

RC1

Runoff modelling is one of the most important issues in hydrology. In this submission, an R package named runoffSIM is developed to establish a relationship between the runoff process and the probability of landslides. In general, the modelling process consists of two stages: first, runoff is calculated, and then landslide probability is evaluated.

Four comments are provided below to facilitate further improvements of the paper:

1. A wide variety of runoff simulation algorithms exist in the literature. In the first stage of the proposed modelling framework, are other existing runoff simulation algorithms applicable to runoff calculation? If yes, the authors are advised to compare the effectiveness of runoffSIM with that of these existing algorithms. If not, the authors should elaborate on the unique advantages of runoffSIM that make it superior or more suitable for the proposed application.
2. The formulation of landslide probability within runoffSIM is only briefly described. Given the central role of landslide analysis in this package, more detailed explanations are required. In particular, the intricate relationships between runoff dynamics and landslide occurrence (e.g., how runoff parameters influence landslide probability) have not yet been sufficiently elaborated.
3. Observational datasets that include both runoff and landslide data are relatively scarce in peer-reviewed studies. Is the dataset used in the current analysis publicly available? The authors are recommended to provide more detailed information about the dataset, such as its source, spatial-temporal coverage, data collection methods, and basic descriptive statistics.
4. Verification is a critical component for evaluating the effectiveness and reliability of runoffSIM. The authors are encouraged to refer to commonly used verification metrics and diagnostic plots in the field of forecast verification (e.g., bias, root mean square error, receiver operating characteristic curves). Additionally, the verification experiments designed to test runoffSIM, as well as the specific metrics employed, should be described in greater detail to enhance the credibility of the results.

Thanks to the reviewer for their feedback. I want to mention that this reviewer's comments have a bit of mix up. They refer to runoffSIM and "runoff" simulation, while this paper presents runoutSIM, an R package for landslide "runout" simulation: these are distinct hydrological and geomorphological processes. That said, assuming this is a mix-up, I will address them as some of the concerns can apply to modelling in general.

Comment 1.

I agree that benchmarking is an important component of model performance assessment, and such comparisons are welcome in future work. The scope of this paper was focused on geospatial software development, tutorial creation, and the implementation of well-established approaches; specifically Perla et al.'s (1980) two-parameter friction model and Gamma's (2000) random walk for flow path simulation, which have been validated and discussed in prior research (e.g. Goetz et al., 2021) and in the case study presented here. The goal was not to show better model performance over existing tools, but to create a flexible open R-based environment that supports model improvement and integration with statistical and machine learning workflows.

Comment 2.

The probability outputs in runoutSIM are not derived from runoff dynamics. They refer to traverse probabilities (the likelihood of a runout path traversing a given grid cell) and connectivity probabilities (the probability that runout from a source cell reaches a downslope feature of interest). Since existing methods were adapted into the R environment, the Methods section provides an in-depth overview of how the algorithms work, while also directing readers to previous works that discuss the modelling components in greater depth (e.g. L100: "For more details on the random walk and PCM modelling components, it is recommended to refer to Wichmann (2017) and Goetz et al. (2021)"). The traverse and connectivity probabilities are described in Sections 2.2.2 and 2.2.4 respectively. Optionally, source area probabilities derived from statistical or machine learning models can be used to weight these outputs, as described in Section 2.4.

Comment 3.

As described in the case study section, the dataset used in this study is a debris flow inventory of 73 manually mapped runout polygons for the Río Olivares basin, Chile, derived from photo-interpretation of high-resolution satellite imagery spanning 2000 to 2019. The

imagery sources, spatial coverage, and mapping approach are described in Section 3.1. The DEM is the publicly available ALOS PALSAR radiometrically terrain corrected product at 12.5 m resolution. All case study data, analysis scripts, and simulation results are publicly available via the Zenodo repository linked in the Code and Data Availability section, supporting full reproducibility.

Comment 4.

Verification is addressed in the paper using the relative runout length error and the area under the receiver operating characteristic curve (AUROC), which are both standard metrics in regional runout model validation (Goetz et al., 2021). Results are reported in Section 3.5.1, including the median error, interquartile range, and observed versus simulated runout length plots (Figure 3).

RC2

Hello Jason,

I have read your manuscript with particular interest.

In my life, I have only suggested accept-as-is to 2 manuscripts I have personally revised.

This is the third time I am doing so. Very well written document, very interesting content, and a useful addition to the literature and tools available out there.

Perhaps I would have expected some reflections on the potential applications of your tool for operational purposes and a summary of your own vision about what a second version would entail. Instead, your discussion is primarily based on technical aspects. However, this is purely a matter of taste, which is why I have not even suggested for a minor revision but rather opted to value your efforts and "reward your contribution for what it is", a manuscript already at the level of a publication, without the need for external requirements, suggestions, etc.

Congratulations again, at least from my side. Really a nice work indeed. I have already shared the code with a student of mine, asking her to test it!

Cheers,

Luigi Lombardo

Thank you Luigi for your interest in the manuscript. I really appreciate your positive and encouraging view of this research.

Excellent comments. I favoured more technical aspects to help support 'practical' adoption of the model. Operationally, in addition to its use for general susceptibility / runout impact assessment and identifying major source areas for mitigation (i.e. with the connectivity probabilities), I also see this tool being developed to integrate short- and medium-term forecast weather data API's to support regional scale runout predictions. The simplified workflow in an accessible programming language makes (in-house) maintenance of the system more feasible than coupling or linking other software environments. While this can use meteorological thresholds where they are well established, the framework is also designed to support threshold-free runout forecasts by conditioning on spatial probabilities of the source area directly.

Future revisions would likely include additions that make it easier to change modelling components, like the model controlling the runout distance, using wrapper functions. For example, if you wanted to use energy line approaches or a different friction model.

I've added the following to section 4.5 to address these:

“The framework of runoutSIM is also designed to support operational deployment. While meteorological thresholds can be applied where they are well established to hard-classify sources areas for an event, the framework is designed to carry spatial likelihoods of initiation through the modelling chain, supporting the possibility of threshold-free runout forecasts.”

And updated the last paragraph in section to include these ideas:

“For research, training, and prototyping, the current R-based implementation should be sufficient. It supports rapid iteration, easy integration with other R packages, and flexible model development. Future improvements may focus on hybrid implementations that balance interpretability with performance, depending on the needs of the application, as well as wrapper functions to make it easier to swap modelling components. For example, wrapper functions that can substitute the current two-parameter friction model with energy line approaches or alternative friction models that would allow users to better match model complexity to the available data and computational requirements.”

All the best,

Jason Goetz