mathrm{pred}}$ for the path Alpenlahner.
From these one can read off that 95% of the simulated avalanche area covers 90% of the documented affected area and just 5% cover more than 30% of the non-affected area. The predicted mean values of the runout and deposit volume equal $2362\,\mathrm{m}$ and $351553\,\mathrm{m}^3$ respectively and are close to the documented data in table 1. We should mention that the predicted maximal velocities are much smaller than the documented one: the 95% confidence interval is $[36.7\,\mathrm{m/s}, 46.7\,\mathrm{m/s}]$ and the maximum equals $49.8\,\mathrm{m/s}$.

3.3.1. Probability map

Another approach to evaluating the avalanche simulation results generated with a sample following the posterior distribution, is to construct a "probability map". For a certain path this is a map which gives the probability that a certain region gets struck by an avalanche given the observed data.

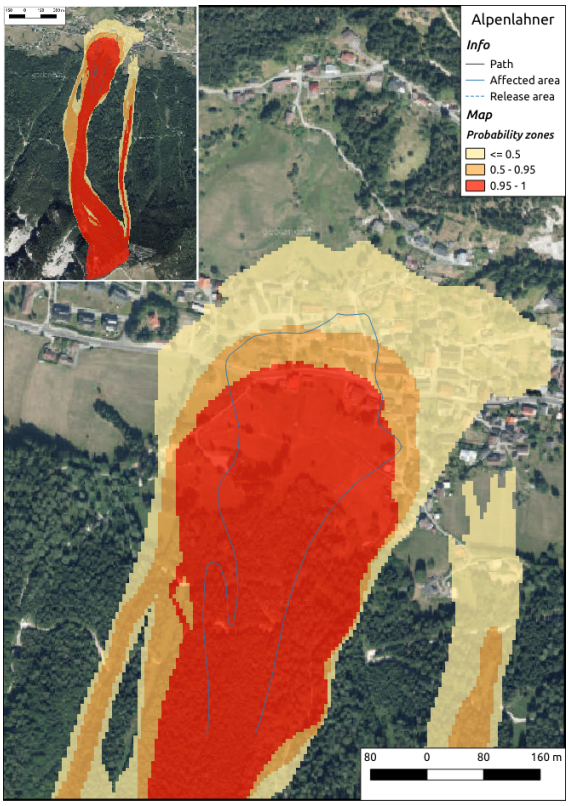

Figure 3: Zoom in on the deposition area for the path Alpenlahner. The different colors show the probability zones and the documented affected area is outlined blue. The picture on the top left illustrates the map for the whole path domain.

We perform avalanche simulations with the parameter sample $\theta^{\text{pred}}_{\text{post}}$ and count how often each square of a rectangular grid overlying the domain gets struck by the a simulation. Getting struck by the avalanche means that the pressure of the avalanche averaged over the square is greater then 3 kPa. Then a probability can be assigned to each square estimated by the relative frequency.

Mathematically one can interpret the probability map as follows: We take $\Omega = [\text{hit}, \text{no hit}]^n$ as underlying discrete probability space, where n is the number of squares of the map. The probability measure P has the hitting probability of the j-th squares as its j-th marginal. Then we can define a random variable (I for impact)

$$I: \Omega \to \{0,1\}^n,$$

$$I_j(w) = \begin{cases} 1, & \text{the square } j \text{ gets struck} \\ 0 & \text{the square } j \text{ does not get struck} \end{cases}$$

and use the results $I_j^{(i)}, i = 1, ..., N$ of N Monte Carlo simulations to approximate the expectation value of the Bernoulli random variable I_j by the relative frequency.
The probability that an arbitrary area j gets struck by the avalanche can be approximated by

$$P(I_j = 1 | y_{\text{obs}}) \approx \frac{1}{N} \sum_{i=1}^{N} I_j^{(i)}(w).$$

In figure 3 one can see that less than 5% of the simulations overflow the affected area and hit large parts of the non-affected area. Most of the simulations (90%) approximate the affected area quite well. Only the lateral extend in the beginning of the runout area appears overestimated by the simulations.

4. SUMMARY

In conclusion we can state that a probabilistic approach can successfully be applied to derive the posterior distribution of the 3-dimensional flow model parameter $\theta = \{\mu, \xi, e_b\}$ and the resulting distribution can be used to generate predictive avalanche simulations. These simulations can be used to assess the quality of the derived model (parameters) and give, moreover, empirical distributions of the estimator of each result variable, not just a point estimator. The resulting predicted result variables, like runout or deposit volume, are close to the documented event apart from the maximum flow velocity. Most of the simulated avalanches are $60\% - 75\%$ slower than the documented avalanche.

