Optimizing and validating the Gravitational Process Path model for regional debris-flow runout modelling

Jason Goetz[1], Robin Kohrs[1], Eric Parra Hormazábal[2], Manuel Bustos Morales[3], María Belén Araneda Riquelme[3], Cristián Henríquez[3], Alexander Brenning[1]

5 [1] Department of Geography, Friedrich Schiller University Jena, German
[2] Institute of Geosciences, University of Potsdam, Germany
[3] Instituto de Geografía, Pontificia Universidad Católica de Chile, Chile. Centre for Sustainable Urban Development, CEDEUS. Centro de Cambio Global UC.

Correspondence to: Jason Goetz (jason.goetz@uni-jena.de)

10 **Abstract.** Knowing the source and runout of debris-flows can help in planning strategies aimed at mitigating these hazards. Our research in this paper focuses on developing a novel approach for optimizing runout models for regional susceptibility modelling, with a case study in the upper Maipo river basin in the Andes of Santiago, Chile. We propose a two-stage optimization approach for automatically selecting parameters for estimating runout path and distance. This approach optimizes the random walk and Perla's two-parameter modelling components of the open-source Gravitational Process Path (GPP) 15 modelling framework. To validate model performance, we assess the spatial transferability of the optimized runout model using spatial cross-validation, including exploring the model's sensitivity to sample size. We also present diagnostic tools for visualizing uncertainties in parameter selection and model performance. Although there was considerable variation in optimal parameters for individual events, we found our runout modelling approach performed well at regional prediction of potential runout areas. We also found that although a relatively small sample size was sufficient to achieve generally good runout 20 modelling performance; larger samples sizes (i.e. ≥80) had higher model performances and lower uncertainties for estimating runout distances at unknown locations. We anticipate that this automated approach using open-source software R and SAGA-GIS will make process-based debris-flow models more readily accessible and thus enable researchers and spatial planners to improve regional-scale hazard assessments.

1 Introduction

25 Knowledge of where debris flows are initiated and how far they travel is crucial for assessing their impact over large regions (Aleotti and Chowdhury 1999; van Westen et al. 2006). Commonly, debris-flow runout modelling for large areas is performed by first delineating source areas and then applying empirical-statistical or process-based numerical methods for simulating the runout characteristics (Blahut et al. 2010a; Horton et al. 2013; Mergili et al. 2019). There is a wide selection of heuristic, statistical and machine learning methods suitable for predicting sources areas in large regions (Chung and Fabbri 1999; Carrara

30   et al. 1999; Brenning 2005; Goetz et al. 2015a; Lombardo et al. 2018). There are also many empirical-statistical and numerical
methods available to model runout patterns – see McDougall (2017) for an overview.

Not all runout methods are suitable for application to large areas. Many of the physically based methods require event-specific
geotechnical and rheological parameters, such as material composition (e.g. bulk density and source depths) and flow
characteristics (e.g., flow discharge rates). These parameters, such as debris-flow volume, can be extremely difficult to obtain

35   for large areas, let alone single unobserved events (Marchi and D'Agostino 2004; Dong et al. 2009). Consequently, runout
modelling at larger scales has been progressing towards applying simplified conceptual models to simulate debris-flow patterns
across different environmental conditions. These models combine spreading algorithms to control runout path with empirical-
statistical or numerical friction models to simulate likely runout paths and distances (Guthrie et al. 2008; Horton et al. 2013;
Wichmann 2017; Mergili et al. 2019). Many of the spreading algorithms, including multiple flow direction models (Holmgren

40   1994), cellular automata (Guthrie et al. 2008) and random walk (Gamma 2000; Mergili et al. 2015), simulate runout paths
using only topographic data.

Calibration of model parameters continues to be one of the main challenges in runout modelling for single-events and over
large areas (Hungr 1995; van Westen et al. 2006; Schraml et al. 2015; McDougall 2017; Mergili et al. 2019). The objective
of model calibration is to determine parameter values that best capture main debris-flow characteristics such as runout distance,

45   velocity, and distribution of deposits (Hungr 1995; McDougall 2017). Approaches for model calibration include adjusting
parameters based on visual inspection (i.e. trial and error; Hungr 1995; Mergili et al. 2012); expert knowledge (Horton et al
2013); posterior analysis (Mergili et al. 2019; Aaron et al. 2019); and optimization algorithms that aim to minimize a cost
function, i.e. a quantitative measure of runout model performance. Some measures of performance include estimates of the
intersection-over-union (Galas et al. 2007), area under the receiver operating characteristic curve (AUROC; Cepeda et al.

50   2010; Mergili et al. 2015) and depth error (Aaron et al. 2019) of simulated and observed debris flows. Since most of these
calibration approaches are for single observed events, they rarely consider how transferable tuned parameter sets are from local
to regional applications.

Assessing spatial transferability is essential for testing the assumption a trained model based on a sample of events captures
the generic debris-flow characteristics across a region (Fabbri et al. 2003). The distribution of training data and the sample size

55   can have a strong influence on the model calibration and performance of regional models (Heckmann et al 2014; Petschko et
al 2014; Rudy et al 2016). For spatially distributed models, spatial transferability can be assessed by exploring model
parameter selection and performance under different spatial partitioning scenarios of training and test data (Wenger and Olden
2012; Brenning 2012; Schratz et al. 2019; Mergili et al. 2019). Although spatial transferability has been well explored for
regional landslide susceptibility models (Brenning 2005; Lombardo et al. 2014; Petschko et al. 2014; Goetz et al. 2015a; Cama

60   et al. 2017; Knevels et al. 2020), such analysis is not common for regional runout modelling.

In this study, we developed an optimization procedure for process-based models applied for regional simulation of debris-flow runout patterns. The performance evaluation focuses on the spatial transferability and sensitivity to sample size of an optimized regional debris-flow runout model. We achieve this by utilizing the open-source statistical software R to add optimization and evaluation functionality to the open-source Gravitational Process Path (GPP) modelling framework (Wichmann 2017).

65 Additionally, this study demonstrates the use of spatial cross-validation and visualization techniques to diagnose uncertainties in the prediction of source areas, runout paths and runout distances, including the sensitivity of optimized parameter selection. The aim of this research is to contribute to improving the development of quantitative techniques for runout model calibration and uncertainty estimation (McDougall 2017). This is especially important in large and inaccessible mountainous areas where various types of mass movements pose unique challenges to the safety of the local population, the integrity of transportation

70 infrastructure, and the reliability of drinking water supplies.

2 Materials and Methods

2.1.1 Study area

Out study area is the upper Maipo river basin (33°03′ to 34°18′ S, 800 m to 6108 m a.s.l.), located in the semi-arid Andes of central Chile. Debris-flow activity in remote and populated areas of the Maipo river basin have caused many deaths and severe

75 disruptions to critical transportation and water supply infrastructure supporting Chile's capital city Santiago (Hauser 2002; Sepúlveda et al. 2015; Moreiras and Sepúlveda 2015).

High-intensity rainfall (Sepúlveda et al 2015), rapid snowmelt (Moreiras et al 2012) and seismic activity (Serey et al 2019) are the main triggers of debris flows in this region. They occur in steep gullies and talus slopes consisting of gravel, small boulders and a fine sandy-silty matrix. Much of this material is from weathered volcanic and sedimentary rocks of the Abanico and

80 Farellones formations in the western Main Cordillera (Sepúlveda et al 2006). A typical runout track will cut through previously formed debris flow channels and alluvial fans, resulting in new erosion and deposition paths (Sepúlveda et al 2015). Rainfall-triggered runout distances in this area have been observed up to 5.5 km, and the thickness of deposits varies from 1 to 2 m in deposition areas (Sepúlveda et al 2015).

2.1.2 The debris-flow inventory

85 Debris flow polygons and source points were mapped based on photointerpretation of high-spatial resolution (0.50 m) satellite imagery (2000 to 2019) from CNES/Airbus and Maxar Technologies available through Google Earth Pro software, field observations, reviewing news articles and the compilation of data collected by public authorities. Each mapped polygon represents a debris-flow track that includes source, runout and deposition area. In total, 541 source points and 521 debris flow polygons were mapped (**Table 1**). Manually mapping all debris flows across the upper Maipo basin is a challenging task due

90   to its large geographical extent and its high abundance of mass movements. Therefore, a mapping strategy was employed that

divided the basin into 58 sub-drainage basins (5,439 km²), 45 (3,936 km²) of which were selected for mapping (**Figure 1**).

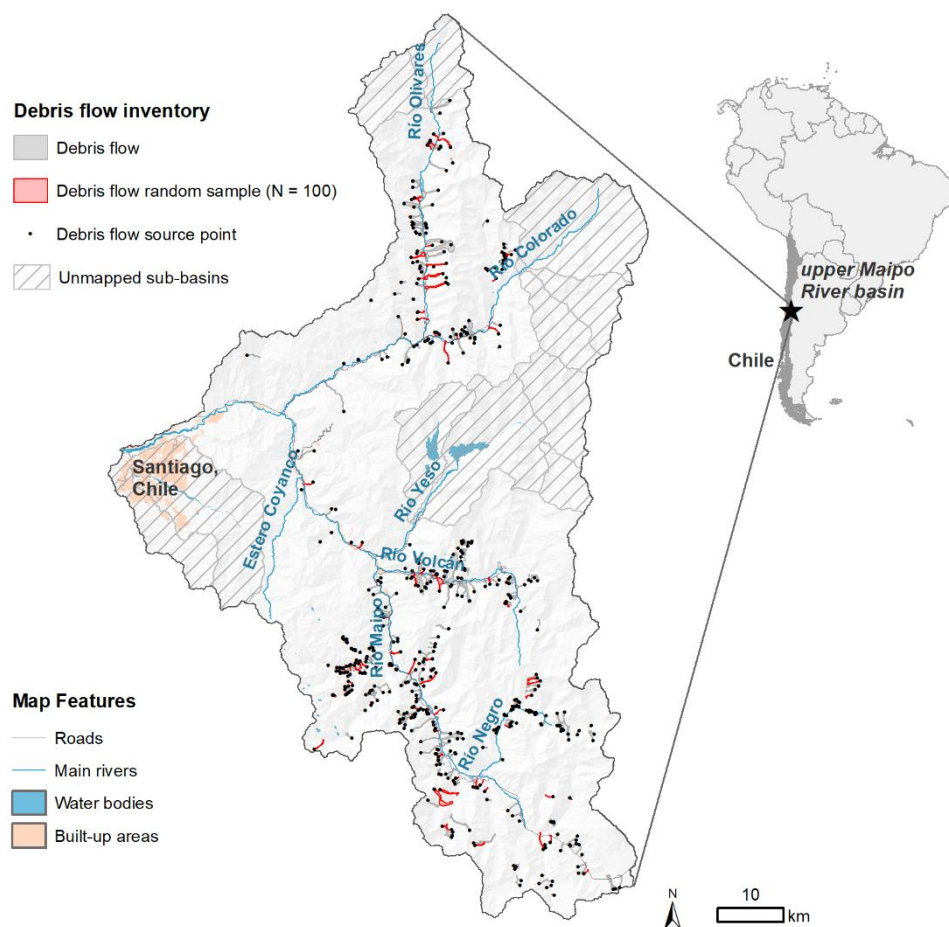



Figure 1: Map providing an overview of the debris flow polygons and source points mapped in the upper Maipo river basin.

