



HAL
open science

Modelling the spatio-temporal repartition of right-truncated data: an application to avalanche runout altitudes in Hautes-Savoie

Aurore A. Lavigne, Nicolas Eckert, Liliane Bel, Michaël Deschâtres, Éric
Parent

► To cite this version:

Aurore A. Lavigne, Nicolas Eckert, Liliane Bel, Michaël Deschâtres, Éric Parent. Modelling the spatio-temporal repartition of right-truncated data: an application to avalanche runout altitudes in Hautes-Savoie. *Stochastic Environmental Research and Risk Assessment*, 2017, 31 (3), pp.629-644. 10.1007/s00477-016-1301-z . hal-01433695

HAL Id: hal-01433695

<https://hal.science/hal-01433695>

Submitted on 31 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modelling the spatio-temporal repartition of right-truncated data: an application to avalanche runout altitudes in Hautes-Savoie

A. Lavigne · N. Eckert · L. Bel · M. Deschâtres · E. Parent

Received: date / Accepted: date

Abstract In this paper, we propose a novel approach for generating avalanche hazard maps based on the spatial dependence of avalanche runout altitudes. The right-truncated data are described with a Bayesian hierarchical model in which the spatio-temporal process is assumed to be the sum of independent spatial and temporal terms. Topography is roughly taken into account according to valley altitude and path exposition, and the spatial dependence is modelled with a Matérn covariance function. An application is performed to the Haute-Savoie region, French Alps. A spatial dependence in runout altitudes is identified, and an effective range of about 10 km is inferred. The temporal trend extracted highlights the increase of avalanche runout altitudes from 1955, attributed to both anthropogenic factors and climate warming. In a cross validation scheme, spatial predictions are provided on undocumented paths using kriging equations. All in all, although our model is unable to take into account small topographic features, it is a first-ever approach that produces very encouraging results. It could be enhanced in future work by incorporating a numerical physically-based code into the modelling.

Keywords Bayesian hierarchical model, geostatistic, snow avalanche, truncated data

A. Lavigne
Université Lille 3 / LEM-CNRS (UMR 9221), F-59653 Villeneuve d'Ascq, France.
E-mail: aurore.lavigne@univ-lille3.fr

N. Eckert
UR ETGR, Irstea / Université Grenoble Alpes, F-38402 St Martin d'Hères, France.

L. Bel
UMR MIA-Paris, AgroParisTech, INRA, Université Paris-Saclay, 75005 Paris, France.

M. Deschâtres
UR ETGR, Irstea / Université Grenoble Alpes, F-38402 St Martin d'Hères, France.

E. Parent
UMR MIA-Paris, AgroParisTech, INRA, Université Paris-Saclay, 75005 Paris, France.

1 Introduction

Hazard maps are useful tools to localize the danger engendered by extreme avalanches in mountains (Jamieson et al, 2008; Barbolini and Keylock, 2002). According to a reference hazard related to a set of suitable variables, “at risk” zones are defined by combining information about avalanche frequency and intensity. Whereas frequency is often considered stationary and rather easily estimated, mapping intensity is the crux of the issue (Ancey et al, 2004; Eckert et al, 2008). Two main approaches are encountered in the literature to evaluate avalanche intensity, namely, statistical methods and dynamic methods. The former are based on direct observations of avalanche intensity, commonly measured by runout distance, *i.e.* the distance covered by the avalanche. In dynamical approaches, runout distributions are derived from factors at the origin of avalanches, such as the accumulated snow depth in the release area, which are propagated according to dynamical models (Barbolini et al, 2000). However, these methods suffer from the small amount of data available, and from many uncertainties in physical laws driving the motion of snow (Jamieson et al, 2008).

Focusing on a region rather than on a path has a major advantage: observations coming from multiple paths can be gathered. However, runout distances cannot be directly compared between paths, since they depend on local climatic and topographical factors. As a consequence, data standardisation is needed. Within the common alpha-beta (Lied and Bakkehøi, 1980) and runout ratio (McClung and Lied, 1987) methods, each runout distance is transformed into a transferred variable, as a function of typical slope values from the paths where it occurred. All transferred variables are then assumed to be independent, identically distributed with their sample size large enough for robust inference. While these methods benefit from all regional information, they may be ineffective when the local path topography is irregular (Smith and McClung, 1997). In addition, they do not make use of any notion of proximity.

Even if, to our knowledge, the spatial dependence of avalanche intensity has never been demonstrated from a statistical point of view, some hints may suggest its reality. Indeed, avalanches arise from topographic and weather factors that are both spatially distributed. For instance, studying the spatial repartition of extreme snowfalls (Gaume et al, 2013) and snow depths (Blanchet and Davison, 2011), two factors impacting directly avalanche activity, have shown that they are spatially dependent. Moreover, the existence of significant spatial dependence in avalanche frequencies in the northern French Alps has been shown in Eckert et al (2007, 2010).

Bayesian hierarchical spatio-temporal models (BHM) (Cressie and Wikle, 2011) are popular tools to analysis complex spatio-temporal structure on many kinds of data; they are encountered in various domains such as epidemiology (Best et al, 2005), ecology (Munoz et al, 2013) and environment (Vanem et al, 2012). In the BHM framework, observations are modelled conditionally to the spatio-temporal process and the conditional definition allows for the splitting of the model into three layers: observation - process - prior. The observa-

tion model takes into account data specificities, whereas the spatio-temporal process is generally modelled with the convenient Gaussian process. Under the Bayesian framework, inference is generally straightforward thanks to the Gibbs algorithm (Gilks et al, 1995b).

The main issue in applying the framework to our case arises from data truncation. Indeed, in the French avalanche database we work with, avalanches are recorded only if their runout altitude is lower than a given threshold. This special case differs from the more common spatial censored data (Stein, 1992), because several repetitions of truncated data are supplied for each site (for us, the avalanche path), and the proportion of non-observed data is unknown and depends on the site. In such a case, it is known that ignoring the truncation will produce biased estimators (A’Hearn, 2004), and so a conditional distribution should be assigned to observations. However, parameter estimation under this distribution is uneasy, because truncation distorts the likelihood surface to create large flat zones of instabilities. As a consequence, the variance of maximum likelihood estimators is high, and modified estimators obtained with penalised likelihood (Cope, 2011), or restricted likelihood (A’Hearn, 2004) have been developed. These two methods aim at introducing a small bias in order to decrease the estimator’s variance. Under the Bayesian framework, Zhou et al (2014) report a good behaviour of the Bayesian estimates, even if non-informative Jeffrey’s priors are used.

In this paper, we develop a spatio-temporal BHM to integrate within a unified framework, all the information about avalanche intensity in Haute-Savoie, the northern most French Alps department. On the observation layer the truncated observations are modelled with a conditional distribution. We reduce the instabilities linked to this distribution by restricting the parameter space to a reasonable domain, with constraints on the mean inter-annual runout altitudes, given by physical laws or expert knowledge. On the process layer, the spatio-temporal covariance is assumed to be separable, and we model the spatial dependence between mean inter-annual runout altitudes with a Matérn covariance function. As a result, we show that spatial dependence exists in runout altitudes. We then take advantage of the identified spatial pattern to predict avalanches on non-documented paths.

The data, their origins and their specificities are described in Section 2. Our specifically designed BHM with a Gaussian truncated distribution as model for observations is described in Section 3, as well as the constraints on parameters. This model is inferred under the Bayesian paradigm. In Section 4, we discuss two points concerning Bayesian inference: the prior choice and the specificities of the Gibbs sampler we propose. In Section 5, we display results regarding the application on avalanche runout altitude. Finally, in Section 6 we explore the predictive abilities of our approach.

2 Avalanche data in Haute-Savoie

A systematic survey of avalanche events began in France, in the early 20th century under the scope of the EPA (*Enquête Permanente des Avalanches*). The aim of the EPA survey was to better understand avalanche activity in order to limit damage caused to forests, roads and inhabited areas. Consequently, a set of paths was chosen and the runout altitude of avalanches occurring on these paths has constantly been recorded since then. EPA data are accessible from the website http://www.avalanches.fr/epa_presentation-donnees/.

In 2002, an observation threshold was further defined for each path. It corresponds to a level above the bottom of the path, such that only the avalanches running down below it are recorded. The thresholds are chosen to be easily detectable by observers such as roads and rivers, but also such that a substantial part of the total avalanche activity, and at least the avalanches arriving close to potential elements at risk, are recorded. Table 1 displays the distribution of the altitudes between valley altitudes and thresholds. On average, the thresholds are located 300 m above the valley altitude, but this distance varies from one path to another within a range of ~ 1000 m. Before 2002, the threshold was implicitly used by most of the observers and we do not know the proportion of recorded and non-recorded avalanches above the threshold. Consequently, here, over the entire of this study period, we consider only the avalanches with a runout altitude below the threshold and treat the data as right-truncated.

We have at our disposal a database of $n = 5331$ avalanches that have been recorded in Haute-Savoie during the 1925-2012 period, indexed by year and path. They occurred on 389 paths, among the 553 paths currently surveyed by the EPA in Haute-Savoie. These 389 paths are located in the eastern part of the department where altitudes are on average high, and most of them are aligned on both sides of valleys, as shown in Fig. 1. In particular, more than 90 paths are concentrated along the Arve valley, where the well-known township of *Chamonix-Mont-Blanc* is situated, and where avalanche danger is known to be very high (Naaim et al, 2010, 2013).

