Check for updates

Computationally-feasible uncertainty quantification in model-based landslide risk assessment

Anil Yildiz*, Hu Zhao and Julia Kowalski

Methods for Model-based Development in Computational Engineering, RWTH Aachen University, Aachen, Germany

Introduction: Increasing complexity and capacity of computational physics-based landslide run-out modelling yielded highly efficient model-based decision support tools, e.g. landslide susceptibility or run-out maps, or geohazard risk assessments. A reliable, robust and reproducible development of such tools requires a thorough quantification of uncertainties, which are present in every step of computational workflow from input data, such as topography or release zone, to modelling framework used, e.g. numerical error.

Methodology: Well-established methods from reliability analysis such as Point Estimate Method (PEM) or Monte Carlo Simulations (MCS) can be used to investigate the uncertainty of model outputs. While PEM requires less computational resources, it does not capture all the details of the uncertain output. MCS tackles this problem, but creates a computational bottleneck. A comparative study is presented herein by conducting multiple forward simulations of landslide run-out for a synthetic and a real-world test case, which are used to construct Gaussian process emulators as a surrogate model to facilitate high-throughput tasks.

Results: It was demonstrated that PEM and MCS provide similar expectancies, while the variance and skewness differ, in terms of post-processed scalar outputs, such as impact area or a point-wise flow height. Spatial distribution of the flow height was clearly affected by the choice of method used in uncertainty quantification.

Discussion: If only expectancies are to be assessed then one can work with computationally-cheap PEM, yet MCS has to be used when higher order moments are needed. In that case physics-based machine learning techniques, such as Gaussian process emulation, provide strategies to tackle the computational bottleneck. It can be further suggested that computational-feasibility of MCS used in landslide risk assessment can be significantly improved by using surrogate modelling. It should also be noted that the gain in compute time by using Gaussian process emulation critically depends on the computational effort needed to produce the training dataset for emulation by conducting simulations.

1 Introduction

Computational landslide run-out models can predict the spatial evolution of depth and velocity of the failed mass, which is crucial for landslide risk assessment and mitigation, especially for flow-like landslides due to their rapid nature (Cepeda et al., 2013; McDougall, 2017). Utilising computational landslide run-out models for model-based decision support requires a well-defined, transparent and modular setup of the complete computational value-chain. Such a chain consists of many links, including a digital representation of the topography, the underlying physics-based process model, a numerical solution scheme, the approach to parameter calibration along with the training data it relies on, and concepts used for sensitivity analyses and uncertainty quantification. Challenges in the technical realisation of such integrated workflows have been successfully addressed in the past (Dalbey et al., 2008; Aaron et al., 2019; Sun X. P. et al., 2021b; Zhao et al., 2021; Aaron et al., 2022; Zhao and Kowalski, 2022). It will be of crucial importance in the future to increase the efficiency, sustainability and, hence, acceptance of such orchestrated workflows for landslide risk assessment by improving their robustness, reliability and computational-feasibility.

It will be particularly important to assess the reliability of landslide risk assessment by quantifying and managing uncertainties throughout the workflow. This is a challenging task, which requires to consider and structure the *landscape of uncertainties* affecting various steps of the decision-making process. Relevant uncertainty originates from uncertain model input—such as the digital elevation model representing the topography (Zhao and Kowalski, 2020) or release area and volume—and rheological parameters (Quan Luna et al., 2013). Furthermore, process uncertainty can result from numerical modelling schemes (Schraml et al., 2015) or calibration methods (Aaron et al., 2019; 2022). All relevant uncertainties in the computational workflow include aleatoric aspects due to the intrinsic randomness of the process, as well as epistemic uncertainty that is of systemic nature, and might be due to a lack of data. A comprehensive, integrated uncertainty analysis within a georisk assessment framework is an important reminder of the limitations of the knowledge about processes involved, and the need to improve data collection and quality (Eidsvig et al., 2014).

Different methods of reliability analysis have been used in the recent decades both in geohazards research and its practical implementation, to assess the output uncertainty in landslide run-out models due to uncertain input parameters, or to quantify the uncertainty of derived metrics such as the Factor of Safety (FoS). Point Estimate Method (PEM) (Przewlocki et al., 2019), First Order Second Moment method (FOSM) (Kaynia et al., 2008), First Order Reliability Method (FORM) (Sun X. et al., 2021a), and Monte Carlo Simulations (MCS) (Cepeda et al., 2013; Liu et al., 2019; Brezzi et al., 2021) are examples of such methods. Dalbey et al. (2008) presented several standard and new methods for characterising the effect of input data uncertainty on model output for hazardous geophysical mass flows.