95   **Table 1: Runout characteristics of the debris flow inventory**

	Median	IQR	Minimum	Maximum
Runout surface area (ha)	3.7	7.5	0.1	358.1
Runout distance (m)	729	823	54	5668
Max. elevation (m.a.s.l.)	2695	714	1256	4768

Note. IQR = interquartile range

2.1.3 Modelling debris-flow source areas

Potential debris-flow source areas were spatially predicted using a generalized additive model (GAM). In general, GAMs demonstrate good performance for susceptibility modeling compared to other commonly used physically-based and machine-

100 learning techniques (Goetz et al. 2011; Goetz et al. 2015a). To improve model generality and avoid overfitting, the GAM smoothing spline parameters were allowed a maximum 5 effective degrees of freedom (Wenger and Olden 2012; Goetz et al. 2015b). The training and test data were based on the common 1:1 sampling strategy (Heckmann et al. 2014) of presence to absence of source points.

The predictor variables of source areas included hillslope angle, elevation, catchment area, plan curvature and distance to

105 faults. These predictor variables generally have a high importance for modeling debris-flow initiation susceptibility as observed in previous works (Blahut et al. 2010b; Goetz et al. 2015a; Angillieri 2020). The publicly available ALOS PALSAR Radiometrically Terrain Corrected (RTC) high-resolution (12.5 m) digital elevation model (DEM; ASF DAAC) was used to derive terrain attributes. Before deriving the terrain attributes, mesh denoising was applied to the DEM to mitigate the propagation of artifacts such as high-frequency noise (Brock et al. 2020) in the prediction of source areas (Sun et al. 2007;

110 Stevenson et al. 2010). We used the implementation of this algorithm in SAGA-GIS. After denoising, an algorithm to fill sinks was applied to the DEM, and the terrain attributes were processed. Distance to faults was calculated as the Euclidean distance from the fault lines (Servicio Nacional de Geología y Minería de Chile 2003; scale 1:1,000,000).

The performance of the source area prediction model was assessed using repeated *k*-fold spatial cross-validation (Brenning 2012). Like cross-validation, spatial cross-validation randomly splits the data (e.g., source points and non-source points) into

115 *k* disjoint subsets, where the model is trained using *k*-1 sets and tested with remaining set during each cross-validation iteration. For *spatial* cross-validation, the data is divided into spatially disjoint sub-areas, in our case using the *k*-means clustering algorithm (Ruß and Brenning 2010). This approach should provide a rigorous estimate of the spatial transferability of a model by attempting to reduce spatial autocorrelation between test and training data (Brenning 2005; Wenger and Olden 2012; Schratz et al. 2019). We estimated model performance by repeating 5-fold spatial cross-validation 1,000 times. Model performance

120 was measured using the AUROC, which is an overall measure of goodness-of-fit that is independent of any particular decision threshold.

2.2 Modelling debris flow runout

The Gravitational Process Path (GPP; Wichmann 2017) model was used to regionally model runout. The GPP model is an open-source framework in the SAGA-GIS software that provides users with various model components to simulate runout

125 path, distance, velocity and deposition of material of mass movements (e.g. snow avalanches, rock falls, and debris flows). Due to the extent and remoteness of the study area, we focus on modelling the likely spatial patterns of runout. That is, we are not modelling flow velocity and depth.

Runout path was modelled using the random walk process path component of the GPP model (Gamma 2000). It is a common approach for debris-flow runout path modelling at medium scales (Mergili et al. 2012; Heckmann and Schwanghart 2013;

130   Mergili et al. 2015). Random walk models the potential path of runout by iteratively simulating (via Monte Carlo simulation) the downslope movement path of debris flows originating in an individual source-area grid cell. This simulation results in a grid with runout frequencies that indicate how many times a grid cell has been traversed. There are three parameters that need to be calibrated to obtain a desired runout path: (1) a slope threshold (°) defining where divergent flow is allowed; (2) the exponent of divergence that controls the amount of divergence, or lateral spreading; and (3) a persistence factor that controls

135   the direction of movement (Wichmann 2017).

Runout distance was constrained using Perla's two-parameter friction model (PCM; Perla et al. 1980) component of the GPP model. The PCM model, which is also a component of the Flow-R model (Horton et al. 2013), has also been used for modelling debris-flow behaviour at medium scales (Mergili et al. 2012; Heckmann and Schwanghart 2013; Mergili et al. 2015). It is a centre-of-mass model where motion is controlled by (1) the sliding friction coefficient and (2) the mass-to-drag ratio. The

140   sliding friction coefficient controls the velocity of movement and the mass-to-drag ratio (m) controls velocity movement over steep terrain (Wichmann 2017).

2.2.1 Optimizing model parameters

For regionally applying the runout model we needed to determine the combination of model parameters that result in the best match to our debris-flow inventory. Determining optimal parameters was based on two criteria: (1) the ability of the model to

145   capture our observed runout paths, and (2) its ability to match the observed runout distances. Therefore, we performed this optimization task using a two-stage approach that first optimizes the random walk model and then the PCM model parameters.

A random sample of 100 debris-flow tracks and corresponding source points was used for optimizing the runout models. This sample of the inventory was chosen to facilitate quality control and reduce the computational complexity of the optimization. Source areas were determined by buffering each source point by 50 m and masking away the buffered area that exceeds the

150   runout trimline; this ensures the source area is contained within a mapped debris flow polygon. A sink-filled version of the original DEM was used for the runout modelling.

For each model component, a grid search in parameter space was used to find parameter sets that achieve optimal model performance across all sampled debris flows. The search ranges were similar to Wichmann's (2017) suggested parameter limits for debris flows (see **Table 2** for value ranges). We additionally tested the use of a spatially varying sliding friction coefficient.

155   The value for this spatially varying sliding friction coefficient μ was calculated as a function of catchment area a (km^2) for maximum runout (Gamma 2000; Wichmann and Becht 2006; Mergili et al. 2012; Wichmann 2017):

$$\mu = 0.13a^{-0.25}. \tag{1}$$

Similar to Mergili et al. (2012) and suggested by Wichmann (2017), the spatially varying sliding friction coefficient was set to a maximum of 0.3 and minimum of 0.045. Runout was computed using 1,000 model iterations.

The AUROC was used as a performance measure for the random walk model. Model performance was rated higher if the random walk model contained observed debris-flow tracks within its simulated paths. After optimizing the random walk model, we fixed these parameters for the PCM model, and optimized the sliding friction coefficient and mass-to-drag ratio parameters for determining runout distance. The performance of the PCM model was measured using the relative error of runout distance. Relative error was used so that each debris flow was weighted equally regardless of its magnitude.

Runout distance was measured in terms of horizontal length of the debris-flow track. This distance was measured as the length of a bounding box containing the observed debris-flow track (**Figure 2**; Niculiță 2016; Taylor et al. 2018). Estimated debris-flow tracks were defined as grid cells with values greater than a median runout frequency (**Figure 2**). In this case, the median value represents the most typical simulated debris-flow track. It also provides a conservative estimate of runout distance, which helps mitigate the chance that the optimized model regionally underestimates runout distances.

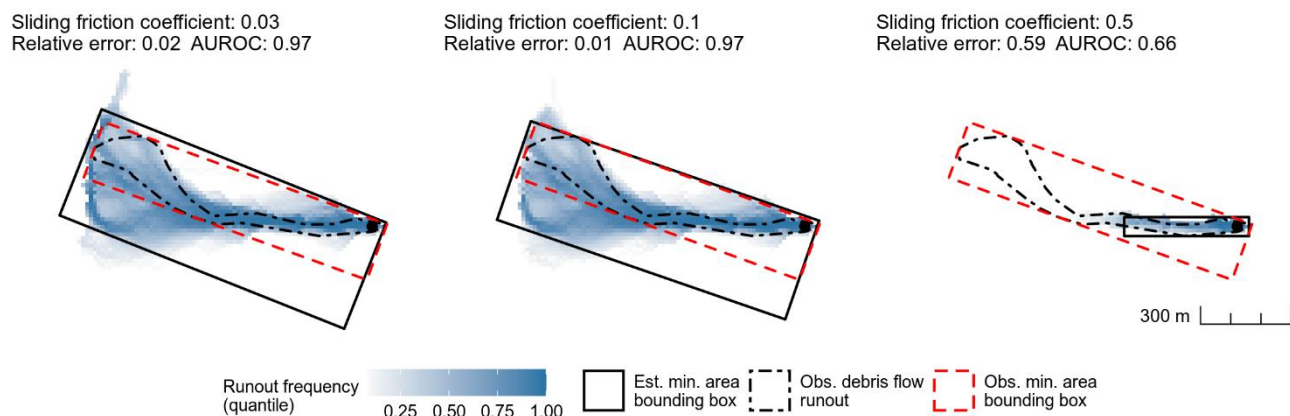



Figure 2: Illustration of runout distance optimization of the sliding friction coefficient using a minimum-area bounding box as a measure of travel distance – mass-drag-ratio fixed at 40 m.

In addition to determining a global optimal parameter set, the best-performing parameters for individual debris-flow events were also explored. We similarly applied a grid search to each event and determined optimal parameter sets based on performances (AUROC and relative error) of each runout model component.

2.2.2 Assessing spatial transferability

We assessed the transferability of optimized runout model (random walk and PCM) parameters by performing 5-fold spatial cross-validation with 1,000 repetitions (**Figure 3**). This approach allows us to explore the sensitivity of grid-search optimized parameter combinations to spatial variation in training and test data. To do this, we observed the frequency of variations in

optimized parameter combinations within all cross-validated iterations. Optimal parameter combinations that occurred more

180 frequently were considered to have a higher degree of transferability; thus, being considered more reliable for application to the entire study area.

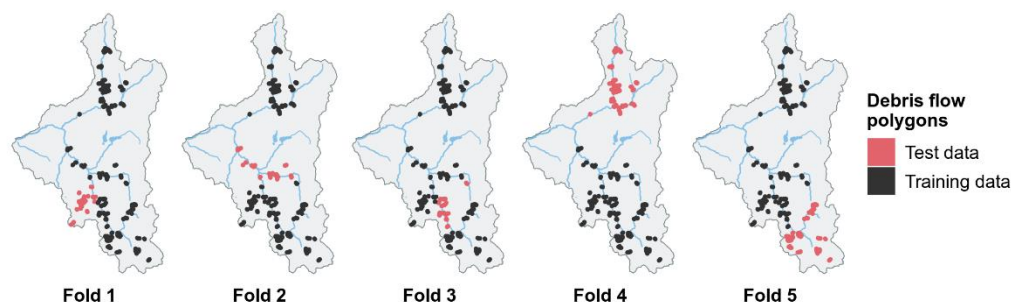



Figure 3: An example realization of random partitions based on k-means clustering of debris-flow polygons for a single repetition of 5-fold spatial cross-validation.

185 We also assessed if there were any spatial patterns in the optimized performance for each model component. That is, were there any spatial trends in model performance that may indicate our model is locally overfitting? We explored for such spatial trends by mapping the distribution of individual debris-flow runout model performances based on the optimized parameters. Additionally, we were concerned if the optimized parameters had a stronger fit to debris flows of a certain magnitude or initiating conditions. Therefore, the potential to overfit to certain debris flow characteristics was assessed by determining

190 Spearman's rank correlation of individual debris-flow performance (for each model component) with observed runout distance and the elevation, catchment area and hillslope angle of the corresponding source points.