Table 1 provides some indicators of the path features, emphasizing the heterogeneity of the paths. Concerning the avalanche record frequency, fewer than 5 avalanches are recorded for 41 % of the paths, but for 6 % of them, the number of records exceeds 50. The within path avalanche activity is also heterogeneous: the standard deviation (sd) of recorded runout altitude is high for some paths, but it generally does not exceed 50 m (for 61 % of the paths). Specifically, for 42 paths, the standard deviation is null, meaning that all the avalanches have exactly the same runout altitude. This result is not surprising, given that the precision of measurement is at least 10 m and surely larger in the past. More generally, this heterogeneity in local runout variability expresses the large range of topographies in the sample of studied paths. Indeed, regular or moderately varying path profiles lead to low variability of runout altitudes, whereas for paths with more complex topography the variability can

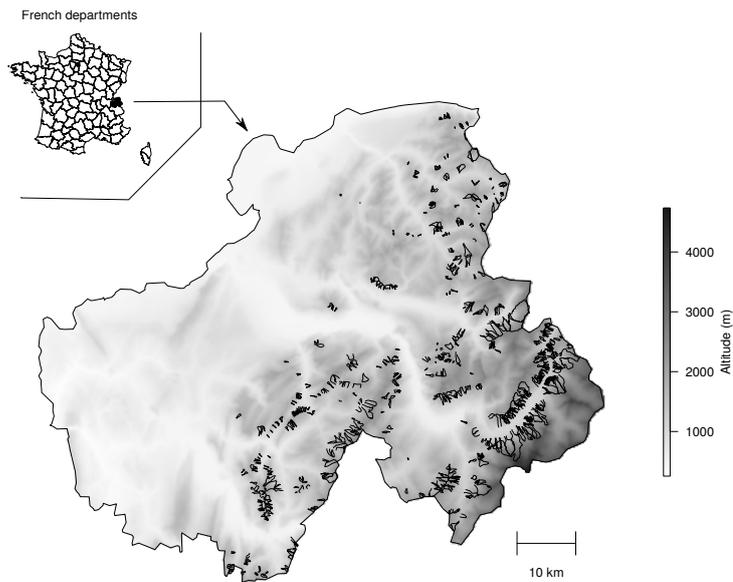


Fig. 1 Localization of the EPA avalanche paths in Haute-Savoie, French Alps. The paths are figured out by a contour line, the gray scale in the background represents the altitude.

Table 1 Repartition of three path features, given in number of paths (Nb.) a) Altitude difference Δ between the observation threshold and the valley. b) Number of avalanches by paths recorded during the period 1925-2012. c) Empirical standard deviation of the runout altitude computed by path.

a									
Δ]0, 100[[100, 200[[200, 400[[400, 600[[600, 1020[
Nb.	44	127	163	42	13				
b									
Nb. of events	1	2-5	6-10	11-20	21-30	31-50	51-100	177	
Nb.	65	97	71	70	35	28	22	1	
c									
sd	NA	0]0, 15[[15, 30[[30, 50[[50, 75[[75, 100[[100, 250[
Nb.	65	42	34	59	65	74	26	24	

be much higher, as demonstrated by long range runout simulations (Meunier and Ancy, 2004; Eckert et al, 2009).

All in all, the avalanche altitude data at our disposal is rather impressive, but with many particular features, which have to be taken into account in an efficient modelling.

3 Hierarchical framework for kriging truncated data

3.1 A spatio-temporal model for right-truncated data

We make use of a hierarchical specification to take into account all the data features at the observation level, and to easily model space and time fluctuations of runout altitudes on a Gaussian latent layer (Cressie and Wikle, 2011).

We opt for Gaussian truncated distributions as an observation model to take into account that the data are observed below a threshold; this choice is discussed in Section 5.5. In addition to the truncation, we chose an heteroscedastic model, *i.e.* path-specific variances in order to tackle the large between path variability of empirical standard deviations.

We denote Y_{cti} the observed runout altitude for avalanche $i \in \{1 \cdots n_{ct}\}$, for the path $c \in \{1, \cdots n_c\}$, during the year $t \in \{1 \cdots n_t\}$, s_c the observation threshold for path c , μ_{ct} the mean of the complete (that is not truncated) Gaussian distribution and σ_c^2 its variance. We assume that conditionally to μ_{ct} and σ_c^2 , the variables Y_{cti} are independent and their probability density function (pdf) is expressed as

$$f(y_{cti}|\mu_{ct}, \sigma_c^2) = \frac{1}{\Phi\left(\frac{s_c - \mu_{ct}}{\sigma_c}\right) \sqrt{2\pi\sigma_c^2}} \exp\left(\frac{-1}{2\sigma_c^2}(y_{cti} - \mu_{ct})^2\right) \mathbf{1}_{y_{cti} < s_c}, \quad (1)$$

where $\Phi(\cdot)$ is the cumulative density function (cdf) of the standard normal distribution.

On the latent layer, we aim at modelling the spatio-temporal fluctuations of the runout altitude. Our choice, discussed in Section 5.5 and based on an empirical analysis of covariance matrices (Appendix A), is made on a simple spatio-temporal factorial decomposition. The annual mean by path μ_{ct} is supposed to be the sum of an annual term B_t and a spatial term C_c . For symmetry purposes between space and time terms, we assume that $\sum_c C_c = \sum_t B_t = 0$ and introduce a constant term α , and so μ_{ct} can be then written as

$$\mu_{ct} = \alpha + B_t + C_c. \quad (2)$$

The temporal term B_t is supposed to be the sum of a second order random walk g_t of variance δ_1 and an independent Gaussian noise ϵ_t of variance δ_0 , re-written as

$$B_t|g_t, \delta_0 \sim \mathcal{N}(g_t, \delta_0). \quad (3)$$

Under the Bayesian paradigm, the posterior mean of g_t approximates a smoothing spline (Wahba, 1978). This non-parametric smooth term has been proved to be convenient to model large-scale inter-annual variability, for instance in avalanche data (Lavigne et al, 2012), or related proxies (Rabatel et al, 2013).

The spatial term aims at reproducing the local variations and the spatial dependence between paths of the mean runout altitude. For the sake of simplicity, we denote as well c the centroid of the contour line of the path c . We assume that C is a Gaussian random field that can be written

$$C_c = \mu_c + A_c, \quad (4)$$

where A_c is a zero mean Gaussian random field with a Matérn isotropic covariance function (Matérn, 1960). Thus, the covariance between A_c and $A_{c'}$ is given by

$$\text{cov}(A_c, A_{c'}) = \begin{cases} \tau^2 \frac{1}{2^{\nu-1} \Gamma(\nu)} \left(\frac{h_{cc'}}{\phi}\right)^\nu K_\nu\left(\frac{h_{cc'}}{\phi}\right) & \text{if } c \neq c' \\ \rho^2 + \tau^2 & \text{otherwise,} \end{cases} \quad (5)$$

where $h_{cc'}$ is the Euclidean distance between the centroids c and c' , K_ν is the modified Bessel function of the second kind of order $\nu > 0$ and Γ is the gamma function. The Matérn class of covariance functions is recommended for its flexibility (Banerjee et al, 2003). Indeed, in addition to the sill τ^2 , the nugget ρ^2 and the range ϕ , the Matérn covariance function is also parametrised by the smoothness parameter ν . This parameter is related to the smoothness of the spatial process at small distances: the larger it is, the smoother the process. In particular, when the smoothness parameter is 0.5 the Matérn covariance function is the exponential covariance function, and in the extreme case where $\nu \rightarrow \infty$, the Matérn covariance is the Gaussian covariance. As it is well-established that this parameter is difficult to infer, we consider that it belongs to the set $\{0.5, 1, 1.5\}$. We do not consider higher values, because in practice, data do not carry enough information to distinguish $\nu = 2$ from $\nu > 2$ (Banerjee et al, 2003).

Two covariates are used to model the mean spatial term μ_c , namely, the altitude valley h_c and the release area exposure o_c , so that the mean μ_c is expressed as

$$\mu_c = a + bh_c + o_c. \quad (6)$$

$o_c \in \{o_N, o_S\}$ is a categorical variable with two levels, S for the south, south-east and southwestern directions and N for the other directions. It is intended to capture the fact that south facing slopes behave differently because of the influence of sun exposure on snow (melting, metamorphism, etc.). For identifiability purposes, we set $o_N = 0$.

3.2 Constraints on the mean runout altitude

In contrast to the usual case, maximum likelihood estimators for the mean and variance of a Gaussian truncated sample are not independent. Thus, the likelihood surface for both parameters is distorted and presents a large flat zone shown in Fig. 2. A couple of estimates close to the maximum likelihood estimate is as likely as a couple of estimates far from the likelihood maximum.

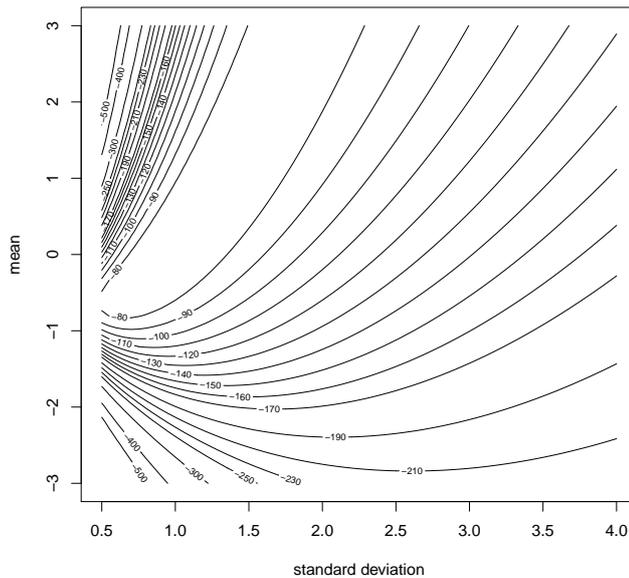


Fig. 2 log-likelihood for a sample of 100 independent observations of a standard Gaussian distribution right-truncated at the value 0.1.

As a consequence, estimation algorithms may fail to converge. We thus add constraints on the mean runout altitude to restrict the parameter space to reasonable values. More specifically, we have at our disposal the following information:

- at least 50% of the data are observed for each path, in accordance with the threshold selection;
- the valley bottom is a lower bound for all avalanches, hence the mean runout altitude is higher than (or ultimately equal to) the valley altitude.

These assertions can be expressed as constraints on the sum $\alpha + C_c$. The first can be written $q_{0.5}^{Y_c} < s_c$, with $q_{0.5}^{Y_c}$ the 50% quantile of the Gaussian distribution $\mathcal{N}(\alpha + C_c, \sigma_c^2)$. It means that the runout altitude mean by path $\alpha + C_c$ must not be higher than the observation threshold s_c . The second is simply written $h_c \leq \alpha + C_c$. These constraints are not very restrictive, but they are essential to make inference tractable.