Further our approach contains several limitations. First, due to the complexity of our simulation concept, we cannot argue the convergence of the Markov chain, i.e. we are not able to formally check if the Markov chain has reached a stationary sate. However, we observed good mixing in the chain and thus high efficiency of the algorithm. Second we consider neither the measurement uncertainty in the release volume nor the measurement uncertainty in the topography, i.e. we do not handle the input variable X as a random variable and do not take it into account in the corresponding error in equation (6). Third we just give a rough estimation of the covariance matrix of the implementation error. A better assessment could be achieved investigating different implementations.
The improvement of the second and third point will be part of future research as well as taking into account multiple and unknown avalanche paths, similarly to the approach by Kofler et al. (2016) presented in the same proceedings.

5. ACKNOWLEDGEMENTS

This research work was carried out with the financial support of the projects EVAMETH (LE 07-13, BMLFUW - Austrian Ministry of Agriculture, Forestry, Environment and Water Management and the European Union) and CoSiCaCV (BMLFUW, WLV - Austrian Service for Torrent and Avalanche Control). We

thank M. Granig (WLV - Austrian Service for Torrent and Avalanche Control) and P. Sampl (AVL LIST GmbH) for their support and cooperation regarding SamosAT.

References

Brooks, S., Gelman, A., Jones, G. L. ., and Meng, X. (2011). *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC.

Brooks, S. and Roberts, G. O. (1998). Convergence assessment techniques for markov chain monte carlo. *Statistics and Computing*, 8(4):319–335.

Eckert, N., Naaim, M., and Parent, E. (2010). Long-term avalanche hazard assesment with a bayesian depth-avarged propagation model. *Journal of Glaciology*, (56):563–586.

Eckert, N., Richard, D., and Parent, E. (2007). Revisiting statistical-topographical methods for avalanche predetermination: Bayesian modelling for runout distance predictive distribution. *Cold Regions Science and Technology*, (49):88–107.

Fischer, J., Kofler, A., Fellin, W., Granig, M., and Kleemayr, K. (2015a). Multivariate parameter optimization for computational snow avalanche simulation. *Journal of Glaciology*, (229).

Fischer, J.-T. (2013). A novel approach to evaluate and compare computational snow avalanche simulation. *Natural Hazards and Earth System Science*, 13(6):1655–1667.

Fischer, J. T., Kofler, A., Fellin, W., Granig, M., and Kleemayr, K. (2015b). Multivariate parameter optimization for computational snow avalanche simulation. *Journal of Glaciology*, 61(229):875–888.

Graham, C. and Talay, D. (2013). *Stochastic Simulation and Monte Carlo Methods: Mathematical Foundations of Stochastic Simulation*. Stochastic Modelling and Applied Probability 68. Springer-Verlag Berlin Heidelberg, 1 edition.

Gregory, P. C. (2005). *Bayesian Logical Data Analysis for the Physical Sciences*. cambridge univercity press.

Kaipio, J. and Somersalo, E. (2005). *Statistical and computational inverse problems*. Applied Mathematical Sciences. Springer.

Kofler, A., Fischer, J., Huber, A., Fellin, W., Rauter, M., Granig, M., Hainzer, E., Tollinger, C., and Kleemayr, K. (2016). Multivariate parameter optimization for operational application of extended kinetic theory in simulation software. In *International Snow Science Workshop, Breckenridge, Colorado, USA - 2016*.

Kowar, R. (2015). Statistical inverse problem. Lecture Notes.

Leichtfried, A. (2010). Empfehlung für die ermittlung der schneemächtigkeit in anbruchgebieten von lawinen für tirol und vorarlberg. Technical report, ZAMG - Zentralanstalt für Meterologie und Geodynamik, ZAMG Fürstenweg 180 A-6020 Innsbruck.

Nelson, R. B. (2006). *An Introduction to Copulas*. Springer Series in Statistics. Springer Science+Business Media.

Nummelin, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators (Cambridge Tracts in Mathematics 83)*. Cambridge Tracts in Mathematics 83. Cambridge University Press.

Roth, H. (2010). Sensitivity analysis with correlated variables. Master's thesis, Universitaet Innsbruck.

Sampl, P. and Granig, M. (2009). Avalanche simulation with SAMOS-AT. In *Proceedings of the International Snow Science Workshop, Davos*.

Shonkwiler, R. W. and Mendivil, F. (2009). *Explorations in Monte Carlo Methods*. Undergraduate Texts in Mathematics. Springer-Verlag New York, 1 edition.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, 1 edition.

Viertel, R. (1990). *Einführung in die Stochastik*. Springer.

Zwinger, T., Kluwick, A., and Sampl, P. (2003). Numerical simulation of dry-snow avalanche flow over natural terrain. *Dynamic Response of Granular and Porous Materials under Large and Catastrophic Deformations, Hutter, K. and Kirchner, N., Springer Verlag*, 11:161–194.