2.2.3 Testing sample-size dependence of performance

We explored how runout model parameter selection, performance and robustness were affected by the number of debris flows used for optimization. This analysis was done by applying repeated spatial cross-validation to data sets of varying training

195 sample sizes. To ensure a fair comparison, the size of the test data for each cross-validation iteration was set to 20 debris flows, which is the maximum test sample size when performing 5-fold cross-validation with a sample of 100 debris flows. We tested training samples sizes from 10 to 80. Model performance for each runout modelling components was summarized using the median and IQR (AUROC and relative error). The optimal parameter set for a given sample size were determined as the parameter combinations that were most frequent.

200 2.2.4 Finding a suitable source area prediction threshold

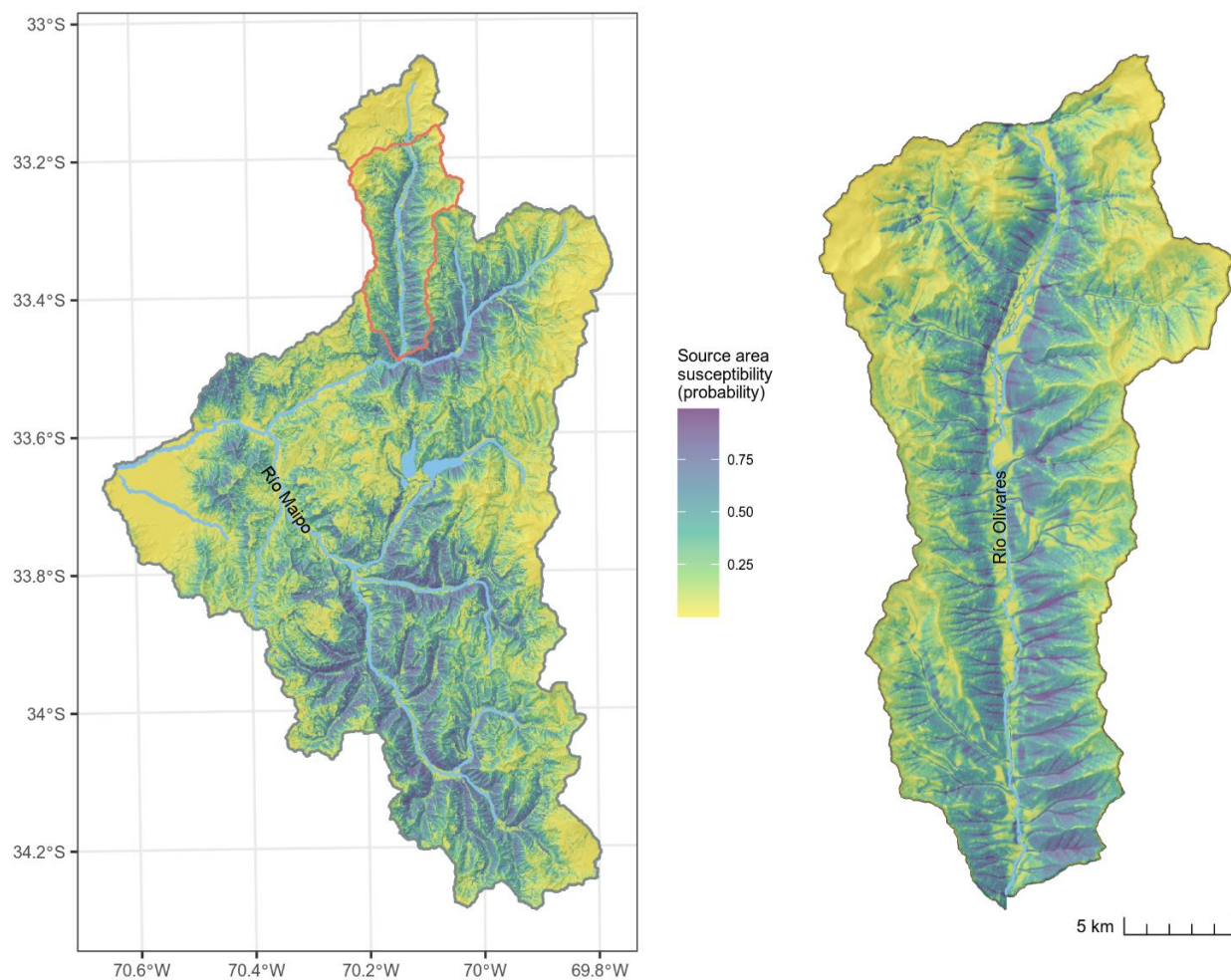
For regionally applying the runout model for susceptibility mapping or exposure analysis, a dichotomous classification of the predicted source areas is required to define the grid cells where simulated debris flows initiate. In the case of basing the source

areas on a susceptibility model, a suitable threshold of the prediction values needs to be selected. In this study, we determine a suitable prediction threshold to classify source areas by searching for the threshold that results in the best-performing runout model for the entire area. Using the optimized model parameters, we tested runout models based on source areas that were delineated using prediction thresholds from 0.5 to 0.95 with a step of 0.05. The performance of each these models was measured using the AUROC. The AUROC was calculated using a sample of 1,000 debris-flow runout locations and 1,000 non-debris flow locations. The source-area prediction threshold resulting in the highest AUROC values was selected for regionally computing a debris-flow runout map for our study area.

2.3 Geocomputing and visualization software

The methods for runout modelling optimization, validation and visualization of the source area prediction and runout modelling were implemented using the open-source statistical software R (ver. 3.6.2; R Core Team 2019) and SAGA GIS (version 6.1; (Conrad et al. 2015) with its Gravitational Process Path model tool (Wichmann 2017). Coupling SAGA GIS with R was done using a combination of the 'RSAGA' (Brenning et al. 2018) and 'Rsagcmd' (Pawley 2019) packages. The GAM was implemented using the 'mgcv' package (Wood 2011). General handling of spatial data in R used the 'sf' (Pebesma 2018), 'sp' (Pebesma and Bivand 2005),'rgeos' (Bivand and Rundel 2019), 'rgdal' (Bivand et al. 2019) and 'raster' (Hijmans 2020) packages; spatial cross-validation was applied using the 'sperrorest' (Brenning 2012) and 'ROCR' (Sing et al. 2005) packages. Parallelization of the optimization and validation procedure used 'foreach' (Microsoft and Weston 2020). Visualization was done using R's 'ggplot2' (Wickham 2009) and 'metR' (Campitelli 2020) packages and ESRI's ArcMap (ver. 10.5).

3 Results
3.1 Source area model performance

The overall performance of the source area prediction based on the GAM was good with a spatially cross-validated median AUROC of 0.80 and an interquartile range (IQR) of 0.001. We found the source-area prediction map was also geomorphologically plausible. Locations most likely to be source areas were within steep terrain associated with channels, gullies and scree slopes. Shallow-flat terrain and areas along ridgelines were modelled as least likely to be source areas (**Figure 4**). This geomorphological knowledge was also expressed in the plots of the GAM spline transformations (**Figure 5**). Relatively steep terrain, slightly concave profile curvature, and areas near faults were modelled as more likely being source areas.

230 **Figure 4: Map of the debris-flow source area prediction.**

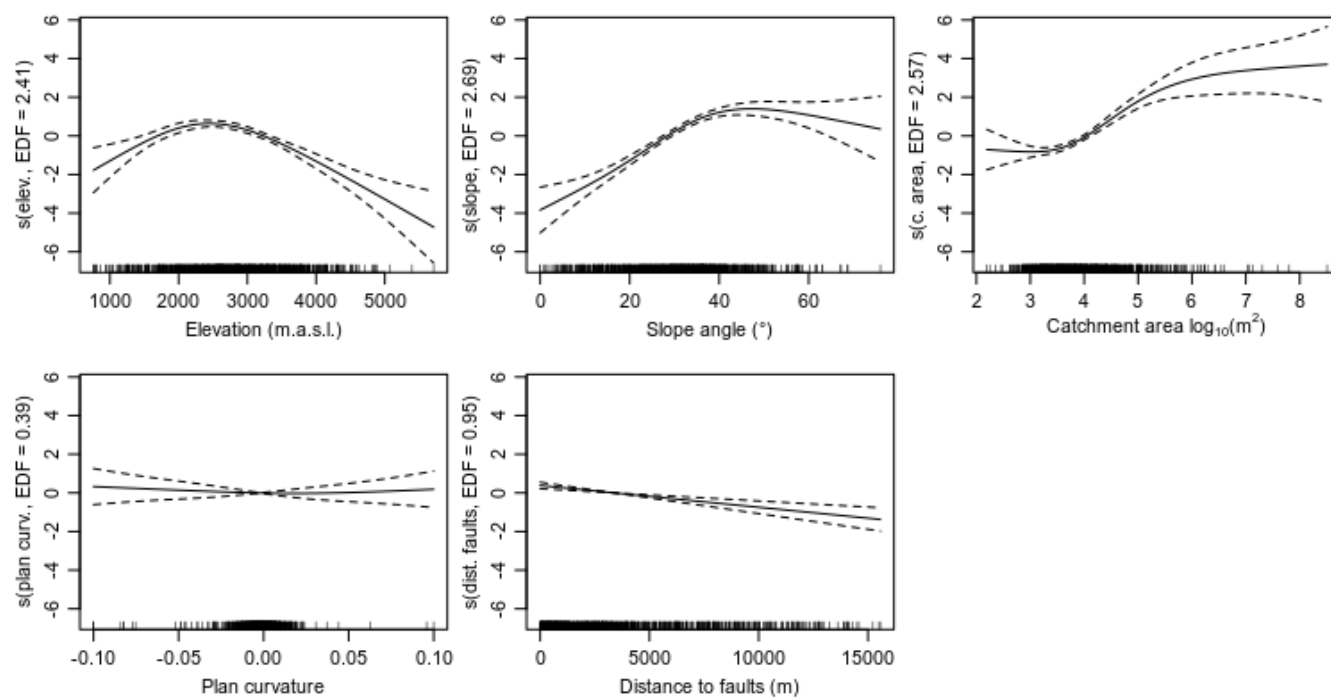



Figure 5: Transformation of predictor variables in the generalized additive model, where the y-axis can be interpreted as the associated likelihood (log-odds) of being a source area. Terms of the form s(predictor) indicate a nonlinear smoothing spline transformation. The effective degrees of freedom (EDF) refer to the flexibility of the smoothers. The dashed lines represent confidence bands at a 95 % level.

3.2 Runout model parameter optimization

The parameter optimization produced runout models with a good spatially cross-validated performance. The optimal parameters for the runout-path model were a slope threshold of 40°, an exponent of divergence of 3.0, and a persistence factor of 1.9 with a median AUROC of 0.94 (IQR = 0.02; **Table 2**). Using these values as plug-in estimates for the PCM runout-distance model component, the optimal sliding friction coefficient and mass-to-drag ratio were 0.11 and 40 m, respectively. The median spatially cross-validated relative length error of the runout-distance model was 0.11, or 11% (IQR = 0.09). We also found that the model based on a global sliding friction parameter estimate performed better than the runout-distance model using a spatially varying sliding friction coefficient (optimal mass-to-drag-ratio = 95 m; **Table 2**), which had a median relative error of 0.19 (IQR 0.09).

Table 2: Runout model grid-search optimization setup and results. Optimization performance was assessed using spatial cross-validation (CV).