PEM is a simple way to determine the expectation (mean), variance, and skewness of a variable that depends on a random input, by evaluating the function at a low number of pre-selected values. Additionally, MCS grants access to the complete probability distribution even in complex problems (Fenton and Griffiths, 2008). However, it requires a large number of model evaluations, so-called

realisations, at randomly selected inputs following a pre-determined statistical distribution. Przewlocki et al. (2019) used PEM to conduct a probabilistic slope stability analysis of a sea cliff in Poland, and compared the moment estimates with results from MSC. Mean and standard deviation values of FoS yielded similar results, and PEM was favoured as it required a lower number of model realisations, hence lower computational costs for a seemingly similar information outcome. Tsai et al. (2015) also obtained similar estimations by comparing PEM and MSC, but pointed out the effects of correlation between input variables. Earlier works also highlighted the limited feasibility of PEM for an increasing number of input variables, as 2^n estimations are required for n input variables (Christian and Baecher, 1999; 2002). Therefore, it is more feasible to handle problems characterised by a low-dimensional input parameter space with PEM, while high-dimensional problems quickly become computationally infeasible.

MCS has also been widely used for practical uncertainty quantification owing to its simplicity of implementation. Liu et al. (2019) and Ma et al. (2022) used MCS to quantify uncertainty in landslide run-out distance due to uncertain soil properties. Brezzi et al. (2021) performed uncertainty quantification of the Sant'Andrea landslide using MCS. They assumed the two friction parameters (Coulomb friction and turbulent friction) in a depth-averaged Voellmy-Salm type approach to follow independent Gaussian distributions, and then studied the induced uncertainty in deposit heights. The major challenge of using MCS for uncertainty quantification in landslide run-out modelling is the high computational cost, as pointed out by many researchers (Dalbey et al., 2008; McDougall, 2017; Aaron et al., 2022; Zhao and Kowalski, 2022).

As PEM relies on a much lower number of sampling points compared to MCS, the probability distribution function (PDF) of the output cannot be reliably approximated for a general non-linear process model, such as a landslide run-out model in complex topography. MCS provides an approximation of the PDF, but it creates a computational bottleneck when the forward-model is complex and subject to long runtimes. Physics-based machine learning, i.e. creating a surrogate by training a data-driven model with results from a physics-based simulation, can overcome the computational bottleneck of highly-throughput tasks such as uncertainty quantification. The surrogate model can be sampled, instead of the sampling from the forward model, and hence the PDF of the quantity of interest can be calculated efficiently. An example of physics-based machine learning techniques proven to be effective in many applications related to geohazards is Gaussian Process Emulation, with successful demonstrations in landslide run-out models (Zhao et al., 2021; Zhao and Kowalski, 2022) and stability of infrastructure slopes (Svalova et al., 2021).

This study aims at demonstrating how an uncertainty quantification workflow can be set up effectively, and how this affects the model-based landslide risk assessment. A test case with a synthetic topography and a test case with a real world problem are designed. Multiple forward simulations of both cases are conducted to construct Gaussian process emulators to facilitate MCS. The objectives are (i) comparing the results of PEM-based simulations and MCS conducted with emulators trained based on datasets from a limited number of simulations in terms of three moments, (ii) investigating the effects of topographic complexity, i.e. synthetic topography vs. the topography of a real-world problem, on the PEM - MCS comparison,

and (iii) demonstrating the applicability of emulation techniques for uncertainty quantification.

2 Materials and methods

2.1 Modelling approach

Existing physics-based landslide run-out models can be divided into three groups: lumped mass models, particle models, and continuum models. Lumped mass models treat the flow mass as a condensed mass point without spatial variation. This process idealisation greatly reduces the complexity of the problem, but the trade-off is losing spatial variation of flow dynamics, such as internal deformation of the flowing mass. Particle models treat the flow mass as an assembly of particles and simulate the movement of each particle and their interactions in order to characterise flow dynamics. They can directly account for three-dimensional flow behaviours, including an internal re-distribution of mass. Computational realisation of a particle model relies on the definition of conceptual particles, whose size is chosen based on available computational resources, and is often much larger than actual material particles in the landslide. This is beneficial from an implementation point of view, yet requires special attention, when formulating necessary interaction forces which are often challenging to justify and validate. Continuum models treat the flow mass as continuum material, for which governing equations are derived from balance laws closed by tailored, complex constitutive relations. Implementing these in a general three-dimensional context is very challenging and uses a lot of computational resources. A majority of particularly relevant continuum models for practical landslide run-out modelling are formulated within a depth-averaging framework. Depth-averaged continuum models balance computational efficiency, accuracy and interpretability. They can account for internal deformation of the flow material, and provide spatial variation of flow dynamics. Eqs. 1–3 describe the governing system of an idealised depth-averaged landslide run-out model:

$$\partial_t h + \partial_x\left(hu_x\right) + \partial_y\left(hu_y\right) = 0, \tag{1}$$

$$\partial_t\left(hu_x\right) + \partial_x\left(hu_x^2 + g_n\frac{h^2}{2}\right) + \partial_y\left(hu_xu_y\right) = g_x h - S_{fx}, \tag{2}$$

$$\partial_t\left(hu_y\right) + \partial_x\left(hu_xu_y\right) + \partial_y\left(hu_y^2 + g_n\frac{h^2}{2}\right) = g_y h - S_{fy}. \tag{3}$$