3.3 Predictions

The spatial dependence in the mean runout altitude helps in making predictions for a set of undocumented paths c_0 . Since a different variance is set by path, we assume that the variance distribution for the predicted paths $\sigma_{c_0}^2$ is the mixture of the posterior distributions of the variances of the modelled path set. We also chose to make predictions for the entire study period,

and thus we integrate over time. Basically, we kriging the spatially dependent term $\mathbf{A} = (A_1, \dots, A_{n_c})'$ and add the variability corresponding to the several sources of uncertainties in order to obtain the predictive distribution $[\mathbf{Y}_{c_0} | \mathbf{Y}]$, where $\mathbf{Y} = (Y_{111}, \dots, Y_{n_c n_t n_c n_t})'$. Considering the hierarchical structure of our model, it is straightforward to simulate predictive runout altitudes for new paths from the MCMC outputs. Details of the method are given in Appendix B.

4 Bayesian inference

4.1 Prior choices

Prior distribution assignment to the parameters is a sensitive task since they must reflect the *a priori* knowledge about the parameters (Carlin and Louis, 1997; Gelman et al, 2004). Here, we use both non-informative and informative prior distributions elicited from expert knowledge. We assume that parameters are independent *a priori*, and most of the time we make recourse to the convenient conjugate priors.

Non-informative vague priors are assigned to the regression parameters, α prior is taken improper, and a zero mean Gaussian distribution with variance 10^7 is chosen for a , b and σ_S . For the variance components of the temporal part, the improper density $1/\sqrt{\delta_0 \delta_1}$ is selected, as advised by Speckman and Sun (2001). Independent inverse gamma distributions with 0.0001 for rate and shape parameters are assigned to τ^2 and ρ^2 . A uniform distribution $[0, \phi_{max}]$ is assigned to the range parameter ϕ . Since, ϕ is highly dependent on the smoothness parameter ν , ϕ_{max} is selected according to the value of ν considered, such that the effective range, *i.e.* the distance beyond which the covariance function does not exceed 5% of the sill, follows a uniform $[0, 100 \text{ km}]$ distribution. Since 100 km is about the distance between the two furthest sites of the study region, this prior can be considered as non-informative.

The n_c variances σ_c^2 , $c \in \{1, \dots, n_c\}$ are assumed to be independent and inverse gamma distributed. To introduce extra data information and help ensure the stability of the algorithm, informative priors are elicited. The hyperparameters of each gamma distribution are set such that the standard deviation mean and variance equal 25 meters, a value considered reasonable by experts. This distribution is then right and left truncated to ensure that no unrealistic variance can be sampled; the standard deviation σ_c must not be smaller than 30 m, representing the observation error, and not larger than 316 m.

4.2 A Gibbs sampler to draw from the posterior

Relying on the conditional independence relations in the BHM, for each value of the smoothness parameter ν , the joint posterior distribution can be written

as

$$\begin{aligned}
[\alpha, \mathbf{C}, \mathbf{B}, \sigma_c^2, \mathbf{g}, a, b, \\
o_S, \mathbf{A}, \delta_0, \delta_1, \phi, \tau^2, \rho^2 | \mathbf{Y}, \nu] \propto & \left(\prod_{c=1}^{n_c} \prod_{t=1}^{n_t} \prod_{i=1}^{n_{ct}} [Y_{cti} | \alpha, B_t, C_c, \sigma_c^2] \right) \\
& \times \left(\prod_{c=1}^{n_c} [C_c | a, b, o_S, A_c, \rho^2] \right) \times \left(\prod_{t=1}^{n_t} [B_t | g_t, \delta_0] \right) \\
& \times [\mathbf{A} | \tau^2, \phi, \nu] \times [\mathbf{g} | \delta_1] \\
& \times [\alpha, \sigma_c^2, \delta_0, \delta_1, \tau^2, \rho^2, \phi, a, b, o_S],
\end{aligned} \tag{7}$$

which is a decomposition of the conditional equations defining the observation model (Eq. 1), the process model (Eqs. 2, 3, 4) and the prior model. A Gibbs sampler is implemented in C++ to draw parameters and latent variables from their joint posterior distribution. In case of a standard linear regression, the algorithm would be straightforward. Here, the truncation, the constraints on the parameters and the spatial covariance function complicate the inference. For the variance components of the spatial term, we rewrite the algorithm supplied by Diggle and Ribeiro (2007) and use a Metropolis-Hastings step to sample the range parameter.

For dealing with the truncation, we follow Griffiths (2004)' algorithm which proposes to add latent variables Z_{cti} sampled in the non-truncated distribution $\mathcal{N}(\mu_{ct}, \sigma_c^2)$. These latent variables can be interpreted as pseudo-data in the case of non-truncation, and conditionally to μ_{ct} and σ_c^2 they are linked to Y_{cti} , by the relationship $P(Z_{cti} < z_{cti}) = P(Y_{cti} < y_{cti})$ leading to

$$z_{cti} = \mu_{ct} + \sigma_c \Phi^{-1} \left(\frac{\Phi\left(\frac{y_{cti} - \mu_{ct}}{\sigma_c}\right)}{\Phi\left(\frac{s_c - \mu_{ct}}{\sigma_c}\right)} \right). \tag{8}$$

It is then straightforward to sample μ_{ct} and σ_c^2 conditionally to Z_{cti} . However, μ_{ct} and σ_c^2 are correlated to Z_{cti} , and to decrease the dependence between μ_{ct} and σ_c^2 into the MCMC sample, we sample directly σ_c^2 from Y_{cti} and not from Z_{cti} . The posterior conditional distribution is then given by

$$\begin{aligned}
[\sigma_c^2 | \mathbf{Y}, \alpha, C_c, \mathbf{B}] \propto & \prod_{t=1}^{n_t} \prod_{i=1}^{n_{ct}} \frac{1}{\Phi\left(\frac{y_{cti} - \mu_{ct}}{\sigma_c}\right)} \left(\frac{1}{\sigma_c^2}\right)^{a_0 + \frac{1}{2}m_c + 1} \\
& \times \exp \left[-\frac{1}{\sigma_c^2} \left(b_0 + \sum_{t=1}^{n_t} \sum_{i=1}^{n_{ct}} (y_{cti} - \mu_{ct})^2 \right) \right],
\end{aligned} \tag{9}$$

where m_c is the number of avalanches which occurred on path c , and where a_0 and b_0 are the hyperparameters of the inverse gamma prior distribution. We draw σ_c^2 according to the pdf of Eq. 9 thanks to the adaptive rejection Metropolis sampling (ARMS) algorithm (Gilks et al, 1995a), an adaptation of the adaptive rejection sampling (Gilks and Wild, 1992) suitable when the pdf is nearly log-concave.

Another difficulty comes from the constraints assigned to the parameters, which are of two types. Firstly, the identifiability constraints stipulate that $\sum_{t=1}^{n_t} B_t = 0$ and $\sum_{c=1}^{n_c} C_c = 0$. A deterministic relation links the components of vectors \mathbf{B} , respectively \mathbf{C} , and we have to draw them in subspaces of dimension $n_t - 1$, respectively $n_c - 1$. For this, we follow the recommendations of

Gelman (2005) and sample B_1, \dots, B_{n_t-1} conditionally to $\sum_{t=1}^{n_t} B_t = 0$, then we set $B_{n_t} = -\sum_{t=1}^{n_t-1} B_t$. Secondly, the mean runout altitude by path, *i.e.* $\alpha + C_c$, is bounded. Specifically, we have $h_c \leq \alpha + C_c < s_c$ for all $c \in \{1, \dots, n_c\}$. These constraints are linear and we follow the algorithm of Rodriguez-Yam et al (2004) to sample the variables α and \mathbf{C} in their conditional complete posterior distribution. This algorithm consists in *i)* removing the dependence within elements of \mathbf{C} or α thanks to a simple linear transformation, *ii)* sampling the independent elements of the transformed variable in the convenient truncated univariate normal distribution. For this step, we use the method and the code proposed by Chopin (2011). Details of the implemented algorithm are given in Appendix C.

After testing the ability of the algorithm to estimate the parameters from simulated data correctly, three chains of 40,000 iterations including a burn-in period of 10,000 iterations were launched for each of the tested models (each of the ν values). In each case, the good mixing of the chains was checked with the Gelman-Rubin convergence diagnostic (Brooks and Gelman, 1998) and confirmed by visual inspection.

5 Results and Discussion

5.1 Choice of the smoothness parameter

The choice of the smoothness parameter of the Matérn covariance function is based on the overall model fit, measured by the expected deviance, where deviance is defined as -2 times the log-likelihood (Gelman et al, 2004). The results for each of the three considered smoothness parameters are given in Table 2. Surprisingly, we obtain the same expected deviance, meaning that for each of the three models, the fit to the data is roughly the same. However, we notice an interesting difference between models. Indeed, the effective range and the sill decrease with the increase of the smoothness parameter, whereas the nugget increases. This suggests that when the smoothness parameter is high, the spatial dependence is weaker and model predictions based on spatial dependence are less accurate. Consequently, in what follows, we decide to assign the value 0.5 to the parameter ν . This choice will be corroborated in section 6 when we compare prediction performances with the three values of this parameter. Table 3 gives the posterior mean, standard deviation and median of all estimated parameters when the exponential covariance ($\nu = 0.5$) is considered.

5.2 Spatial repartition

The μ_{ct} posterior variance is given by $\text{Var}(C_c) + \text{Var}(B_t)$. The ratio $\frac{\text{Var}(C_c)}{\text{Var}(C_c) + \text{Var}(B_t)}$ equals 98.7% indicating that spatio-temporal variations of runout-altitudes are mainly due to spatial variations. The large variations of the spatial component

ν	0.5	1	1.5
ϕ	3209 m	1652 m	1262 m
Effective range	9460 m	6591 m	5982 m
sill τ^2	4669 m ²	3968 m ²	3738 m ²
nugget ρ^2	2478 m ²	2842 m ²	3019 m ²
Expected Deviance	4.718.10 ⁴	4.718.10 ⁴	4.718.10 ⁴

Table 2 Comparison of the parameters posterior mean for the three Matérn variogram models.

are explained for 81% by the covariates, the valley altitude, h_c and the path exposition o_c . The value of the valley altitude coefficient ($b=0.83$ with standard deviation 0.03) indicates that its influence is dominating. Its introduction into the modelling enables the comparison of the runout altitude irrespective of the valley altitude. The influence of the path exposition is not so clear. Avalanches on south exposed paths generally stop higher (7.20 m), but the variability on the parameter is large (the standard deviation is 9.35 m), and we cannot fully conclude to a real effect of the exposition in the runout altitude, even if the obtained result is physically consistent. In fact, it is expected that for south exposed slopes, there is generally less snow and it is less cold than in north exposed slopes; the runout distance is therefore shorter and the runout altitude higher.