Model component	Model Parameter	Grid search value range	Grid search steps	Optimal value	Spatial CV performance
Runout path (Random walk)	Slope threshold	20-40°	2°	40°	**AUROC**
	Exponent of divergence	1.3-3.0	0.17	3.0	Median: 0.94
	Persistence factor	1.5-2.0	0.05	1.9	IQR: 0.02
Runout distance (PCM model)	Sliding friction coefficient	0.04-0.6	0.01	0.11	**Relative error**
					Median: 0.11
	Mass-to-drag ratio	20-150 m	5 m	40 m	IQR: 0.09
Runout distance (PCM model)	Sliding friction coefficient	$0.13a^{-0.25}$	-	-	**Relative error**
					Median: 0.19
	Mass-to-drag ratio	20-150 m	5 m	95 m	IQR: 0.06

250   *Note.* IQR = interquartile range; a = catchment area (km^2)

By visualizing the runout-distance optimization results across grid search space, we can observe model performance and sensitivity to different parameter combinations. In this case, we observed only slight model performance differences for sliding friction coefficients just under 0.2 to 0.04 and mass-to-drag ratios from 20 to 150 m (**Figure 6a**). In general, the values in this band across grid search space would result in good performances with median relative errors ≤ 0.15. However, in terms of

255   controlling the spread of error (**Figure 6b**), sliding coefficients from about 0.05 to 0.15 and mass-to-drag ratios from 20 to 95 m had the lowest IQRs (≤ 0.2). This result illustrates a generally low sensitivity in runout distance model for a wide range of parameter combinations.

Exploring the optimal combination of parameter values using spatial cross-validation provides insights into performance reliability of the optimized model. Given different spatial combinations of testing and training data, we found that the optimal

260   combinations of parameters were associated with high performance values (AUROC = 0.94, relative error = 0.11) and high relative frequencies (61-69%) of occurring in each cross-validation iteration (**Figure 7**). Additionally, the spatially cross-validated optimal parameter sets were generally clustered in grid search space, showing there was little variation in optimized parameters depending on the spatial partitions (**Figure 7**).

265

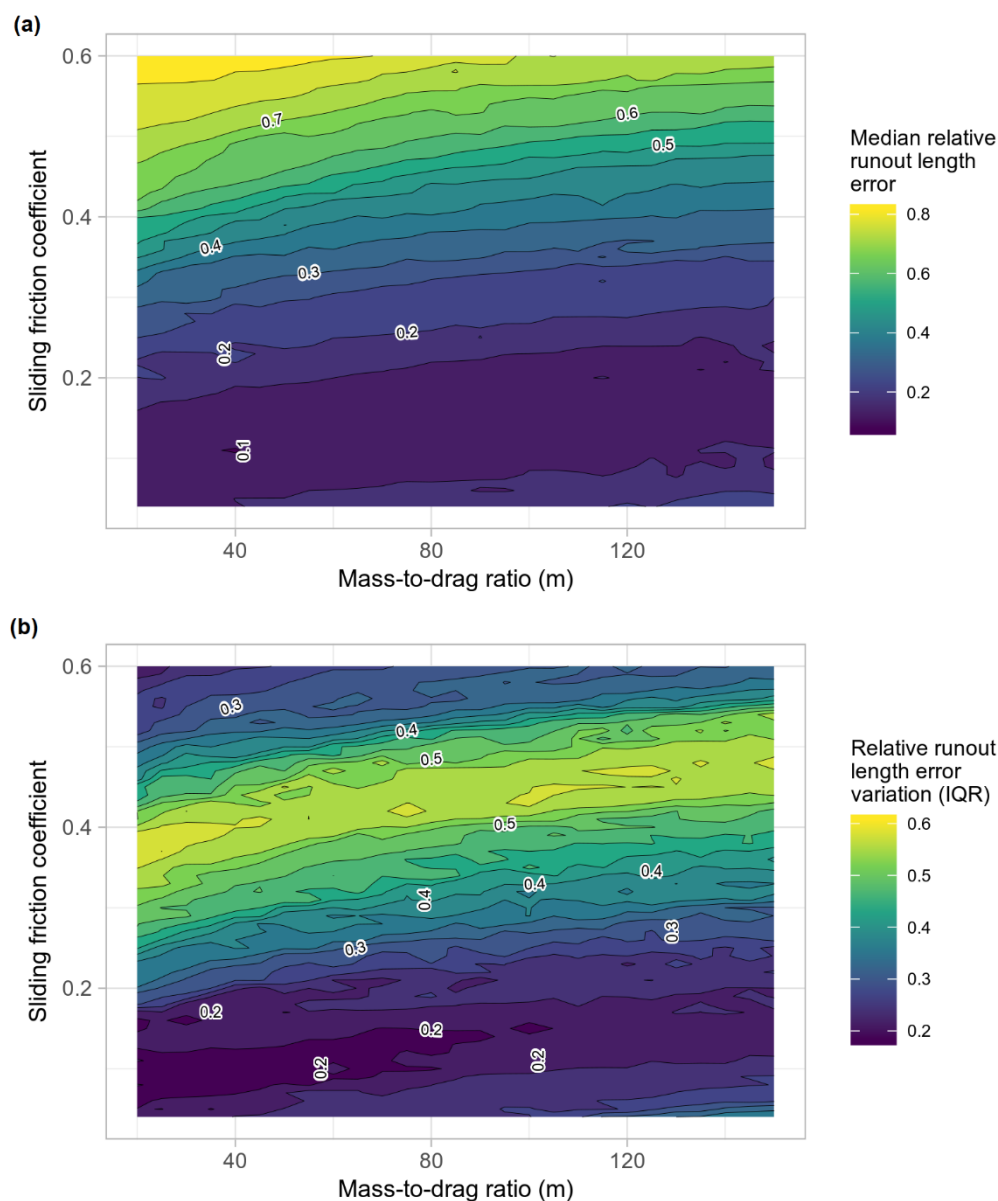



Figure 6: Density contour plots of parameter optimization of sliding friction coefficient and mass-to-drag ratio in the PCM model, illustrating median relative runout length errors (a) and the associated variations in relative error (b) using the interquartile range (IQR).

270

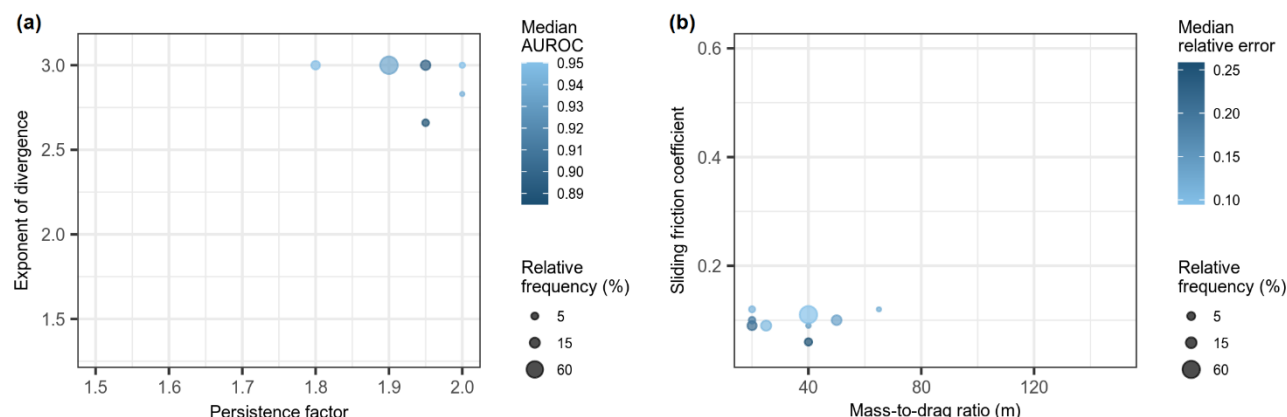



Figure 7: Model performance and frequency of optimal parameters for the runout path (a; given an optimized slope threshold = 40) and runout distance models (b) estimated using 1,000-repeated 5-fold spatial-cross-validation. Relative frequency is the percentage of all repeated spatial-cross validation iterations where a given parameter combination was optimal.

275 **3.3 Thresholding source areas for runout analysis**

The best threshold for delineating source areas from the GAM prediction for runout modelling was 0.7, which results in runout affecting 22% of the study area. This threshold had the peak AUROC value of 0.83. The performance of the runout model drastically decreased with thresholds > 0.8 and gradually decreased towards a thresholds of 0.5: lower thresholds spatially predicted more runout area (**Figure 8**). The resulting runout prediction map (**Figure 9**) appears to be geomorphologically plausible. The locations of the runout deposit areas are similar to what we have observed in the satellite imagery. Additionally,

280 lateral spreading is generally low in narrow gullies and much broader on talus slopes, which is what we would expect for debris flows occurring in our study area.

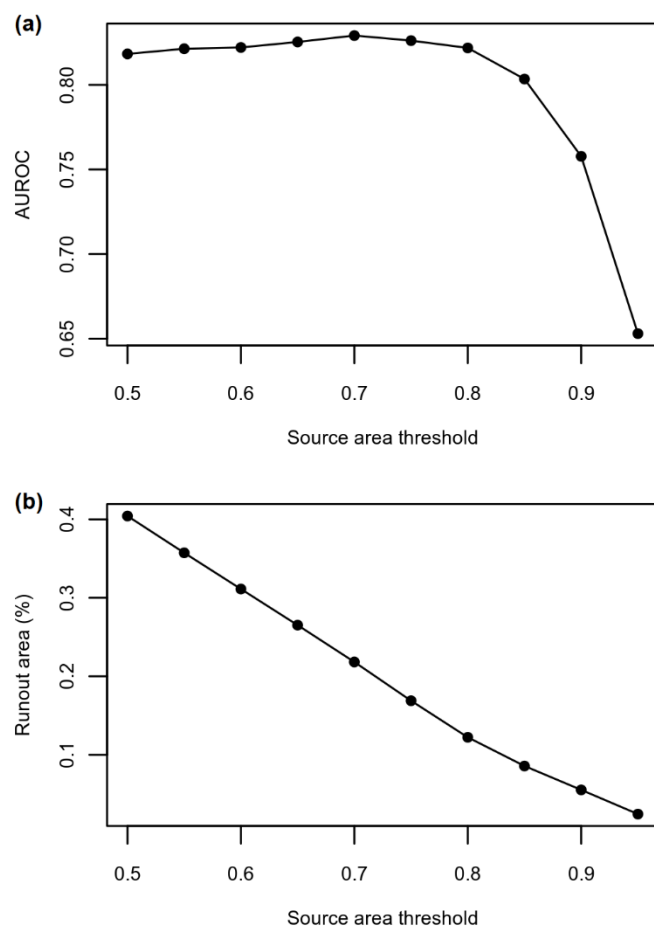



Figure 8: Performance of debris flow runout model for different source area thresholds (a) and percentage of study area impacted by runout (b). Optimal threshold was a threshold of 0.70 (AUROC = 0.83) with runout affecting 22% of the study area.