The equations are derived from the mass balance (Eq. 1) and momentum balance (Eqs. 2, 3). The flow height h and the depth-averaged surface tangent flow velocities u_x and u_y are the state variables. g_x, g_y, and g_n are components of the gravitational acceleration along surface tangent and in normal directions. The friction terms S_{fx} and S_{fy} depend on the chosen basal rheology. In terms of the Voellmy rheological model, they are defined as:

$$S_{fi} = \frac{u_i}{\|\mathbf{u}\|}\left(\mu g_n h + \frac{g}{\xi}\|\mathbf{u}\|^2\right), \quad i \in \{x, y\}, \tag{4}$$

where $\|\mathbf{u}\|$ denotes the magnitude of the flow velocity; μ and ξ are the dry-Coulomb friction coefficient and turbulent friction coefficient respectively.

Many numerical solvers for depth-averaged landslide run-out models have been developed in the past decades, and McDougall (2017) provides a comprehensive review. A GIS-based open source computational tool developed by Mergili et al. (2017), r.avaflow v2.3, is used in this study. It implements a high-resolution total variation diminishing non-oscillatory central differencing scheme to solve Eqs. 1–4 given topographic data, initial mass distribution, and Voellmy parameters. It runs on Linux systems, and employs the GRASS GIS software, together with the programming languages Python and C, and the statistical software R. A Digital Elevation Model (DEM) and a release height map are given as input as raster files.

2.2 Case studies

2.2.1 Synthetic case

A simple topography has been created in Python—similar to the topography generation in AvaFrame (D'Amboise et al., 2022) — which is denoted as *synthetic case* herein. Topography consists of a parabolic slope starting at $x = 0$ at an altitude of 1332 m, and connecting to a flat land at $x = 3000$ at an altitude of 0 m. Extent of the area in x and y directions are 5000 m and 4000 m, respectively, whereas the resolution is 20 m. Release zone was defined as an elliptic cylinder, of which the centre is located at $(x, y) = (600, 2000)$ with a minor axis of 100 m and a major axis of 200 m. Height at each point within the ellipse was defined as 20 m, which generates a total release volume of 1.432×10^6 m^3. Figure 1A illustrates the maximum flow height map from a simulation of synthetic case, while Figure 1B shows the deposit height from the same simulation.

2.2.2 Acheron rock avalanche

A real-world case is chosen to compare the differences in uncertainty quantification with the synthetic case. Radiocarbon testing dates the occurrence of Acheron rock avalanche—located near Canterbury, New Zealand—approximately 1100 years before present, and it may have been triggered by seismic activity (Smith et al., 2006). The deposit area was estimated to be $.72 \times 10^6$ m^2 using a GPS outline of the deposit, while the deposit volume was estimated as 8.9×10^6 m^3 using an estimated mean depth derived from observed and estimated thicknesses for different morphological zones (Smith et al., 2012). DEM file and the release height map are obtained from Mergili and Pudasaini (2014–2021)[1], which gives an initial release volume of 6.4×10^6 m^3. Figure 2 presents the shaded relief of the area, together with a map of maximum flow height (see Figure 2A) and deposit height (see Figure 2B) from a random simulation of Acheron rock avalanche.

2.3 Gaussian process emulation

The main problem of applying MCS for uncertainty quantification in landslide run-out modelling is its high computational cost. The runtime of the landslide run-out model is one of the key drivers of computational cost in classical MCS, as it scales with the number of

---

1 Mergili, M., Pudasaini, S.P., 2014–2021. r.avaflow—The mass flow simulation tool. https://www.avaflow.org. Accessed on 2022-07-12.

FIGURE 1
(A) Maximum flow height and **(B)** deposit height from a simulation of the synthetic case.