Figure 3 displays the purely spatially structured random term A_c . We distinguish clearly a spatial pattern. For instance in the Arve valley (highlighted with a black solid line) avalanches seem to stop lower on the west side than on the east side. The estimates of the variance parameters confirm the presence of spatial dependence. The sill is estimated to 4669 m², whereas the nugget is only 2479 m² and the effective range is about 9.6 km. The uncertainty regarding the three parameters (sill, range and nugget) are shown in Fig. 4, which represents the posterior mean semivariogram and two hundred realizations. Whereas the variability on the sum of sill and nugget is high, for the range parameter, nearly all of the realizations agree, suggesting that the spatial dependence does not originate from embedded-scale phenomena.

This effective range obtained here for runout altitudes is shorter than the effective range of 42 km found in the study of avalanche frequencies on the French Alps (Lavigne et al, 2015). Two factors account for this difference. First, the avalanche frequency is primarily sensitive to climate variations whereas the runout altitude is, in addition, strongly constrained by local topography (paths more or less channelized, land cover, etc.), leading to range reduction. Second, the frequency data considered in Lavigne et al (2015) were aggregated by township, attenuating the effect of local topography even more, and leading to a range probably longer than if counts by paths had been considered.

Nevertheless the significant value for the range we obtain here demonstrates the presence of spatial dependence in the runout altitude after crude correction of the topographic effects. This result is of crucial importance, since it has

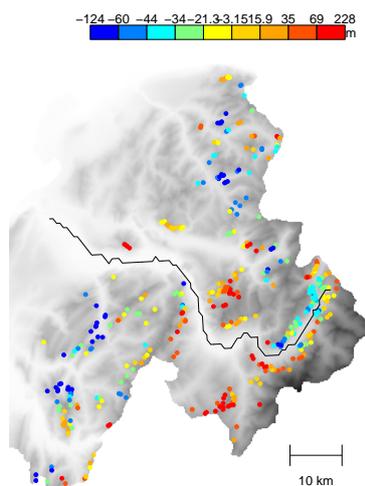


Fig. 3 Posterior mean of the random field \mathbf{A} of the spatial term. In black solid line, the Arve river.

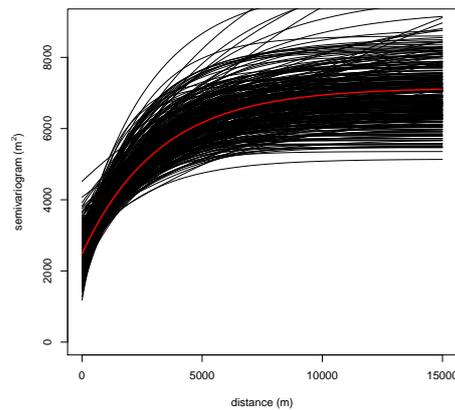


Fig. 4 Red: mean posterior variogram, obtained with sill, range and nugget posterior mean; black: posterior variogram realizations. (m^2)

never been obtained before to our knowledge and proposes an alternative to the standard approaches based on topographic variables (Lied and Bakkehøi, 1980; McClung and Lied, 1987; Sigurðsson et al, 1998). More precisely, it suggests that it is conceivable to transfer information within a 10 km neighbourhood, as an alternative to the introduction of additional variables required by standard methods for avalanche hazard mapping, which are not always easily available.

5.3 Temporal evolution

The temporal term represents only 1.3% of the variability of the spatio-temporal term μ_{ct} , the largest part being explained by spatial variability. However, significant year-to-year fluctuations and a smooth trend can be extracted, as shown in Fig. 5. The year-to-year fluctuations are essentially captured by the unstructured temporal term whose variance (δ_0) is estimated to 138 m^2 . These random fluctuations mainly result from the annual variations of extreme weather conditions which influence avalanching.

At a broader time scale, the inter annual variations are modelled by the structured temporal term \mathbf{g} . From 1925 to 1955, the mean runout altitude decreases slowly at a rate of about 0.6 m/year , a trend difficult to interpret since, even if climate reconstructions at this time are not very precise, they do not seem to show colder temperature or higher snow records (Beniston et al, 1997; Corona et al, 2010). Furthermore, we do not have any other reconstruction of runout altitude fluctuations at this early period, in France or other countries.

On the other hand, the 1955-1975 period is marked by a sudden increase of the mean runout altitude, increasing from 1200 to 1245 m in only 20 years.

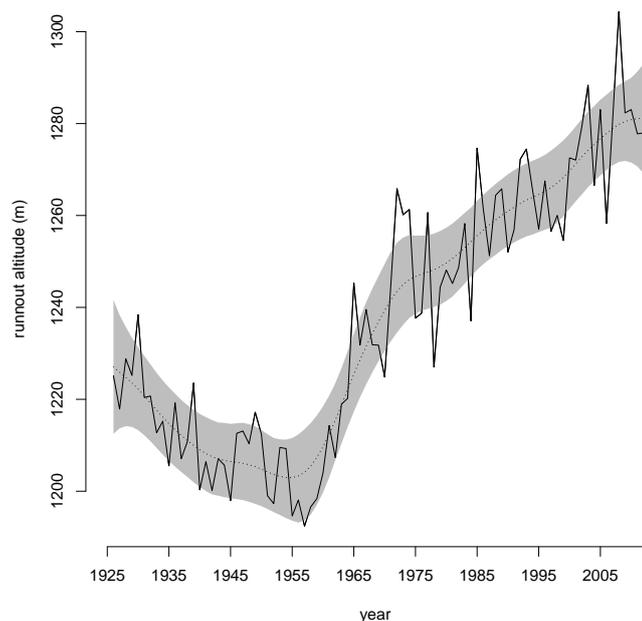


Fig. 5 Posterior mean of the temporal term (solid line), posterior mean of the smooth term (dotted line), and 95% credible interval of the smooth term (grey area).

This change is not reported by Eckert et al (2013) when both the Savoie and the Haute-Savoie departments are under study simultaneously with a simpler modelling approach. The divergence may be explained by the strong relative importance in the latter study of Savoie in terms of total avalanche activity, and by the numerous constructions of avalanche defense structures made at this period in Haute-Savoie to reduce avalanche danger in ski resorts. Therefore, the sudden increase we highlight here should have mostly direct anthropic origins and does probably not reflect a climatic trend.

From around 1975 to 2005, a slower increase (0.95 m/year) is observed, up to around 1280 m, coherent with what is observed for the Savoie and Haute-Savoie departments together, and also with climate warming and the subsequent decrease of snow depths at low and moderate altitudes.

Finally, over the most recent years, this trend stops in accordance with winters which are somewhat harsher (Eckert et al, 2013).

5.4 High variability in the between path variance

For most of the paths (91%), the runout altitude variance is small and does not exceed 5000 m², but few paths can have a much higher variance reaching 25000 m². Understanding the origin of this between path variability is a key step for

Table 3 Parameter estimates: posterior mean, posterior standard deviation and posterior median

parameters	mean	standard deviation	median
ϕ	3209.40 m	2102.98 m	2719.85 m
τ^2	4669.10 m ²	1230.73 m ²	4540.29 m ²
ρ^2	2478.59 m ²	651.17 m ²	2485.95 m ²
δ_0	137.69 m ²	33.47 m ²	133.76 m ²
δ_1	0.85 m ²	1.06 m ²	0.49 m ²
α	1237.86 m	1.22 m	1237.84 m
a	-935.56 m	31.03 m	-936.02 m
b	0.83	0.03	0.83
o_S	7.20 m	9.35 m	7.13 m

making predictions on new paths. Here, the largest variance reaches 26425 m² (standard deviation 162.6 m) and corresponds to quite a well documented path with 10 observations, but for which two avalanches stop on a gentler slope before the valley. Thus, the specific topography of the site (two distinct runout zones) is responsible for the high variance observed. In other cases of large variance, the paths are still well documented, and the presence of forest in the slope at the bottom of the path seems to be the cause for the runout altitude dispersion, because trees are able to stop the less dynamic events (Bartelt and Stockli, 2001). Taking into account covariates in the variance model linked to topography or land cover would therefore be the next step to improve prediction. However, at this stage, it is impossible to obtain this information automatically, and spatial prediction will therefore be performed by employing regional distribution, *i.e.* without considering local specificities in runout distance dispersion around their mean.

5.5 Model choice

The BHM chosen may seem rather simple since it combines a truncated Gaussian distribution on the observation layer and an additive spatio-temporal process. However, we think that introducing more complexity in the modelling is not appropriate given the data quality and our objectives.

In the context of extreme event prediction, generalised extreme values (GEV) distributions are traditionally used to describe avalanche runout distances (McClung, 2001; Keylock, 2005). More generally, skewed distributions are popular for fitting environmental data. For instance the gamma distribution is broadly used in hydrology (Zalina et al, 2002), and the lognormal or Birnbaum-Saunders distributions (Leiva et al, 2015; Garcia-Papani et al, 2016) often model air quality measurements. Our runout altitude records do not allow us to study the distribution tail, because for many avalanche paths, the runout zone is flat, leading to a left censorship. The GEV distributions have thus been ruled out. This choice is confirmed by a preliminary analysis,

in which the adjustment of the normal, lognormal, Weibull, and gamma truncated distributions are compared for paths on which more than 5 avalanches are recorded. The four distributions are adjusted for each path, and the AICs are averaged over the paths. The best distributions seem to be the normal (average AIC equals 236.0) and the lognormal (average AIC = 236.2). However the Weibull distribution (AIC = 237.1) performs quite well also. In contrast, the adjustment of the gamma distribution is worse (AIC = 282.8). Considering the ease of use of the normal distribution, we chose it rather than the lognormal, keeping in mind that the model is less accurate for tail distribution. The additive spatio-temporal structure chosen (Eq. 2) enables us to directly apply the Matérn covariance and the second order random walk, two efficient techniques developed for spatial and temporal data, respectively. Moreover, thanks to the spatio-temporal separability, the computational cost is considerably reduced. However, this simple model does not allow interactions between time and space, a constraint which is not always fulfilled in practice (Fuentes, 2006). In order to investigate the presence of interactions between space and time, a preliminary study of the empirical spatio-temporal covariance matrices (Cressie and Wikle, 2011) has been carried out (see Appendix A). It shows that assuming space-time interactions is not relevant, essentially because spatial dependence is much higher than temporal dependence. This point is conceptually corroborated from a physical point of view, as soon as we recognise that in this relatively small region the climate is constant. We expect that the same variations of the climate will have the same effects on runout altitudes for all the paths of the study region.