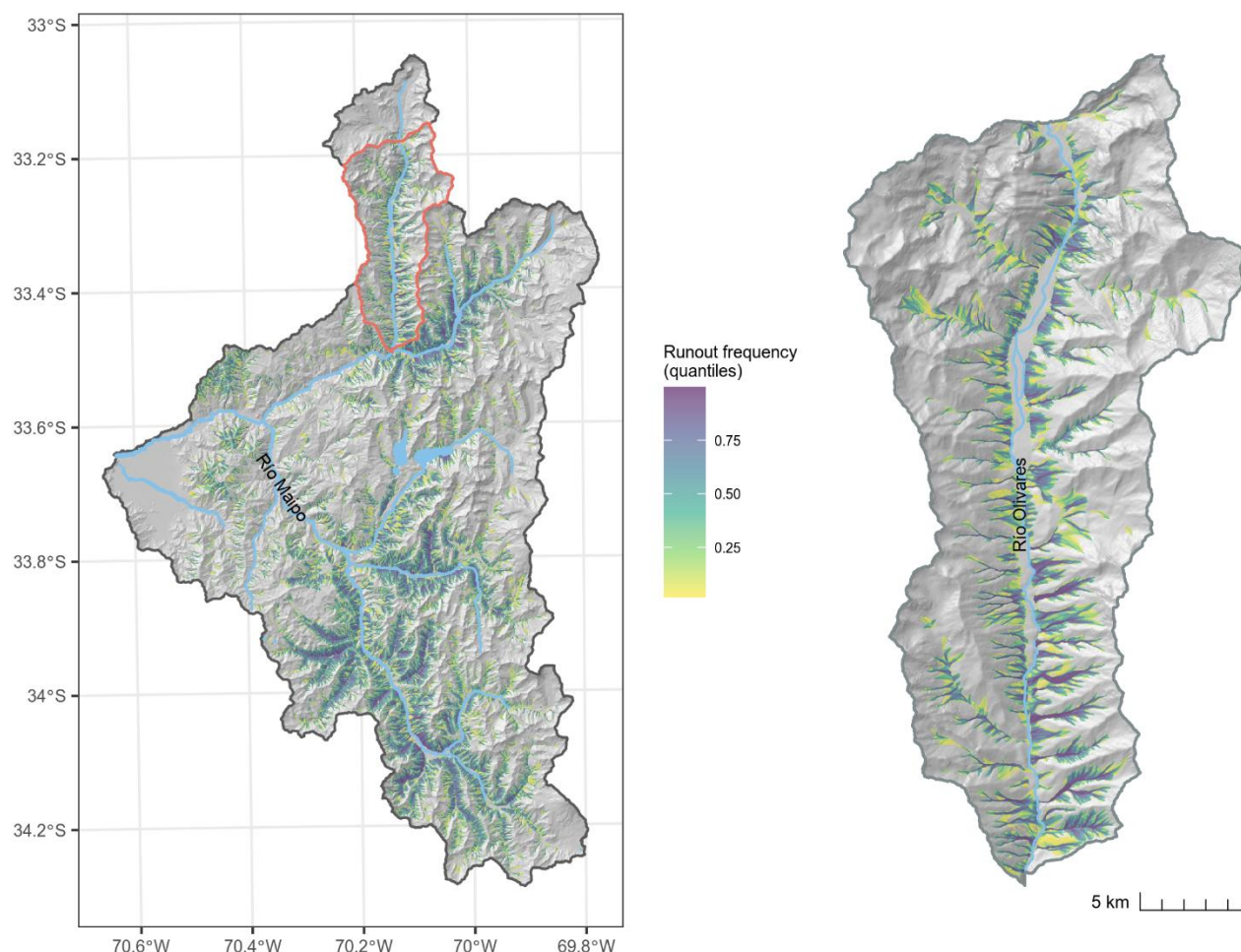



Figure 9: Map of optimized debris-flow runout model based on global parameters and source area threshold of 0.70.

3.4 Exploring patterns in model performance

290   We observed no clear spatial pattern of individual debris-flow performances of the runout model components (**Figure 10**), which is evidence that there was no local overfitting to a particular region of the study area. The distribution of individual AUROC performances of the runout path model were generally high for most of the region, showing that optimization of the path model performs well across the study area. One debris flow was not captured within the runout path model (AUROC < 0.5; **Figure 10**; **Figure 11**). Having a closer look at this occurrence, we found that the modelled runout paths failed to follow

295   the flow direction of the observed debris flow and source point.

The runout distance model had generally more spatial heterogeneity in performance (**Figure 10**), but again no clear spatial patterns in the distribution of relative errors were observed. We investigated the debris-flows that did not perform well for

modelling runout distance (relative error > 0.8) by looking at the satellite imagery used for mapping. It was found that these cases were related to misclassifying stream erosion in relatively shallow and long-hillslope channels as debris flows.

300

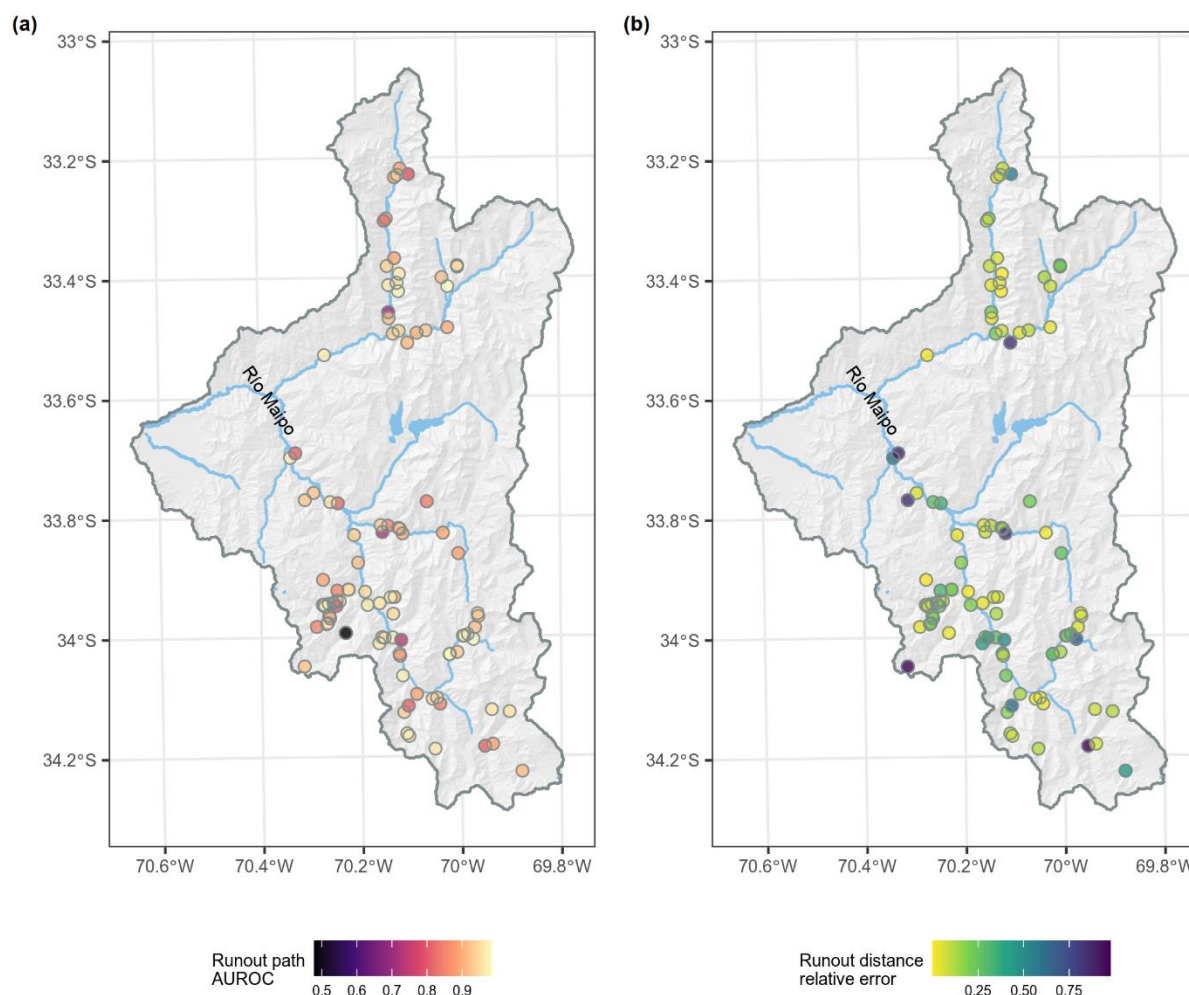



Figure 10: Maps of the performances of the runout path (a) and runout distance models (b) for individual debris-flows based on the global optimal parameters.

Overall, the regionally optimized runout modelled distances fit well with our mapped observations (**Figure 11**). It tended to slightly overestimate debris-flow travel lengths with a median runout distance error of 16.6 m (**Figure 11**). This error is just
305 slightly over a single grid cell size of the DEM. The highest runout distance errors were due to misclassified debris-flows, as previously mentioned. The optimization of the runout model avoided overfitting to higher or smaller runout distance and elevations of source points. That is, we observed no clear relationship between runout model optimization performance to the observed runout length and the elevation, catchment area or hillslope angle of the training source points (**Figure 11**). Spearman
correlations of estimated and observed runout distance were -0.36 and -0.17 for runout path (AUROC) and runout distance
310 (relative error model) performance, respectively. The runout model error correlations with elevation, catchment area, and
hillslope angle of the source points were -0.21, 0.11 and 0.29 respectively.

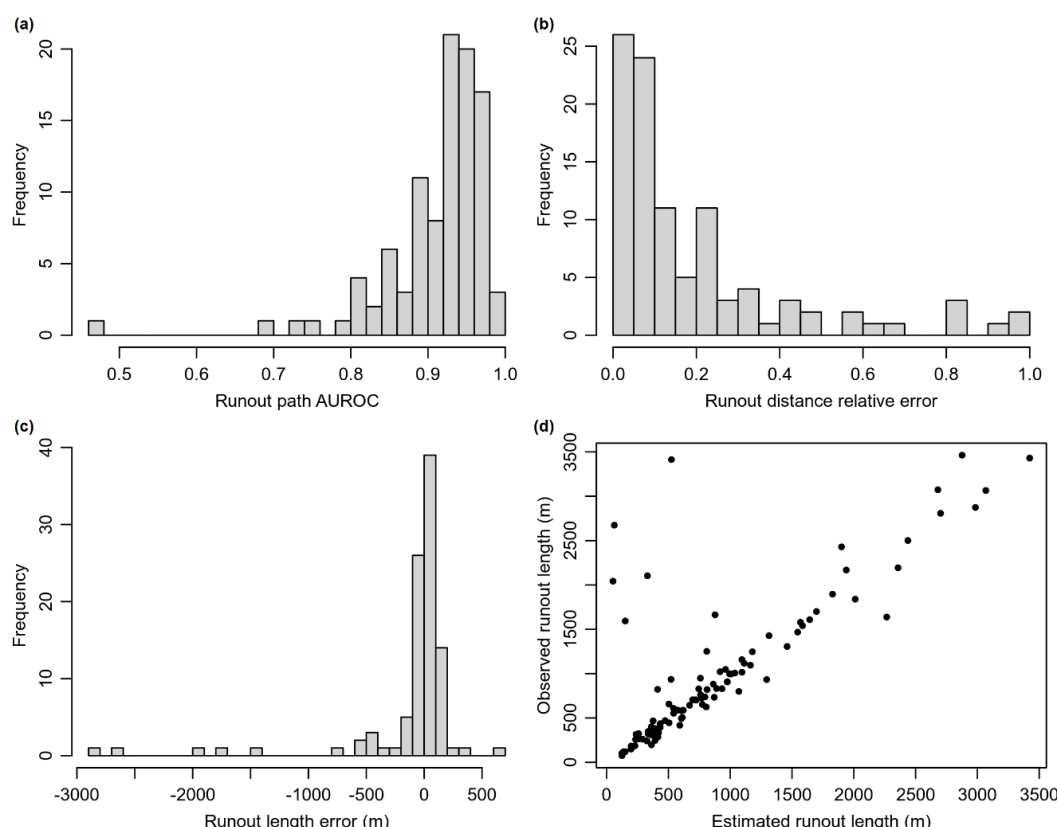



Figure 11: Histogram of performance of the runout path (a) runout distance models based on relative error (b) and actual runout
315 **length error (c), and a plot of observed runout lengths versus modelled runout lengths (d).**

3.5 Optimization for individual debris-flows

The optimal parameters for individual debris-flows were also computed for general comparison to the regional model. We
found that the optimized runout model parameters were highly variable for individual debris flows (**Figure 12**). The relative
errors were low (median = 0.002, IQR = 0.007), except for a few debris flows that failed to optimize (**Figure 12**). There was
320 no clear spatial pattern in optimal sliding friction coefficient and mass-to-drag ratio parameter combinations across the study
area, which shows that runout characteristics were quite diverse for individual debris-flows across this large study area (**Figure**

13). We also observed no strong relationships between terrain attributes and optimal parameters. Spearman's correlations of sliding friction coefficient and mass-to-drag ratio with elevation, slope and catchment area were ≤|0.29|.