FIGURE 2
(A) Maximum flow height and **(B)** deposit height from a simulation of Acheron rock avalanche.
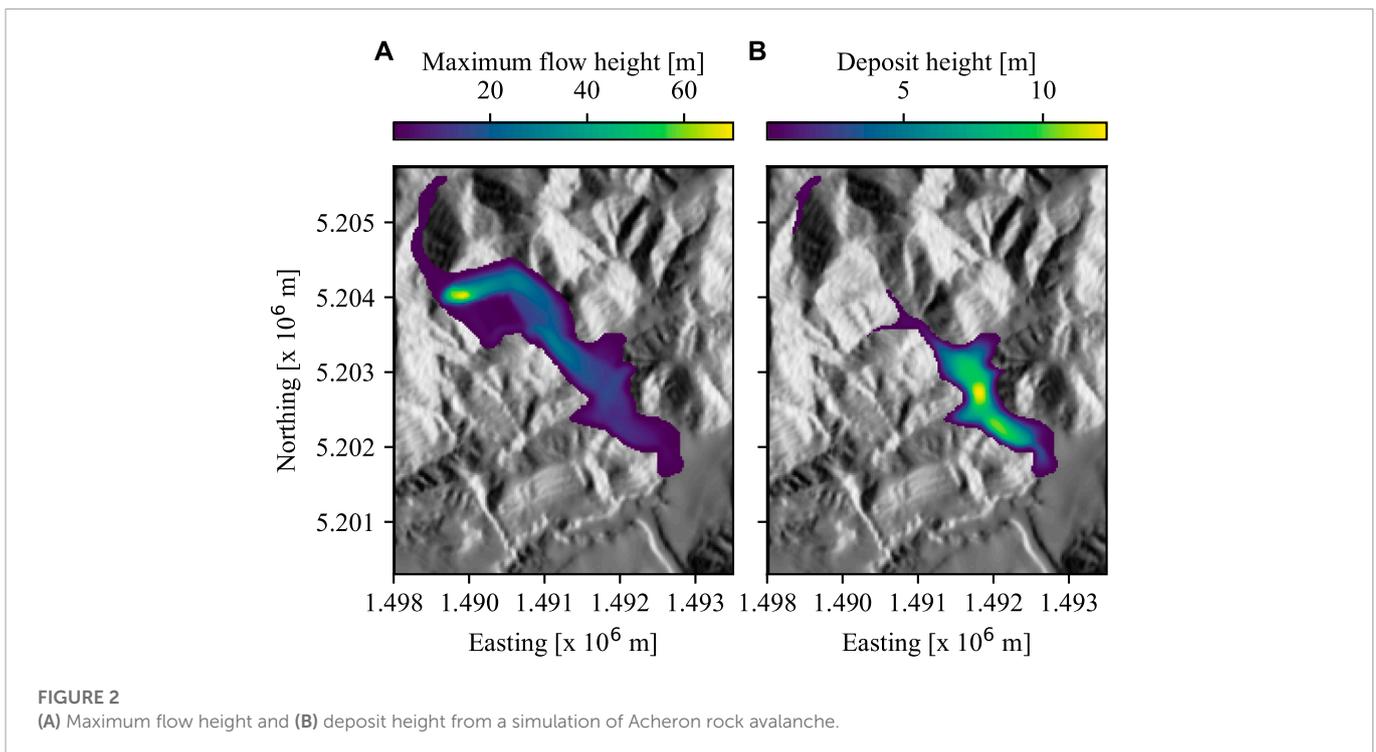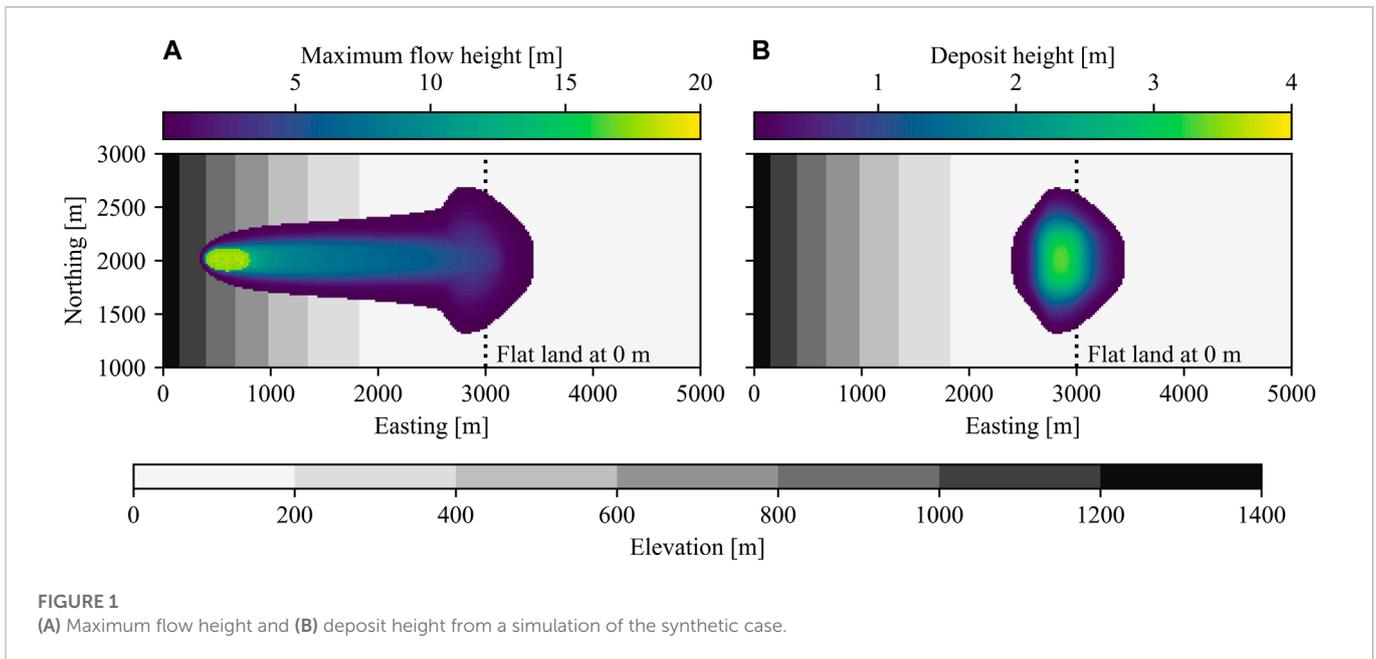
forward evaluations. Gaussian process emulation has been used in recent years to build cheap-to-evaluate emulators to replace expensive-to-evaluate computational models in the framework of uncertainty quantification, such as Sun X. P. et al. (2021b), Zeng et al. (2021), and Zhao and Kowalski (2022). A Gaussian process emulator is a statistical approximation of a simulation model, built based on input and output data of a small number of simulation runs. Once an emulator is