6 Prediction performances

We use the prediction method detailed in Section 3 to assess the ability of the model to predict runout altitude distributions on new undocumented sites. We consider a cross-validation framework, in which the 51 paths with more than 30 observations are used as a test sample. They are split in 10 groups and, alternatively, each group is removed from the study. For each path of the group, draws are generated from the predictive distribution. We then compare this predictive distribution to the empirical distribution originating directly from the data, with numerical and graphical tools.

First, the Kullback-Leibler (KL) divergence of the predictive from the empirical distribution is examined. This non-symmetric measure of difference between two probabilities can be interpreted as the loss in information by using the predictive distribution rather than the empirical one. Thus, the KL divergence is always positive, and null when the empirical and predictive distributions are identical. Here, we compute an empirical KL divergence thanks to the formula provided by Pérez-Cruz (2008), which has good asymptotic properties. Figure 6 maps the predicted paths, in colors, and their associated KL divergence. The best predicted paths, with the smallest KL divergences, are located along the Arve valley, a well documented region with paths of

rather regular topography. For one path highlighted in Fig. 6, predictions are particularly bad: its KL divergence reaches 3.17, whereas the divergence of the other paths does not exceed 2.1. By studying this site in detail, we note that it is located at a higher altitude, with a very high observation threshold and avalanches runout at about 2200 m a.s.l (above sea level). Hence, this elevated path does not look like the standard paths of our database, which partially explains its bad prediction.

As predictions are carried out on the basis of spatial dependence, predictive paths with numerous well documented paths nearby should be better predicted than isolated paths. To check this assertion, we investigate the correlation coefficient between the KL divergence and the number of observations within a given radius around the predicted path (Tab. 4). The path mentioned above is excluded from this study. When the radius is small (2 km) no correlation is observed, but when it is greater (5-15 km), there is a significant negative correlation meaning that the loss in information by using the model rather than the empirical distribution decreases when the amount of information around a predicted path increases. This consistent result is encouraging, since it means that our approach relying on the spatial dependence between avalanche runout altitudes is an efficient means to predict the distribution on a new path.

Figure 7 shows the probability-probability plot of the predictive and empirical distributions. It illustrates the variability of the prediction ability of our method. Indeed, for some paths, the curve is close to the first bissector, but in some other cases, it can be very far from it. Two hypotheses could be at the origin of these relatively bad predictions (i) the choice of the covariance model, or (ii) the lack of topographic variables in the model.

In order to check the first hypothesis, we compare prediction abilities of three Matérn covariance models with smoothness parameters equal to 0.5, 1 and 1.5. The prediction ability is measured by summing the KL divergence of all of the 51 paths. It equals 59.13 when the exponential covariance model is considered, whereas it reaches 59.78 (resp. 60.33) when ν equals 1 (resp. 1.5). This small difference in favour of the exponential model, suggests that, considering the data at our disposal, the covariance model does not seem to be the cause of the rather bad predictions, and that sticking on the exponential choice is sensible.

In order to identify the relation between local topography and predictions, four paths are investigated more deeply. Their predictive and empirical cumulative probability distributions are given in Fig. 8, while topographic details are supplied in Tab. 5. In the three cases of bad prediction (2, 3 and 4), a topographic particularity exists in the predicted path or in its surroundings (the slope is irregular with flat or steep parts). On the contrary, for the case of rather good prediction (1), the path is regular and ends up far from the valley bottom. This suggests that taking into account complementary topographic features would improve the prediction for bad predicted paths. So far, only dynamical models based on a detailed topography of the path are able to deal with these irregularities, at the price of a computational burden and only if a reliable, detailed record of local activity is available. However, this leaves

distance (km)	Correlation coefficient	pvalue	Confidence interval
2	-0.24	0.10	$[-0.49; 0.04]$
5	-0.39	$5.7 \cdot 10^{-3}$	$[-0.60; -0.12]$
10	-0.45	$1.1 \cdot 10^{-3}$	$[-0.65; -0.19]$
15	-0.47	$6.2 \cdot 10^{-4}$	$[-0.66; -0.22]$

Table 4 Correlation coefficient between the Kullback-Leibler divergence and the number of observation at 2, 5, 10, and 15 km around the predicted path, pvalue of the test of the null assumption of no correlation and confidence interval.

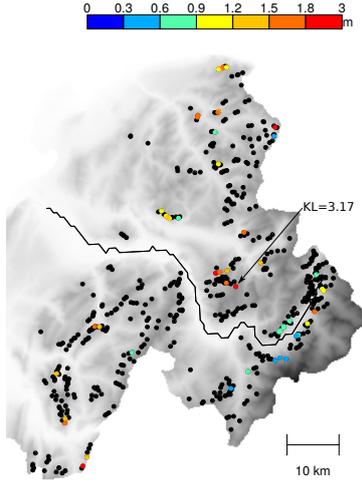


Fig. 6 Kullback-Leibler divergence for the 51 predictions paths (colored). The paths in black are in the learning sample.

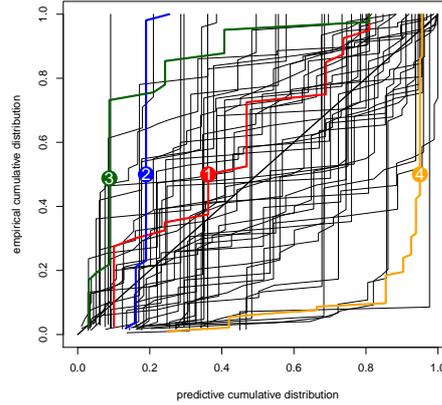


Fig. 7 Probability-probability plot of the predictive and the empirical distributions. Each curve represents one of the 51 paths where the runout altitude is predicted (validation sample). Number labels identify the four paths described in Tab. 5.

the door open for further developments of our regional approach integrating dynamical simulations to cope for the specific topography of each path.

7 Conclusion and outlooks

This study is the first model-based spatio-temporal analysis of avalanche runout altitudes. The model we propose takes into account a major specificity of the data, namely, their truncation. Space and time are considered as additive effects and the model is developed within a hierarchical framework under the Bayesian paradigm, so that additional extra-data information can be injected through constraints on the parameters and elicited prior distributions. Finally, with due recourse to MCMC methods, we have designed an efficient inference scheme.

Important quantities for the avalanche hazard assessment field have been inferred. The most relevant are the spatial covariance parameters, which demon-

Nb.	EPA name EPA map	KL div.	Prediction quality	Descriptive features of the path and of its surroundings
1	74276008 AF65	0.67	Good	Small regular path with a runout area in the forest far from the valley bottom. It is surrounded by paths with a similar topography.
2	74280041 AK63	1.26	Observations all at the same altitude	Small path on a forest area which ends up in a flat-bottomed valley.
3	74270013 AN63	2.09	Over-estimated mean	Observed avalanche runouts at ≈ 900 m, at the bottom of a steep slope. Runout altitudes are predicted around 1000 m, in the middle of the steep slope. Nearby, another path presents a flat part at 1000 m a.s.l., with a 1080 m mean runout altitude.
4	74063018 AD67	1.86	Under-estimated mean	Path with a specific topography: a flat part between 1350 and 1300 m a.s.l., then the slope is steeper and the path is channeled down to 1200 m a.s.l.

Table 5 Details of four predicted paths: Kullback-Leibler divergence between the predicted and empirical distribution (KL div.), prediction quality and descriptive features of the path, *i.e.* topographic elements which can explain the good or bad prediction. For more details, the topographic maps where paths are located can be download from ftp://avalanchesftp.grenoble.cemagref.fr/epaclpa/EPA_feuilles_carte/alpes/, the EPA map gives the name of the map to select, and the paths can be localized with their names (the first five numbers correspond to the postal code of the township and the three last to the path number).

strate the presence of a spatial dependence, with an effective range of about 10 km. In comparison to other quantities such as avalanche frequencies, this effective range is clearly shorter. The difference may be explained by the factor at the origin of the spatial structure. In our case, the topography has been identified as the major control variable, whereas in the frequency studies, climate may play a larger role.