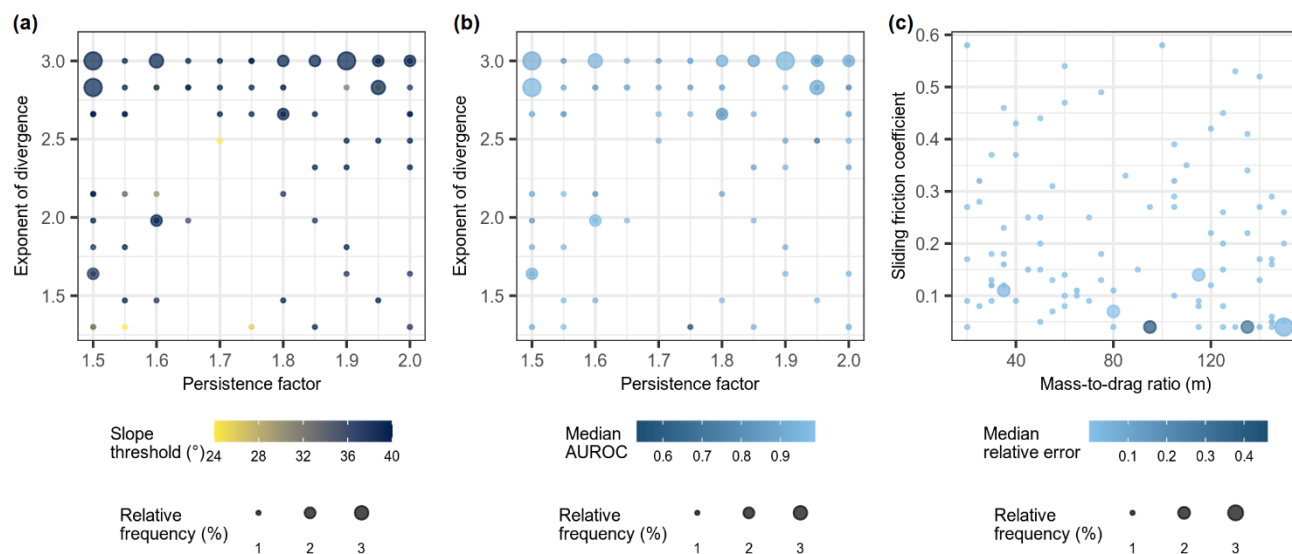



Figure 12: **Performance and frequency of runout path (a, b) and distance (c) optimal parameters determined for individual debris flows. Relative frequency is the percentage of all repeated spatial-cross validation iterations where a given parameter combination was optimal.**

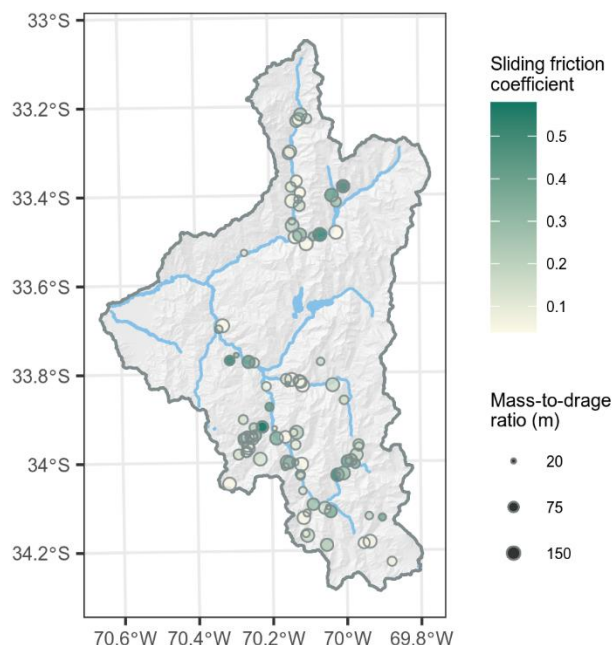
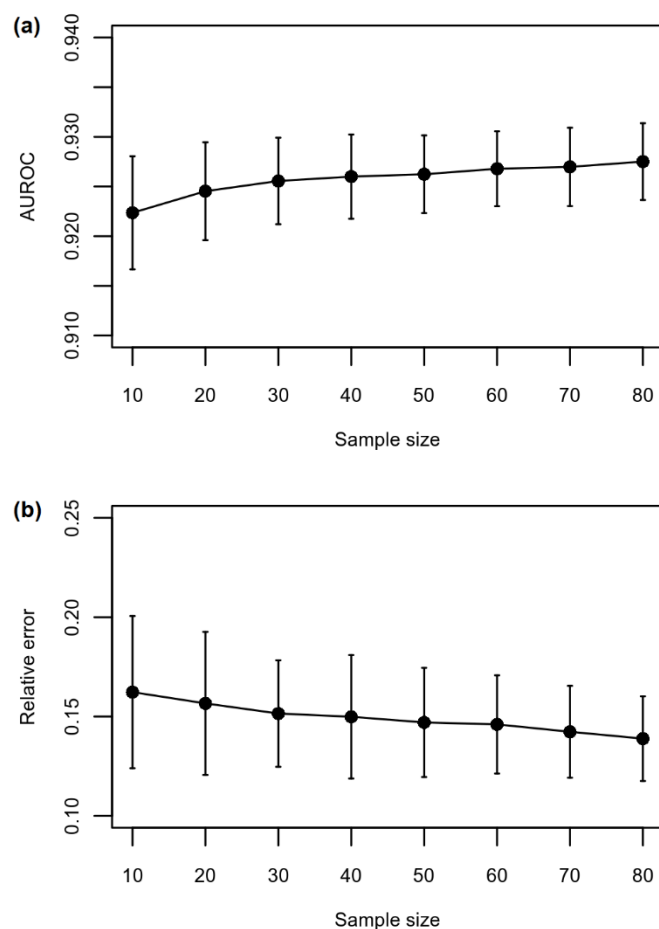



Figure 13: **Map of runout distance model optimal parameters determined for individual debris flows.**

330

3.6 Runout model robustness to sample size

Runout model performance and variation tends to improve slightly when using larger sample sizes (**Figure 14**). Large sample sizes also resulted in more consistently selected model parameters across spatial cross-validation iterations (**Figure 15**). The larger spread across grid search space of the relative frequency of optimal parameter combinations for smaller sample sizes

335 illustrates that we may be less likely to find the best model parameters that generalize well for a large region (**Figure 15**). In general, as we increased sample size, we reduced sampling variability and narrowed the number of optimal combinations of parameters in grid search space, meaning we became more confident that the selected parameters would transfer better to estimate runout distance in adjacent areas.

340 **Figure 14: Comparison of runout path (a) and distance (b) performances for different model training samples sizes. The error bars indicate the standard deviation in performances.**

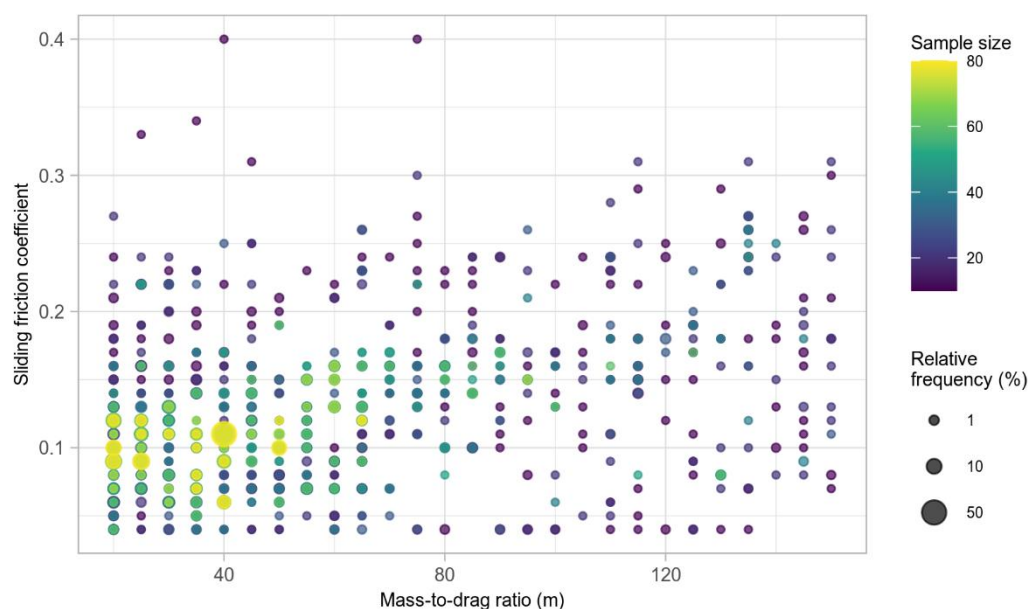



Figure 15: Relative frequencies of optimal parameter combinations for different training sample sizes computed using spatial cross-validation.

4 Discussion

4.1 Model performance and transferability

Assessing the spatial transferability of runout models is essential when extending their use from single or local events to regionally modelling runout across unknown space. Overall, this study demonstrates that our novel optimization approach performed well at regionally modelling the spatial distribution of runout path and distances across the upper Maipo river valley basin. A key component of the success of our modelling approach was its ability to generalize. The transferability of a regional runout model can be affected by the generalization ability of both the source area prediction and the optimized process-based runout model.

4.1.1 Source area modelling

In the development of our source area prediction model, we aimed to produce a simplified empirical model that would result in good performance and transferability. Wenger and Olden (2012) recommended to control the flexibility of the GAM smoothing parameters to avoid overfitting that would lead to poorer model transferability. Therefore, similarly to previous landslide susceptibility studies using the GAM (Goetz et al. 2011; Goetz et al. 2015b; Bordoni et al. 2020) we limited the degrees of freedom for smoothing-spline fitting.

A detailed model based on a large set of predictors can impede its ability to transfer to other locations (Tuanmu et al. 2011). As we and others have demonstrated, good predictive performance of susceptibility models of debris flows can be achieved

with a relatively small set of predictors, which are primarily DEM-derived terrain attributes (Blahut et al. 2010b; Heckmann et al. 2014; Goetz et al. 2015a). Additionally, by fitting our models with multi-temporal data, we may be more likely to achieve better transferability in time (Tuanmu et al. 2011; Knevels et al. 2020). Event-specific inventories may not be large enough for regional optimization and risk the potential of overfitting source conditions spatially varying conditions of that event (e.g.,

365   precipitation and snowmelt patterns). A smaller sample may also lead to not capturing the range of terrain conditions across the study area required for robust empirical modelling of source area locations (Petschko et al. 2014; Rudy et al. 2016).