constructed, it provides prediction of simulation output at a new input point almost instantly, together with an assessment of the prediction uncertainty. This emulator-induced uncertainty can be taken into account in the framework of uncertainty quantification.

Let $y = f(\mathbf{x})$ denote a simulator where y represents a scalar output depending on a p-dimensional input $\mathbf{x}$. Assuming that the simulator is a realisation of a Gaussian process with a mean function $m(\cdot)$ and a

kernel function $k(\cdot, \cdot)$, namely

$$f(\cdot) \sim \mathcal{N}(m(\cdot), k(\cdot, \cdot)), \qquad (5)$$

A Gaussian process emulator can be built based on the input-output data $\mathcal{D}$ of n simulation runs, namely $\mathcal{D} = \{\mathbf{x}_i, y_i\}_{i=1}^n$. At any new input point $\mathbf{x}^*$, the approximated output y^* follows a Gaussian distribution given by

$$y^* \sim \mathcal{N}\left(m'\left(\mathbf{x}^*\right), k'\left(\mathbf{x}^*, \mathbf{x}^*\right)\right), \qquad (6)$$

$$m'\left(\mathbf{x}^*\right) = m\left(\mathbf{x}^*\right) + \mathbf{k}^T\left(\mathbf{x}^*\right)\mathbf{K}^{-1}(y_1 - m(\mathbf{x}_1), \dots, y_n - m(\mathbf{x}_n))^T, \qquad (7)$$

$$k'\left(\mathbf{x}^*, \mathbf{x}^*\right) = k\left(\mathbf{x}^*, \mathbf{x}^*\right) - \mathbf{k}^T\left(\mathbf{x}^*\right)\mathbf{K}^{-1}\mathbf{k}\left(\mathbf{x}^*\right). \qquad (8)$$

The symbol $\mathbf{K}$ denotes the $n \times n$ covariance matrix of which the (i,j)-th entry $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ and the symbol $\mathbf{k}(\mathbf{x}^*) = \left[k(\mathbf{x}^*, \mathbf{x}_1), \dots, k(\mathbf{x}^*, \mathbf{x}_n)\right]^T$.

RobustGaSP package developed by Gu et al. (2019) is used in this study to build emulators. It provides robust Gaussian process emulation for both single-variate (Gu et al., 2018), i.e. one simulation producing one scalar output, and multi-variate simulators (Gu and Berger, 2016), i.e. one simulation producing high-dimensional output. Training and validation datasets of each case were generated from outputs of both cases simulated with r. avaflow 2.3. 100 simulations for training, and 20 additional simulations for validation of vector emulators were run for synthetic case and Acheron rock avalanche separately. Dry-Coulomb friction coefficient (μ), turbulent friction coefficient (ξ), and release volume (v_o) were chosen as uncertain input variables. The friction coefficients are rather conceptual than physical (Fischer et al., 2015) and rely on back-analysing past events where field data are available. The calibrated results are often ranges (Zhao et al., 2021) or probability density functions (Quan Luna et al., 2013). In this study, ranges of friction coefficients were chosen as $\mu = [.02, .3]$, $\xi = [100, 2200]$ m/s^2 according to Zhao et al. (2021). The release volume of a future landslide event is hardly predictable. Uncertainty in v_o were defined by multiplying the release height map with a coefficient k, which is assumed to vary between .5 and 1.5 in this study. Parameter combinations were generated using Latin Hypercube sampling by maximising the minimum distance between points. Datasets were extracted from simulation results after training, and scalar outputs, e.g. impact area, deposit area, deposit volume, and vector outputs, e. g flow height or flow velocity in each cell, were defined.