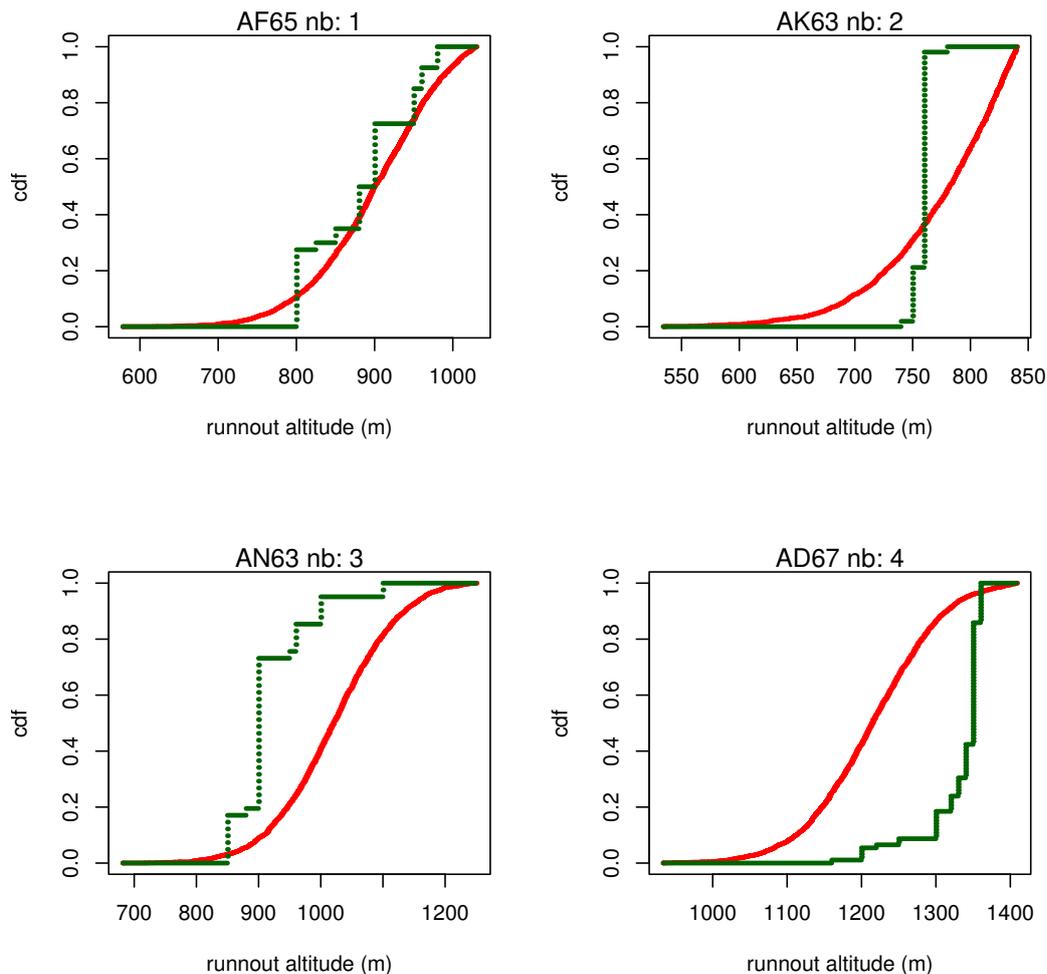


Fig. 8 Predictive (red line) and empirical (green line) cumulative probability distributions for the four paths detailed in Tab. 5.

The temporal term extracted from the model presents high year-to-year fluctuations and a global increase of the runout altitude over recent decades. The results roughly reflect with those obtained in the northern Alps by Eckert et al (2013), but a 1955-1975 period of high increase has appeared here, presumably caused by an intense construction of avalanche defense structures in Haute Savoie at this time. The moderate increase over more recent decades is coherent with the temperature increase reported in the Northern Alps (Durand et al, 2009), and may be related to global warming.

Relying on spatial dependence, a consistent framework for prediction based on kriging equations has been proposed. It is a major advance with regards to

existing approaches in the avalanche literature, in which the site effect is only addressed with topographic variables.

However, since this work is a first-ever approach, certain limitations remain. The variability between paths is high, and ground factors such as the presence of forest or the path shape have been identified as good candidates to explain it. Unfortunately, it is not possible to automatically extract these variables easily and quickly at a regional scale. Consequently, at this stage, the path-specific variances in the model are an obstacle to a conclusive predictive framework, and is partially responsible for some poor predictions. Introducing dynamical calculation in the hierarchical framework would be an attractive possibility for further improvements. Indeed, it could solve the two major difficulties identified: distribution tail and local topography.

Finally, to achieve the ultimate goal, *i.e.* hazard mapping, spatialized return levels are necessary. The results obtained for the runout altitude in a spatial context should now be coupled with frequency estimates to supply them as function of space and, possibly, time.

Acknowledgements The authors thank the ANR research program MOPERA (Modélisation probabiliste pour l'Étude du Risque d'Avalanche - www.avalanches.fr/mopera-projet/) for funding this work.

References

- A'Hearn B (2004) A restricted maximum likelihood estimator for truncated height samples. *Economics & Human Biology* 2(1):5–19
- Ancey C, Gervasoni C, Meunier M (2004) Computing extreme avalanches. *Cold Regions Science and Technology* 39(2):161–180
- Banerjee S, Gelfand A, Carlin B (2003) Hierarchical modeling and analysis for spatial data. Crc Press
- Barbolini M, Keylock C (2002) A new method for avalanche hazard mapping using a combination of statistical and deterministic models. *Natural Hazards and Earth System Science* 2(3/4):239–245
- Barbolini M, Gruber U, Keylock C, Naaim M, Savi F (2000) Application of statistical and hydraulic-continuum dense-snow avalanche models to five real European sites. *Cold Regions Science and Technology* 31(2):133–149
- Bartelt P, Stockli V (2001) The influence of tree and branch fracture, overturning and debris entrainment on snow avalanche flow. *Annals of Glaciology* 32(1):209–216
- Beniston M, Diaz H, Bradley R (1997) Climatic change at high elevation sites: an overview. *Climatic Change* 36(3-4):233–251
- Best N, Richardson S, Thomson A (2005) A comparison of Bayesian spatial models for disease mapping. *Statistical methods in medical research* 14(1):35–59
- Blanchet J, Davison A (2011) Spatial modeling of extreme snow depth. *The Annals of Applied Statistics* 5(3):1699–1725

- Brooks S, Gelman A (1998) Some Issues for Monitoring Convergence of Iterative Simulations. In: In Proceedings of the Section on Statistical Computing. ASA
- Carlin B, Louis T (1997) Bayes and empirical Bayes methods for data analysis. *Statistics and Computing* 7(2):153–154
- Chopin N (2011) Fast simulation of truncated Gaussian distributions. *Statistics and Computing* 21(2):275–288
- Cope EW (2011) Penalized likelihood estimators for truncated data. *Journal of Statistical Planning and Inference* 141(1):345–358
- Corona C, Guiot J, Edouard J, Chalié F, Büntgen U, Nola P, Urbinati C (2010) Millennium-long summer temperature variations in the European Alps as reconstructed from tree rings. *Climate of the Past* 6(3):379–400
- Cressie N, Wikle C (2011) *Statistics for spatio-temporal data*. Wiley
- Diggle P, Ribeiro PJ (2007) *Model-based geostatistics*. Springer Science & Business Media
- Durand Y, Laternser M, Giraud G, Etchevers P, Lesaffre B, Mérindol L (2009) Reanalysis of 44 Yr of Climate in the French Alps (1958-2002): Methodology, Model Validation, Climatology, and Trends for Air Temperature and Precipitation. *Journal of Applied Meteorology and Climatology* 48(3):429–449
- Eckert N, Parent E, Belanger L, Garcia S (2007) Hierarchical Bayesian modelling for spatial analysis of the number of avalanche occurrences at the scale of the township. *Cold Regions Science and Technology* 50(1):97–112
- Eckert N, Parent E, Naaim M, Richard D (2008) Bayesian stochastic modelling for avalanche predetermination: from a general system framework to return period computations. *Stochastic Environmental Research and Risk Assessment* 22(2):185–206
- Eckert N, Parent E, Faug T, Naaim M (2009) Bayesian optimal design of an avalanche dam using a multivariate numerical avalanche model. *Stochastic Environmental Research and Risk Assessment* 23(8):1123–1141
- Eckert N, Parent E, Kies R, Baya H (2010) A spatio-temporal modelling framework for assessing the fluctuations of avalanche occurrence resulting from climate change: application to 60 years of data in the northern French Alps. *Climatic Change* 101(3-4):515–553
- Eckert N, Keylock C, Castebrunet H, Lavigne A, Naaim M (2013) Temporal trends in avalanche activity in the French Alps and subregions: from occurrences and runout altitudes to unsteady return periods. *Journal of Glaciology* 59(213):93–114
- Fuentes M (2006) Testing for separability of spatial-temporal covariance functions. *Journal of statistical planning and inference* 136(2):447–466
- Garcia-Papani F, Uribe-Opazo M, Leiva V, Aykroyd R (2016) Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment* pp 1–20
- Gaume J, Eckert N, Chambon G, Naaim M, Bel L (2013) Mapping extreme snowfalls in the French Alps using max-stable processes. *Water Resources Research* 49(7):1079–1098

- Gelman A (2005) Analysis of variance—why it is more important than ever. *The Annals of Statistics* 33(1):1–53
- Gelman A, Carlin J, Stern H, Rubin D (2004) *Bayesian Data Analysis*. Chapman & Hall/CRC
- Gilks W, Wild P (1992) Adaptive rejection sampling for Gibbs sampling. *Applied Statistics* 41(2):337–348
- Gilks W, Best N, Tan K (1995a) Adaptive rejection Metropolis sampling within Gibbs sampling. *Applied Statistics* pp 455–472
- Gilks W, Richardson S, Spiegelhalter D (1995b) *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Interdisciplinary Statistics, Taylor & Francis
- Griffiths W (2004) A Gibbs’ sampler for the parameters of a truncated multivariate normal distribution. In: Becker R, Hurn S (eds) *Contemporary issues in economics and econometrics: theory and application*, Edward Elgar Pub, pp 75–91
- Jamieson B, Margreth S, Jones A (2008) Application and limitations of dynamic models for snow avalanche hazard mapping. In: *Proceedings of the ISSW*, pp 730–739
- Keylock C (2005) An alternative form for the statistical distribution of extreme avalanche runout distances. *Cold regions science and technology* 42(3):185–193
- Lavigne A, Bel L, Parent E, Eckert N (2012) A model for spatio-temporal clustering using multinomial probit regression: application to avalanche counts. *Environmetrics* 23:522–534
- Lavigne A, Eckert N, Bel L, Parent E (2015) Adding expert contributions to the spatiotemporal modelling of avalanche activity under different climatic influences. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* DOI 10.1111/rssc.12095
- Leiva V, Ferreira M, Gomes M, Lillo C (2015) Extreme value Birnbaum–Saunders regression models applied to environmental data. *Stochastic Environmental Research and Risk Assessment* 30:1–14
- Lied K, Bakkehøi S (1980) Empirical calculations of snow-avalanche run-out distance based on topographic parameters. *Journal of Glaciology* 26:165–177
- Matérn B (1960) *Spatial variation*, vol 49. Meddelanden från Statens Skogsforskningsinstitut [2nd Edition (2013), Vol 36, Springer Science & Business Media]
- McClung D (2001) Extreme avalanche runout: a comparison of empirical models. *Canadian Geotechnical Journal* 38(6):1254–1265
- McClung D, Lied K (1987) Statistical and geometrical definition of snow avalanche runout. *Cold Regions Science and Technology* 13(2):107–119
- Meunier M, Ancely C (2004) Towards a conceptual approach to predetermining long-return-period avalanche run-out distances. *Journal of Glaciology* 50(169):268–278
- Munoz F, Pennino MG, Conesa D, López-Quílez A, Bellido JM (2013) Estimation and prediction of the spatial occurrence of fish species using Bayesian latent Gaussian models. *Stochastic Environmental Research and Risk As-*