4.1.2 Runout analysis

By optimizing the runout distance model using the median relative error as a metric, we managed to reduce the impact of possible outliers in our training and test data. We additionally reduced our chances of overfitting the regional model to larger

370   debris flow events, which was crucial for a model to maintain a generalization that makes it transferable across large areas.

Plotting the individual optimized models and exploring correlations between terrain attributes and optimal parameters allowed us to see if there were any broad trends in the parameter selection. In our study, we observed high variability in optimal parameters the PCM model. While applying a trial and error approach, Mergili et al. (2012) also observed different optimal parameter combinations when modelling runout for a couple of debris flows just north of our study area. It seems apparent

375   that determining runout parameter values for regional modelling without an optimization procedure would be difficult. It is also no surprise that the errors of the individually optimized debris-flows were low (**Figure 12**). Whether meaningful or not, the optimization approach should be able to match the runout distance with very low error. Poor individually optimized events could be attributed to locally poor DEM quality and/or complex or ambiguous events. It was interesting that despite the variety in optimal parameters, we were still able to produce a good spatially distributed model of runout patterns using a global set of

380   parameters. This may indicate that the combination of random walk and the process based PCM model dictates a general runout pattern that is insensitive to values within a broad and nearly optimal range of physically reasonable parameters. For regional applications, such models that are already well adapted to generalize runout patterns are likely to transfer well to unknown locations, as we observed in this study by assessing runout model transferability.

4.2 Regionally optimizing runout models

385   As with any optimization problem, using a suitable cost function is critical to ensure the model parameters are optimized to solve a very a specific problem. It can be difficult to define a single metric that simultaneously measures both performance of path and distance simulations. As shown in this study, it may also not be necessary. By using the two-stage approach to optimization, where we first optimize the runout path model and then plugin those values to optimize the runout-distance model component, we can considerably reduce computational complexity, while ensuring that our final model result explicitly

390   optimizes for runout distance – a key characteristic for spatially predicting areas potentially impacted by debris flow runout. In our case, we only needed to solve two separate problems with 3 (random walk) and 2 (PCM) unknown parameters, as opposed to solving simultaneously for 5 unknowns. We used an exhaustive grid search to optimize the runout model

components because it allowed us to visualize and assess model performance across parameter space. However, a drawback of this optimization method is that it can be computationally slow to explore all candidate parameter combinations. If speed is a requirement for regional runout analysis, then a faster method like random search (Bergstra and Bengio 2012) may be preferred (Schratz et al. 2019).

In terms of model improvement, analysis of the spatial distribution of optimal parameters may lead to better parameter optimization based on terrain or geological characteristics. The spatially varying values used in this study were based on modelling for alpine regions (Gamma 2000), and do not account for the potential variability in sliding conditions between different catchments (Guthrie 2002). This challenge may be overcome by using regional modelling strategies already applied for landslide initiation susceptibility modelling, where the study area is divided into geologically similar regions, runout model optimization is performed for each region and then combined (e.g., Petschko et al. 2014).

5 Conclusions

Modelling the spatial pattern of debris flow runout in large, mainly remote areas, with sparse data, requires model calibration and validation methods that ensure spatial transferability. In this study, we demonstrated that our integrated two-stage optimization approach is one such method suitable for regionally simulating debris-flow runout patterns. Through this analysis, we observed a general insensitivity in runout distance performance of the PCM model to a range of parameters, which is an indication of the model's inherent capability to characterize general runout behaviour across our study area. Future improvements to our approach may include building a model that allows for more spatial variation in optimized parameters, especially in regions with better availability of soil physical information. Overall, our open-source modelling approach, which enhanced the GPP model by adding support for automated model calibration and transferability assessment, provides an accessible and extendible modelling framework for such advances in regional runout modelling.

Code availability

To go with this paper, we developed the runoptGPP R package for optimizing mass movement runout models using the random walk and PCM model components of the GPP tool in SAGA-GIS. It is available for download from the GitHub repository: https://github.com/jngtz/runoptGPP (DOI: 10.5281/zenodo.4428050). This package contains tutorials on how to apply the optimization methods regionally or for single runout events. The GitHub repository also contains the R code used to conduct our analysis.

Data availability

The data is available for download from Zenodo https://doi.org/10.5281/zenodo.4428080.

Acknowledgements

This research was funded by CETAQUA Chile on behalf of Aguas Andinas from the project titled "Mass movement processes in the upper Maipo basin: Modelling the susceptibility and probable sediment transfers", which was done in collaboration between the Friedrich Schiller University Jena and the Institute of Geography Pontificia Universidad Católica de Chile.

425 **References**

Aaron, Jordan; McDougall, Scott; Nolde, Natalia (2019): Two methodologies to calibrate landslide runout models. In *Landslides* 16 (5), pp. 907–920. DOI: 10.1007/s10346-018-1116-8.

Aleotti, P.; Chowdhury, R. (1999): Landslide hazard assessment: summary review and new perspectives. In *Bull Eng Geol Env* 58 (1), pp. 21–44. DOI: 10.1007/s100640050066.

430 Angillieri, María Yanina Esper (2020): Debris flow susceptibility mapping using frequency ratio and seed cells, in a portion of a mountain international route, Dry Central Andes of Argentina. In *CATENA* 189, p. 104504. DOI: 10.1016/j.catena.2020.104504.

ASF DAAC: ALOS PALSAR Radiometric Terrain Corrected high resolution digital elevation model. Includes Material ©JAXA/METI (2011). DOI: 10.5067/Z97HFCNKR6VA.

435 Bergstra, James; Bengio, Yoshua (2012): Random search for hyper-parameter optimization. In *Journal of Machine Learning Research* 13 (1), pp. 281–305.

Bivand, R.; Keitt T.; Rowlingson B. (2019): rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-8. https://CRAN.R-project.org/package=rgdal.

Bivand, R.; Rundel, C. (2019): rgeos: Interface to Geometry Engine - Open Source ('GEOS'). R package version 0.5-2.
440 https://CRAN.R-project.org/package=rgeos.

Blahut, J.; Horton, P.; Sterlacchini, S.; Jaboyedoff, M. (2010a): Debris flow hazard modelling on medium scale: Valtellina di Tirano, Italy. In *Nat. Hazards Earth Syst. Sci.* 10 (11), pp. 2379–2390. DOI: 10.5194/nhess-10-2379-2010.

Blahut, Jan; van Westen, Cees J.; Sterlacchini, Simone (2010b): Analysis of landslide inventories for accurate prediction of debris-flow source areas. In *Geomorphology* 119 (1-2), pp. 36–51. DOI: 10.1016/j.geomorph.2010.02.017.

445 Bordoni, Massimiliano; Galanti, Yuri; Bartelletti, Carlotta; Persichillo, Maria Giuseppina; Barsanti, Michele; Giannecchini, Roberto et al. (2020): The influence of the inventory on the determination of the rainfall-induced shallow landslides susceptibility using generalized additive models. In *CATENA* 193, p. 104630. DOI: 10.1016/j.catena.2020.104630.

Brenning, A. (2005): Spatial prediction models for landslide hazards: review, comparison and evaluation. In *Nat. Hazards Earth Syst. Sci.* 5 (6), pp. 853–862. DOI: 10.5194/nhess-5-853-2005.

450 Brenning, A.; Bangsm D.; Becker, M. (2018): RSAGA: SAGA Geoprocessing and Terrain Analysis. R package version 1.3.0 https://CRAN.R-project.org/package=RSAGA.

Brenning, Alexander (2012): Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. In : 2012 IEEE International Geoscience and Remote Sensing Symposium. IGARSS 2012 - 2012 IEEE International Geoscience and Remote Sensing Symposium. Munich, Germany, 2012-07-22 - 2012-07-27: IEEE, 455 pp. 5372–5375.

Brock, Jonas; Schratz, Patrick; Petschko, Helene; Muenchow, Jannes; Micu, Mihai; Brenning, Alexander (2020): The performance of landslide susceptibility models critically depends on the quality of digital elevation models. In *Geomatics, Natural Hazards and Risk* 11 (1), pp. 1075–1092. DOI: 10.1080/19475705.2020.1776403.

Cama, M.; Lombardo, L.; Conoscenti, C.; Rotigliano, E. (2017): Improving transferability strategies for debris flow 460 susceptibility assessment: Application to the Saponara and Itala catchments (Messina, Italy). In *Geomorphology* 288, pp. 52–65. DOI: 10.1016/j.geomorph.2017.03.025.

Campitelli, E. (2020): metR: Tools for Easier Analysis of Meteorological Fields. R package version 0.7.0 https://CRAN.R-project.org/package=metR.

Carrara, A.; Guzzetti, F.; Cardinali, M.; Reichenbach, P. (1999): Use of GIS technology in the prediction and monitoring of 465 landslide hazard. In *Nat Hazards* 20 (2/3), pp. 117–135. DOI: 10.1023/A:1008097111310.

Cepeda, José; Chávez, José Alexander; Cruz Martínez, Celina (2010): Procedure for the selection of runout model parameters from landslide back-analyses: application to the Metropolitan Area of San Salvador, El Salvador. In *Landslides* 7 (2), pp. 105–116. DOI: 10.1007/s10346-010-0197-9.

Chung; Fabbri (1999): Probabilistic prediction models for landslide hazard mapping. In *Photogrammetric Engineering & 470 Remote Sensing* 65 (12), pp. 1389–1399.

Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L. et al. (2015): System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. In *Geosci. Model Dev.* 8 (7), pp. 1991–2007. DOI: 10.5194/gmd-8-1991-2015.

Dong, Jia-Jyun; Lee, Chyi-Tyi; Tung, Yu-Hsiang; Liu, Chia-Nan; Lin, Kuang-Ping; Lee, Jiin-Fa (2009): The role of the sediment budget in understanding debris flow susceptibility. In *Earth Surf. Process. Landforms* 34 (12), pp. 1612–1624. DOI: 475 10.1002/esp.1850.

Fabbri, Andrea G.; Chung, Chang-Jo F.; Cendrero, Antonio; Remondo, Juan (2003): Is Prediction of Future Landslides Possible with a GIS? In *Nat Hazards* 30 (3), pp. 487–503. DOI: 10.1023/B:NHAZ.0000007282.62071.75.

Galas, S.; Dalbey, K.; Kumar, D.; Patra, A.; Sheridan, M. (2007): Benchmarking TITAN2D mass flow model against a sand flow experiment and the 1903 Frank Slide. In Ken Ho, Victor Li (Eds.): The 2007 International Forum on Landslide Disaster management. Proceedings of the 2007 International Forum on Landslide Disaster Management, Hong Kong, 10-12 December 2007 / edited by Ken Ho, Victor Li. Hong Kong: Geotechnical Division, pp. 899–918.

Gamma, P. (2000): dfwalk – Ein Murgang-Simulationsprogramm zur Gefahrenzonierung. In *Geographica Bernensia, University of Bern, Switzerland* G66, 14pp.

Goetz, J. N.; Brenning, A.; Petschko, H.; Leopold, P. (2015a): Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. In *Computers & Geosciences* 81, pp. 1–11. DOI: 10.1016/j.cageo.2015.04.007.

Goetz, J. N.; Guthrie, R. H.; Brenning, A. (2015b): Forest harvesting is associated with increased landslide activity during an extreme rainstorm on Vancouver Island, Canada. In *Nat. Hazards Earth Syst. Sci.* 15 (6), pp. 1311–1330. DOI: 10.5194/nhess-15-1311-2015.