2.4 Uncertainty analysis

2.4.1 Point estimate method

Any scalar output $y(\mathbf{s}, t)$ — either aggregated, such as impact area, deposit area or deposit volume, or point-wise flow height, flow velocity, or flow pressure—in space ($\mathbf{s}$) and time (t), from the landslide run-out simulations can be expressed as a function of three uncertain input variables as shown in Eq. 9.

$$y(\mathbf{s}, t) = f(\mu, \xi, v_o) \qquad (9)$$

Assuming that the input variables are not correlated and skewness is 0, locations of sampling points of each variable correspond to mean

± standard deviation. This results in eight sampling points with equal weights for three input variables to evaluate the output function y and calculate the three moments of the output, i.e. mean, variance, skewness. PEM were executed three times with different coefficient of variation (*COV*), i.e. 10%, 25% and 50%. Table 1 shows the assumed mean, which was chosen as the central point of the ranges defined in Section 2.3, and standard deviations for the input variables used in the analyses, PEM 1, PEM 2, and PEM 3.

2.4.2 Monte Carlo Simulations

Parameter combinations for MCS were sampled from truncated multivariate normal distribution using the R package, *tmvtnorm* (Wilhelm and Manjunath, 2010). Points of truncation were chosen as the ranges given in Section 2.3. Three sets of MCS were conducted similarly to the PEM, i.e. mean value is the central point of the range and *COV* is chosen arbitrarily as 10% (MCS 1), 25% (MCS 2), and 50% (MCS 3) to represent different levels of uncertainty in input variables. 10,000 parameter sets were generated for each MCS analysis, and the outputs are estimated using the emulators defined in Section 2.3.

3 Results

Simulation outputs used in this study can be found in Yildiz et al. (2022a), and the general workflow, as well as the scripts to reproduce the figures can be found in the Git repository presented in Yildiz et al. (2022b).

3.1 r.avaflow simulations

A total of 100 simulations of the synthetic case and a separate 100 simulations for Acheron rock avalanche have been used to calculate the scalar outputs, and to extract vector outputs. Quantities of interest (*QoI*) derived from the simulations are impact area, deposit area and deposit volume. In addition to the derived ones, direct simulation outputs, i.e. maximum flow height ($h_{\max}$) and maximum flow velocity ($v_{\max}$) at a predefined cell, are extracted from the simulations. Point of extraction was chosen arbitrarily as $(x, y) = (1000, 2000)$ for the synthetic case, and as (1490100, 5204100) for Acheron rock avalanche. These points were picked among those which had a value higher than a threshold, e.g. a maximum flow height of .1 m, from all simulations.

Evolution of a simulation in synthetic case can be summarised as a rather constrained flow, with a limited lateral spread at the upper section of the slope, and a more pronounced lateral spread close to the transition to the flat land (See Figure 1A). As no stopping criteria was defined, the failed mass accumulates mostly around the toe of the slope (See Figure 1B). Mean values ± standard deviations of the impact area, deposit area and deposit volume were $(2.39 \pm .37) \times 10^6$ m^2, $(1.22 \pm .22) \times 10^6$ m^2 and $(1.37 \pm .41) \times 10^6$ m^3, respectively. Ranges of $v_{\max}$ and $h_{\max}$ at $(x, y) = (1000, 2000)$ are 24.9–63.4 m/s and 7.18–14.4 m, with mean values of 47.7 m/s and 11.0 m, respectively.

Flow path of Acheron rock avalanche can be generalised based on the simulations conducted in this paper as an initially relatively straight path, followed by a sharp turn and extending into the valley (See Figure 2A). Similar to synthetic case, 100 simulations of Acheron rock avalanche were run, and the same scalar outputs were calculated or extracted. Mean values ± standard deviations of the impact area, deposit area and deposit volume were $(2.78 \pm .91) \times 10^6$ m^2, $(1.45 \pm$

TABLE 1   Mean and standard deviation values of the input variables used in the Point Estimate Method (PEM) analysis of the synthetic case and Acheron rock avalanche. Same values of Dry-Coulomb and turbulent friction coefficients are used for both cases.

	Unit	PEM 1 (COV = 10%)		PEM 2 (COV = 25%)		PEM 3 (COV = 50%)	
		mean	Std. dev.	mean	Std. dev.	mean	Std. dev.
Dry-Coulomb friction coefficient	-	.160	.016	.160	.040	.160	.080
Turbulent friction coefficient	m/s^2	1150	115.0	1150	287.5	1150	575.0
Release volume - Synthetic	$x\ 10^6\ m^3$	1.432	.143	1.432	.358	1.432	.716
Release volume - Acheron	$x\ 10^6\ m^3$	6.40	.64	6.40	1.60	6.40	3.20

TABLE 2   Coefficient of determination, R^2, mean absolute percentage error, *MAPE*, and normalised root mean squared error *nRMSE* for the emulators trained with scalar outputs from synthetic case and Acheron rock avalanche.