- essment 27(5):1171–1180
- Naaim M, Faug T, Naaim F, Eckert N (2010) Return period calculation and passive structure design at the Tacconnaz avalanche path, France. *Annals of Glaciology* 51(54):89–97
- Naaim M, Durand Y, Eckert N, Chambon G (2013) Dense avalanche friction coefficients : influence of physical properties of snow. *Journal of Glaciology* 59:771–782
- Pérez-Cruz F (2008) Kullback-Leibler divergence estimation of continuous distributions. In: *Information Theory, 2008. ISIT 2008. IEEE International Symposium on, IEEE*, pp 1666–1670
- Rabatel A, Letréguilly A, Dedieu J, Eckert N (2013) Changes in glacier equilibrium-line altitude (ELA) in the western Alps over the 1984-2010 period: evaluation by remote sensing and modeling of the morpho-topographic and climate controls. *The Cryosphere* 7:1455–1471
- Rodriguez-Yam G, Davis R, Scharf L (2004) Efficient Gibbs sampling of truncated multivariate normal with application to constrained linear regression. Unpublished Manuscript
- Sigurðsson S, Jónasson K, Arnalds P (1998) Transferring avalanches between paths. In: Hestnes E (ed) *Proceedings of the anniversary conference 25 years of snow avalanche research*, Voss, 12–16 May, Norwegian Geotechnical Institute, publication 203, pp 259–263
- Smith M, McClung D (1997) Characteristics and prediction of high-frequency avalanche runout. *Arctic and Alpine Research* 29(3):352–357
- Speckman P, Sun D (2001) Bayesian nonparametric regression and autoregression priors
- Stein ML (1992) Prediction and inference for truncated spatial data. *Journal of Computational and Graphical Statistics* 1(1):91–110
- Vanem E, Huseby A, Natvig B (2012) A Bayesian hierarchical spatio-temporal model for significant wave height in the North Atlantic. *Stochastic environmental research and risk assessment* 26(5):609–632
- Wahba G (1978) Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society Series B (Methodological)* 40(3):364–372
- Zalina M, Desa M, Nguyen V, Kassim A (2002) Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science & Technology* 45(2):63–68
- Zhou X, Giacometti R, Fabozzi FJ, Tucker AH (2014) Bayesian estimation of truncated data with applications to operational risk measurement. *Quantitative Finance* 14(5):862–888

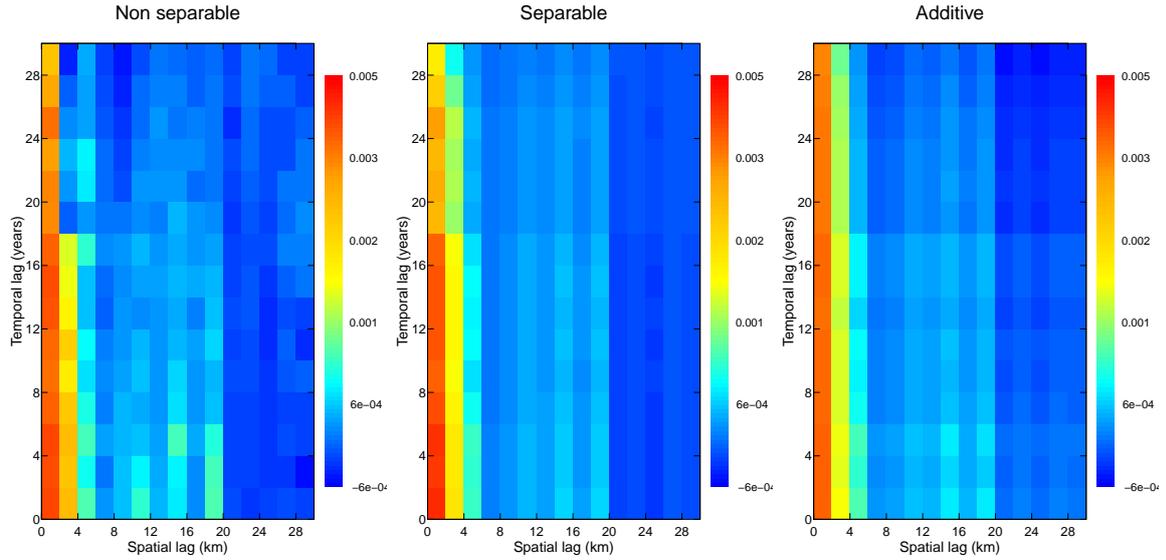


Fig. 9 Empirical non-separable (left), separable (center) and additive (right) spatio-temporal covariance matrices. The mean is only spatially indexed, and is computed from the valley altitude.

Appendices

A Empirical study of the spatio-temporal covariance structure

In order to investigate the presence of spatio-temporal dependence in runout altitudes, empirical covariance matrices are estimated in the non-separable, separable and additive cases (Fig. 9). Let Y_{ict} be the runout altitude of avalanche i on path c the year t . The empirical spatio-temporal covariance at spatial lag $[h_{k-1}, h_k]$ and time lag $[\tau_{l-1}, \tau_l]$ is given by:

$$\hat{C}_Y(h_k, \tau_l) = \frac{1}{|N(h_k, \tau_l)|} \sum_{(ict, i'c't') \in N(h_k, \tau_l)} (Y_{ict} - \hat{\mu}_Y(c))(Y_{i'c't'} - \hat{\mu}_Y(c'))$$

with $N(h_k, \tau_l)$ the set of avalanche pairs $(ict, i'c't')$ such that the distance between sites c and c' is in the interval $]h_{k-1}, h_k]$ and the lag time $t - t'$ belongs to the interval $[\tau_{l-1}, \tau_l]$. $|N(h_k, \tau_l)|$ is the number of pairs. The reference mean $\hat{\mu}_Y(c)$ is calculated by linear regression on the valley altitude, the best covariate: $\hat{\mu}_Y(c) = 6.3 + 7.410^{-4}h_c$. From the empirical non-separable covariance matrix, the additive and separable covariance matrices have been estimated by least squares. In details, for each spatial and time lag, estimates of γ_l^T and γ_k^S are found such that:

$$\hat{C}_Y(h_k, \tau_l) = \gamma_l^T + \gamma_k^S + \epsilon_{kl}$$

for the additive covariance matrix, and:

$$\hat{C}_Y(h_k, \tau_l) = \gamma_l^T \gamma_k^S + \epsilon_{kl}$$

for the separable covariance matrix, where ϵ_{kl} is the error.

Figure 9 shows that the non-separable covariance matrix is similar to the separable and additive ones, essentially because the spatial dependence is much higher than the temporal dependence. This pleads further in favor of our choice of a simple model with additive effects.

B Decomposition of the predictive distribution

First we integrate the annual predictive distribution $[\mathbf{Y}_{c_0t}|\mathbf{Y}]$ for time t and paths c_0 over the time to make prediction for the entire study period:

$$[\mathbf{Y}_{c_0}|\mathbf{Y}] = \sum_{t=1}^{n_t} \frac{1}{n_t} [\mathbf{Y}_{c_0t}|\mathbf{Y}].$$

By using the hierarchical structure of the model we can express this distribution conditionally to the process level:

$$[\mathbf{Y}_{c_0t}|\mathbf{Y}] = \int [\mathbf{Y}_{c_0t}|\mathbf{C}_{c_0}, B_t, \alpha, \sigma_{c_0}^2][\mathbf{C}_{c_0}, B_t, \alpha, \sigma_{c_0}^2|\mathbf{Y}]d(\mathbf{C}_{c_0}, B_t, \alpha, \sigma_{c_0}^2). \quad (\text{B.1})$$

Let $\boldsymbol{\theta}_1 = (B_t, \alpha, \sigma_{c_0}^2)'$, we decompose the distribution $[\mathbf{C}_{c_0}, \boldsymbol{\theta}_1|\mathbf{Y}]$ in the same way to make the spatial term \mathbf{A}_{c_0} appear,

$$[\mathbf{C}_{c_0}, \boldsymbol{\theta}_1|\mathbf{Y}] = \int [\mathbf{C}_{c_0}, \boldsymbol{\theta}_1|\mathbf{A}_{c_0}, a, b, o_S, \rho^2][\mathbf{A}_{c_0}, a, b, o_S, \rho^2|\mathbf{Y}]d(\mathbf{A}_{c_0}, a, b, o_S, \rho^2). \quad (\text{B.2})$$

Finally, we express $[\mathbf{A}_{c_0}, \boldsymbol{\theta}_2|\mathbf{Y}]$, with $\boldsymbol{\theta}_2 = (a, b, o_S, \rho^2)'$, conditionally to \mathbf{A} rather than \mathbf{Y} :

$$[\mathbf{A}_{c_0}, \boldsymbol{\theta}_2|\mathbf{Y}] = \int [\mathbf{A}_{c_0}, \boldsymbol{\theta}_2|\mathbf{A}, \phi, \tau^2][\mathbf{A}, \phi, \tau^2|\mathbf{Y}]d(\mathbf{A}, \phi, \tau^2). \quad (\text{B.3})$$

The joint Gaussian distribution for the vector $[\mathbf{A}, \mathbf{A}_{c_0}|\phi, \tau^2]$ is:

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{A}_{c_0} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tau^2 \begin{bmatrix} \boldsymbol{\Sigma}_{\phi}^{\cdot\cdot} & \boldsymbol{\Sigma}_{\phi}^{\cdot 0} \\ \boldsymbol{\Sigma}_{\phi}^{0\cdot} & \boldsymbol{\Sigma}_{\phi}^{00} \end{bmatrix} \right),$$

where $\boldsymbol{\Sigma}_{\phi}^{\cdot\cdot}$ and $\boldsymbol{\Sigma}_{\phi}^{00}$ are the variance-covariance matrices of \mathbf{A} and \mathbf{A}_{c_0} , respectively, and where $\boldsymbol{\Sigma}_{\phi}^{\cdot 0}$ and $\boldsymbol{\Sigma}_{\phi}^{0\cdot}$ are covariance matrices between elements of \mathbf{A} and \mathbf{A}_{c_0} . Then the formulas for conditional Gaussian distributions give:

$$\begin{aligned} E(\mathbf{A}_{c_0}|\mathbf{A}, \phi, \tau^2) &= \boldsymbol{\Sigma}_{\phi}^{0\cdot} \boldsymbol{\Sigma}_{\phi}^{\cdot\cdot -1} \mathbf{A} \\ V(\mathbf{A}_{c_0}|\mathbf{A}, \phi, \tau^2) &= \boldsymbol{\Sigma}_{\phi}^{00} - \boldsymbol{\Sigma}_{\phi}^{0\cdot} \boldsymbol{\Sigma}_{\phi}^{\cdot\cdot -1} \boldsymbol{\Sigma}_{\phi}^{\cdot 0}. \end{aligned}$$

By combining Eq. (B.2) and (B.3) within Eq. (B.1), and using the Bayes' rule we obtain

$$\begin{aligned} [\mathbf{Y}_{c_0,t}|\mathbf{Y}] &= \int [\mathbf{Y}_{c_0t}|\mathbf{C}_{c_0}, \boldsymbol{\theta}_1] \\ &\quad [\mathbf{C}_{c_0}|\mathbf{A}_{c_0}, \boldsymbol{\theta}_2][\mathbf{A}_{c_0}|\boldsymbol{\theta}_3] \\ &\quad [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3|\mathbf{Y}]d(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3), \end{aligned} \quad (\text{B.4})$$

with $\boldsymbol{\theta}_3$ the vector $(\mathbf{A}, \phi, \tau^2)'$.