Goetz, Jason N.; Guthrie, Richard H.; Brenning, Alexander (2011): Integrating physical and empirical landslide susceptibility models using generalized additive models. In *Geomorphology* 129 (3-4), pp. 376–386. DOI: 10.1016/j.geomorph.2011.03.001.

Guthrie, R. H. (2002): The effects of logging on frequency and distribution of landslides in three watersheds on Vancouver Island, British Columbia. In *Geomorphology* 43 (3-4), pp. 273–292. DOI: 10.1016/S0169-555X(01)00138-6.

Guthrie, Richard H.; Deadman, Peter J.; Cabrera, A. Raymond; Evans, Stephen G. (2008): Exploring the magnitude–frequency distribution: a cellular automata model for landslides. In *Landslides* 5 (1), pp. 151–159. DOI: 10.1007/s10346-007-0104-1.

Hauser, A. (2002): Rock avalanche and resulting debris flow in Estero Parraguirre and Río Colorado, Region Metropolitana, Chile. In : Catastrophic Landslides: Effects, Occurrence, and Mechanisms, vol. 15: Geological Society of America (Reviews in Engineering Geology), pp. 135–148.

Heckmann, T.; Gegg, K.; Gegg, A.; Becht, M. (2014): Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. In *Nat. Hazards Earth Syst. Sci.* 14 (2), pp. 259–278. DOI: 10.5194/nhess-14-259-2014.

Heckmann, Tobias; Schwanghart, Wolfgang (2013): Geomorphic coupling and sediment connectivity in an alpine catchment — Exploring sediment cascades using graph theory. In *Geomorphology* 182, pp. 89–103. DOI: 10.1016/j.geomorph.2012.10.033.

505 Hijmans, R. J. (2020): raster: Geographic Data Analysis and Modeling. R package version 3.0-12 https://CRAN.R-project.org/package=raster.

Holmgren, Peter (1994): Multiple flow direction algorithms for runoff modelling in grid based elevation models: An empirical evaluation. In *Hydrol. Process.* 8 (4), pp. 327–334. DOI: 10.1002/hyp.3360080405.

Horton, P.; Jaboyedoff, M.; Rudaz, B.; Zimmermann, M. (2013): Flow-R, a model for susceptibility mapping of debris flows
510 and other gravitational hazards at a regional scale. In *Nat. Hazards Earth Syst. Sci.* 13 (4), pp. 869–885. DOI: 10.5194/nhess-13-869-2013.

Hungr, Oldrich (1995): A model for the runout analysis of rapid flow slides, debris flows, and avalanches. In *Can. Geotech. J.* 32 (4), pp. 610–623. DOI: 10.1139/t95-063.

Knevels, Raphael; Petschko, Helene; Proske, Herwig; Leopold, Philip; Maraun, Douglas; Brenning, Alexander (2020): Event-
515 Based Landslide Modeling in the Styrian Basin, Austria: Accounting for Time-Varying Rainfall and Land Cover. In *Geosciences* 10 (6), p. 217. DOI: 10.3390/geosciences10060217.

Lombardo, L.; Cama, M.; Maerker, M.; Rotigliano, E. (2014): A test of transferability for landslides susceptibility models under extreme climatic events: application to the Messina 2009 disaster. In *Nat Hazards* 74 (3), pp. 1951–1989. DOI: 10.1007/s11069-014-1285-2.

520 Lombardo, Luigi; Opitz, Thomas; Huser, Raphaël (2018): Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. In *Stoch Environ Res Risk Assess* 32 (7), pp. 2179–2198. DOI: 10.1007/s00477-018-1518-0.

Marchi, Lorenzo; D'Agostino, Vincenzo (2004): Estimation of debris-flow magnitude in the Eastern Italian Alps. In *Earth Surf. Process. Landforms* 29 (2), pp. 207–220. DOI: 10.1002/esp.1027.

525 McDougall, Scott (2017): 2014 Canadian Geotechnical Colloquium: Landslide runout analysis — current practice and challenges. In *Can. Geotech. J.* 54 (5), pp. 605–620. DOI: 10.1139/cgj-2016-0104.

Mergili, M.; Krenn, J.; Chu, H.-J. (2015): r.randomwalk v1, a multi-functional conceptual tool for mass movement routing. In *Geosci. Model Dev.* 8 (12), pp. 4027–4043. DOI: 10.5194/gmd-8-4027-2015.

Mergili, Martin; Fellin, Wolfgang; Moreiras, Stella M.; Stötter, Johann (2012): Simulation of debris flows in the Central Andes
530 based on Open Source GIS: possibilities, limitations, and parameter sensitivity. In *Nat Hazards* 61 (3), pp. 1051–1081. DOI: 10.1007/s11069-011-9965-7.

Mergili, Martin; Schwarz, Leonhard; Kociu, Arben (2019): Combining release and runout in statistical landslide susceptibility modeling. In *Landslides* 16 (11), pp. 2151–2165. DOI: 10.1007/s10346-019-01222-7.

Microsoft; Weston, S. (2020): foreach: Provides Foreach Looping Construct. R package version 1.4.8. https://CRAN.R-project.org/package=foreach.

Moreiras, Stella M.; Sepúlveda, Sergio A. (2015): Megalandslides in the Andes of central Chile and Argentina (32°–34°S) and potential hazards. In *Geological Society, London, Special Publications* 399 (1), pp. 329–344. DOI: 10.1144/SP399.18.

Niculiţă, Mihai (2016): Automatic landslide length and width estimation based on the geometric processing of the bounding box and the geomorphometric analysis of DEMs. In *Nat. Hazards Earth Syst. Sci.* 16 (8), pp. 2021–2030. DOI: 10.5194/nhess-16-2021-2016.

Pawley, S. (2019): Rsagacmd: Linking R with the Open-Source 'SAGA-GIS' Software. R package version 0.0.2 https://CRAN.R-project.org/package=Rsagacmd.

Pebesma, E.; Bivand, R. (2005): Classes and methods for spatial data in R. In *R News* 5 (2). Available online at https://cran.r-project.org/doc/Rnews/.

Pebesma, Edzer (2018): Simple Features for R: Standardized Support for Spatial Vector Data. In *The R Journal* 10 (1), p. 439. DOI: 10.32614/RJ-2018-009.

Perla, R.; Cheng, T. T.; McClung, D. M. (1980): A Two–Parameter Model of Snow–Avalanche Motion. In *J. Glaciol.* 26 (94), pp. 197–207. DOI: 10.3189/S002214300001073X.

Petschko, H.; Brenning, A.; Bell, R.; Goetz, J.; Glade, T. (2014): Assessing the quality of landslide susceptibility maps – case study Lower Austria. In *Nat. Hazards Earth Syst. Sci.* 14 (1), pp. 95–118. DOI: 10.5194/nhess-14-95-2014.

R Core Team (2019): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org/.

Rudy, Ashley C.A.; Lamoureux, Scott F.; Treitz, Paul; van Ewijk, Karin Y. (2016): Transferability of regional permafrost disturbance susceptibility modelling using generalized linear and generalized additive models. In *Geomorphology* 264, pp. 95–108. DOI: 10.1016/j.geomorph.2016.04.011.

Ruß, Georg; Brenning, Alexander (2010): Data Mining in Precision Agriculture: Management of Spatial Information. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell et al. (Eds.): Computational Intelligence for Knowledge-Based Systems Design, vol. 6178. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 350–359.

560   Schraml, K.; Thomschitz, B.; McArdell, B. W.; Graf, C.; Kaitna, R. (2015): Modeling debris-flow runout patterns on two alpine fans with different dynamic simulation models. In *Nat. Hazards Earth Syst. Sci.* 15 (7), pp. 1483–1492. DOI: 10.5194/nhess-15-1483-2015.

Schratz, Patrick; Muenchow, Jannes; Iturritxa, Eugenia; Richter, Jakob; Brenning, Alexander (2019): Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. In *Ecological Modelling* 406, 565   pp. 109–120. DOI: 10.1016/j.ecolmodel.2019.06.002.

Sepúlveda, Sergio A.; Moreiras, Stella M.; Lara, Marisol; Alfaro, Alejandro (2015): Debris flows in the Andean ranges of central Chile and Argentina triggered by 2013 summer storms: characteristics and consequences. In *Landslides* 12 (1), pp. 115–133. DOI: 10.1007/s10346-014-0539-0.

Servicio Nacional de Geología y Minería de Chile: Geología de Chile Escala 1:1,000,000 (v.1.0), Publicacion Geologica 570   Digital, No. 4.

Sing, Tobias; Sander, Oliver; Beerenwinkel, Niko; Lengauer, Thomas (2005): ROCR: visualizing classifier performance in R. In *Bioinformatics (Oxford, England)* 21 (20), pp. 3940–3941. DOI: 10.1093/bioinformatics/bti623.

Stevenson, John A.; Sun, Xianfang; Mitchell, Neil C. (2010): Despeckling SRTM and other topographic data with a denoising algorithm. In *Geomorphology* 114 (3), pp. 238–252. DOI: 10.1016/j.geomorph.2009.07.006.

575   Sun, Xianfang; Rosin, Paul; Martin, Ralph; Langbein, Frank (2007): Fast and effective feature-preserving mesh denoising. In *IEEE transactions on visualization and computer graphics* 13 (5), pp. 925–938. DOI: 10.1109/TVCG.2007.1065.

Taylor, Faith E.; Malamud, Bruce D.; Witt, Annette; Guzzetti, Fausto (2018): Landslide shape, ellipticity and length-to-width ratios. In *Earth Surf. Process. Landforms* 43 (15), pp. 3164–3189. DOI: 10.1002/esp.4479.

Tuanmu, Mao-Ning; Viña, Andrés; Roloff, Gary J.; Liu, Wei; Ouyang, Zhiyun; Zhang, Hemin; Liu, Jianguo (2011): Temporal 580   transferability of wildlife habitat models: implications for habitat monitoring. In *Journal of Biogeography* 38 (8), pp. 1510–1523. DOI: 10.1111/j.1365-2699.2011.02479.x.

van Westen, C. J.; van Asch, T.W.J.; Soeters, R. (2006): Landslide hazard and risk zonation—why is it still so difficult? In *Bull Eng Geol Environ* 65 (2), pp. 167–184. DOI: 10.1007/s10064-005-0023-0.

Wenger, Seth J.; Olden, Julian D. (2012): Assessing transferability of ecological models: an underappreciated aspect of 585   statistical validation. In *Methods in Ecology and Evolution* 3 (2), pp. 260–267. DOI: 10.1111/j.2041-210X.2011.00170.x.

Wichmann, V.; Becht, M. (2006): Rockfall modelling: methods and model application in an alpine basin (Reintal, Germany). In J. Böhner, K. R. McCloy, J. Strobl (Eds.): SAGA – Analysis and Modelling Applications, vol. 115, pp. 105–116.

Wichmann, Volker (2017): The Gravitational Process Path (GPP) model (v1.0) – a GIS-based simulation framework for gravitational processes. In *Geosci. Model Dev.* 10 (9), pp. 3309–3327. DOI: 10.5194/gmd-10-3309-2017.

590   Wickham, Hadley (2009): ggplot2. Elegant graphics for data analysis. New York: Springer (Use R!).

Wood, Simon N. (2011): Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. In *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73 (1), pp. 3–36. DOI: 10.1111/j.1467-9868.2010.00749.x.

595