Output	Synthetic case			Acheron rock avalanche		
	R^2 [-]	MAPE [%]	nRMSE [%]	R^2 [-]	MAPE [%]	nRMSE [%]
Impact area	.999	.468	.606	.997	1.294	1.900
Deposit area	.996	.864	1.173	.988	2.530	3.100
Deposit volume	.999	.553	.717	.999	.541	.752
Maximum velocity	.999	.592	.731	.967	2.802	3.473
Maximum height	.989	1.609	1.996	.965	2.084	2.828

.42) x 10^6 m^2 and (6.28 ± 1.83) x 10^6 m^3, respectively. v_{max} and h_{max} at (x,y) = (1490100, 5204100) have mean values of 32.2 m/s and 42.2 m, with standard deviations of 6.2 m/s and 6.4 m.

3.2 Emulation

Gaussian process emulation has been used in this study in order to facilitate the prediction of many scalar outputs for MCS analysis. Once the training datasets consisting of scalar outputs described in **Section 3.1** are generated, the trained emulators are first validated with leave-one-out cross-validation technique. **Table 2** presents the validation results of scalar emulators, i.e. mean absolute percentage error (*MAPE*), normalised root-mean-square error (*nRMSE*), and coefficient of determination (R^2) values for both cases. Emulators trained for synthetic case produced very low percentage-based errors, i.e. maximum *MAPE* and *nRMSE* was 1.609% and 1.996%, respectively, while the R^2 values were around .99. The lowest prediction quality was obtained for maximum flow height at (x,y) = (1000, 2000) in synthetic case. Emulators trained with derived scalar outputs, i.e. impact area, deposit area and deposit volume, of Acheron rock avalanche yielded similar model performance as the synthetic case. Relatively lower R^2 values, approximately .97, and higher *MAPE* and *nRMSE* values were obtained for v_{max} and h_{max} at (x,y) = (1490100, 5204100). Lowest prediction performance for Acheron rock avalanche was for maximum flow velocity at the given point.

Prediction quality of the vector emulators both for synthetic case and Acheron rock avalanche has been evaluated using testing data from additional 20 simulations. $P_{CI}(95\%)$, defined by Gu and Berger (2016), is chosen as the diagnostic for vector emulators. It represents the proportion of testing outputs that lie in emulator-based 95% credible intervals. $P_{CI}(95\%)$ of the vector emulator for point-wise maximum flow height is 83.8%

and 85.3% for synthetic case and Acheron rock avalanche respectively; $P_{CI}(95\%)$ of the vector emulator for point-wise maximum flow velocity is 86.3% and 89.1% for synthetic case and Acheron rock avalanche respectively.

3.3 Uncertainty analysis

Uncertainty of model outputs were investigated by conducting PEM and MCS with three different *COV*, a comparison has been made in terms of the three moments, i.e. mean, variance, and skewness, of the scalar outputs. **Figure 3** presents the comparison for the synthetic case. If the uncertainty of the model outputs is assessed *via* PEM, similar mean values are obtained even though the *COV* is varied from 10% up to 50%. Differences were higher for v_{max} and h_{max} (See **Figures 3D, E**). An increase of variance with increasing *COV* is evident for all outputs produced by PEM, whereas no clear relationship can be defined for skewness. For example, the skewness of impact area and deposit area increased with increasing *COV*, whereas the deposit volume, v_{max} and h_{max} had nearly no skewness if analysed by PEM.