C Algorithm details

We detail the algorithm to draw \mathbf{C} under the constraint $\sum_{c=1}^{n_c} C_c = 0$, the same method is used to draw \mathbf{B} under the constraint $\sum_{t=1}^{n_t} B_t = 0$.

Let \mathbf{X}_C the $n \times n_c$ matrix such that $\mathbf{X}_{C_{ij}} = 1$ if the avalanche i occurs on path j , 0 either, \mathbf{D} the diagonal variance-covariance matrix of \mathbf{Z} , $\mathbf{D} = \text{diag}(\sigma_{c_i}^2)_{i=1,n}$, with c_i the path label for avalanche i , and $\mathbf{1}_c$ the n_c length column vector of ones.

The complete posterior distribution of \mathbf{C} is given by,

$$\begin{aligned} &\mathbf{C}|\mathbf{Z}, \mathbf{B}, \alpha, \mathbf{D}, E[\mathbf{C}], \rho^2 \sim \mathcal{N}(\mathbf{m}, \boldsymbol{\Sigma}) \\ \text{with } \begin{cases} \mathbf{m} = (\mathbf{X}'_C \mathbf{D}^{-1} \mathbf{X}_C + \frac{1}{\rho^2} \mathbf{I}_c)^{-1} (\mathbf{X}'_C \mathbf{D}^{-1} (\mathbf{Z} - \alpha - \mathbf{X}_B \mathbf{B}) + \frac{1}{\rho^2} E(\mathbf{C})) \\ \boldsymbol{\Sigma} = (\mathbf{X}'_C \mathbf{D}^{-1} \mathbf{X}_C + \frac{1}{\rho^2} \mathbf{I}_c)^{-1} \end{cases} \end{aligned}$$

At a second step we write the joint distribution of $\mathbf{C}|\mathbf{Z}, \mathbf{B}, \alpha, \mathbf{D}, E[\mathbf{C}], \rho^2$, noted $\mathbf{C}|\gamma$ and its constraint $\mathbf{1}_c \mathbf{C} = 0$,

$$\begin{bmatrix} \mathbf{C}|\gamma \\ \mathbf{1}'_c \mathbf{C} \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{m} \\ \mathbf{1}'_c \mathbf{m} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Sigma} \mathbf{1}'_c \\ \mathbf{1}'_c \boldsymbol{\Sigma} & \mathbf{1}'_c \boldsymbol{\Sigma} \mathbf{1}'_c \end{bmatrix} \right).$$

The conditional distribution is then:

$$\begin{aligned} \mathbf{C}|\mathbf{Z}, \mathbf{B}, \alpha, \mathbf{D}, E[\mathbf{C}], \rho^2, \mathbf{1}'_c \mathbf{C} = 0 &\sim \mathcal{N}(\mathbf{m}_0, \boldsymbol{\Sigma}_0) \\ \text{with } \begin{cases} \mathbf{m}_0 = \mathbf{m} - \frac{\boldsymbol{\Sigma} \mathbf{1}'_c \mathbf{m}}{\mathbf{1}'_c \boldsymbol{\Sigma} \mathbf{1}'_c} \\ \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} - \frac{\boldsymbol{\Sigma} \mathbf{1}'_c \mathbf{1}'_c \boldsymbol{\Sigma}}{\mathbf{1}'_c \boldsymbol{\Sigma} \mathbf{1}'_c} \end{cases} &. \end{aligned} \quad (\text{C.1})$$

Only the $n_c - 1$ first components of \mathbf{C} are sampled, C_{n_c} is given by $C_{n_c} = -\sum_{c=1}^{n_c-1} C_c$.

Then we constrain the mean runout altitude by path, $\alpha + C_c$. We have $h_c < \alpha + C_c < s_c$ for all $c \in \{1, \dots, n_c\}$. In a Gibbs sampler, \mathbf{C} and α are drawn successively in their complete conditional distributions. We describe the method to take into account the constraint for \mathbf{C} , α is sampled in a similar way. The $n_c - 1$ first components of \mathbf{C} must fulfill the constraints:

$$\begin{cases} h_1 - \alpha < C_1 < s_1 - \alpha \\ \vdots \\ h_{n_c-1} - \alpha < C_{n_c-1} < s_{n_c-1} - \alpha \\ h_{n_c} - \alpha < -\sum_{c=1}^{n_c-1} C_c < s_{n_c} - \alpha, \end{cases} \quad (\text{C.2})$$

that we can write matrix-wise $\mathbf{V}\mathbf{C} < \mathbf{v}$ with \mathbf{V} the $k \times n_c - 1$ matrix of constraints. Here $k = 2n_c$, since there are two constraints for each line of the system Eq. (C.2), one bears on s_c , the other one on h_c . The posterior distribution of \mathbf{C} is thus a multivariate truncated distribution whose mean and variance for the complete version are given by Eq. (C.1) and whose support $S = \{\mathbf{x} \in \mathbb{R}^{n_c-1} : \mathbf{V}\mathbf{x} < \mathbf{v}\}$. We follow the two-step algorithm of Rodriguez-Yam et al (2004). In the first step, we remove the dependence between the components of \mathbf{C} by considering the vector $\mathbf{w} = \mathbf{L}\mathbf{C}$, with \mathbf{L} such that $\mathbf{L}\boldsymbol{\Sigma}_0\mathbf{L}' = \mathbf{I}_{n_c-1}$ and \mathbf{I}_{n_c-1} the identity matrix. It has the following truncated distribution:

$$\begin{aligned} \mathbf{w}|\mathbf{Z}, \mathbf{B}, \alpha, \mathbf{D}, E[\mathbf{C}], \rho^2, \mathbf{1}'_c \mathbf{C} = 0 &\sim \mathcal{N}_S(\mathbf{L}\mathbf{m}_0, \mathbf{I}_{n_c-1}) \\ \text{with support: } S = \{\mathbf{w} \in \mathbb{R}^{n_c-1} : \mathbf{V}\mathbf{L}^{-1}\mathbf{w} < \mathbf{v}\}. & \end{aligned}$$

In the second step, each component of \mathbf{w} is drawn successively conditionally to the others in a Gibbs sampling. In order to avoid the computation of the entire vector of constraints $\mathbf{V}\mathbf{L}^{-1}\mathbf{w}$ at each draw, we may notice that for each constraint $i \in \{1, \dots, k\}$, and each component of \mathbf{w} $j_0 \in \{1, \dots, n_c - 1\}$

$$\begin{aligned} (\mathbf{V}\mathbf{L}^{-1}\mathbf{w})_i &= \sum_{j=1}^{n_c-1} \left(\sum_{l=1}^{n_c-1} \mathbf{V}_{il} \mathbf{L}_{lj}^{-1} \right) w_j \\ &= \underbrace{\sum_{j \neq j_0}^{n_c-1} \left(\sum_{l=1}^{n_c-1} \mathbf{V}_{il} \mathbf{L}_{lj}^{-1} \right) w_j}_{u_i^{-j_0}} + \underbrace{\left(\sum_{l=1}^{n_c-1} \mathbf{V}_{il} \mathbf{L}_{lj_0}^{-1} \right) w_{j_0}}_{u_i^{j_0}}. \end{aligned}$$

The constraint i for the j_0 component of \mathbf{w} is written $u_i^{j_0} w_{j_0} < v_i - u_i^{-j_0}$. To update w_{j_0} , one only has to compute $u_i^{j_0}$ for each constraint. This step demands $k(n_c - 1)$ elementary operations, instead of $k(n_c - 1)^2$ operations when $\mathbf{V}\mathbf{L}^{-1}$ is computed naively.

Finally we implement the following algorithm for sampling the $n_c - 1$ components of \mathbf{C} from their truncated Gaussian distributions:

- Compute the vector $\mathbf{u} = \mathbf{V}\mathbf{C}$;
- Initiate the matrix \mathbf{L} , its inverse \mathbf{L}^{-1} , and the vector $\mathbf{w} = \mathbf{L}\mathbf{C}$;

- For each j in $\{1, \dots, n_c - 1\}$
 - Compute the k -dimension vectors $\mathbf{u}^j = (u_1^j, \dots, u_k^j)$, and $\mathbf{u}^{-j} = \mathbf{u} - \mathbf{u}^j w_j$;
 - Find the interval $[a_j, b_j]$ in which w_j is drawn to satisfy the k constraints;
 - Sample w_j in the normal truncated distribution $\mathcal{N}((\mathbf{L}\mathbf{m}_0)_j, 1)$ with support $[a_j, b_j]$. Here, we use the method and the code proposed by Chopin (2011);
 - Update \mathbf{u} as $\mathbf{u} = \mathbf{u}^{-j} + \mathbf{u}^j w_j$;
- Set \mathbf{C} as $\mathbf{C} = \mathbf{L}^{-1}\mathbf{w}$.