Similar to PEM, MCS at all *COV* produced similar mean values and increasing variance with increasing *COV* for all scalar outputs. No overall trend can be observed in skewness of the outputs generated *via* MCS. If both techniques are compared, no significant difference is present in mean values. PEM produced higher variances especially at the highest *COV*, while—similar to previous comparisons—no generalisations can be made for skewness. It should be noted that there is a change from a slightly positive skewness to negative skewness for v_{max} and h_{max}, if the method is switched from PEM to MCS. Other scalar outputs had rather arbitrary changes between methods, even though slightly lower values are observed for impact area and deposit area.
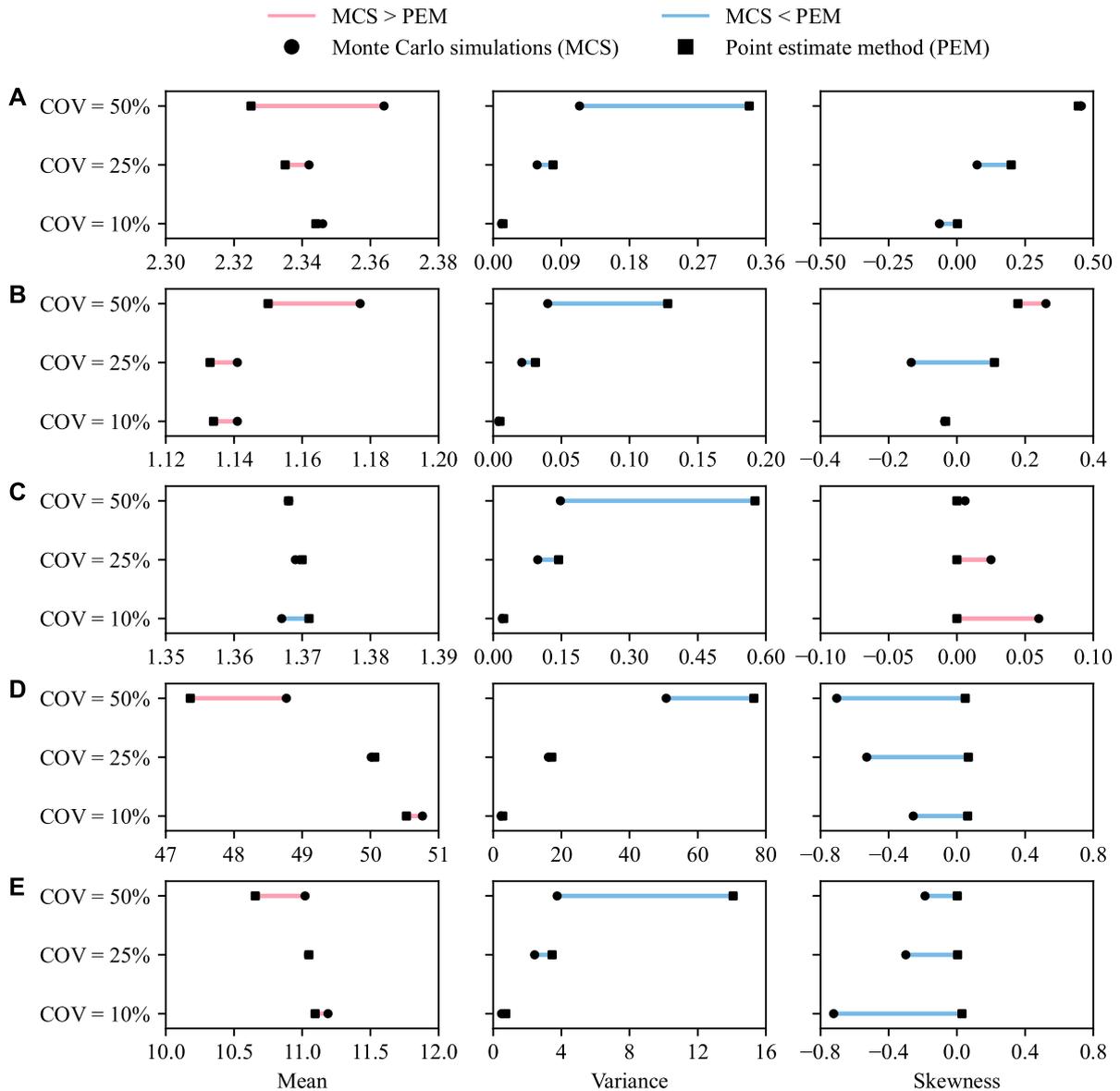
FIGURE 3
Three moments (mean, variance and skewness) of **(A)** impact area (in x10^6 m^2), **(B)** deposit area (in x10^6 m^2), **(C)** deposit volume (in x10^6 m^3), **(D)** maximum flow velocity (in m/s) and **(E)** maximum flow height (in m) at (x,y) = (1000, 2000) obtained from synthetic case with Monte Carlo simulations (MCS) and point estimate method (PEM).

Figure 4 presents three moments of the same scalar outputs from Acheron rock avalanche. Similar patterns to data from synthetic case are observed in **Figure 4**. Mean values of the outputs are similar between different methods and levels of *COV*. Variance increases in both methods with increasing *COV*, while the values obtained from MCS is lower than those from PEM.

Figure 5 illustrates the synthetic case results of the MCS analysis at *COV* of 50% in terms of the five scalar outputs–the impact area, deposit area, deposit volume, v_{max} and h_{max} at (x,y) = (1000, 2000). The first three columns show the relationships of the scalar outputs with respect to the three uncertain parameters.

The last column shows the histograms of each scalar output. **Figure 6** presents the corresponding results of Acheron rock avalanche. Observations from **Figures 5**, **6** can be generalised as follows.

- the impact area and deposit area decrease with increasing dry-Coulomb friction coefficient and increases with increasing release volume (see **Figures 5A**, **B**, **6A**, **B**),
- the deposit volume is proportional to the release volume and has almost no dependence on the dry-Coulomb and turbulent friction coefficients (see **Figures 5C**, **6C**),
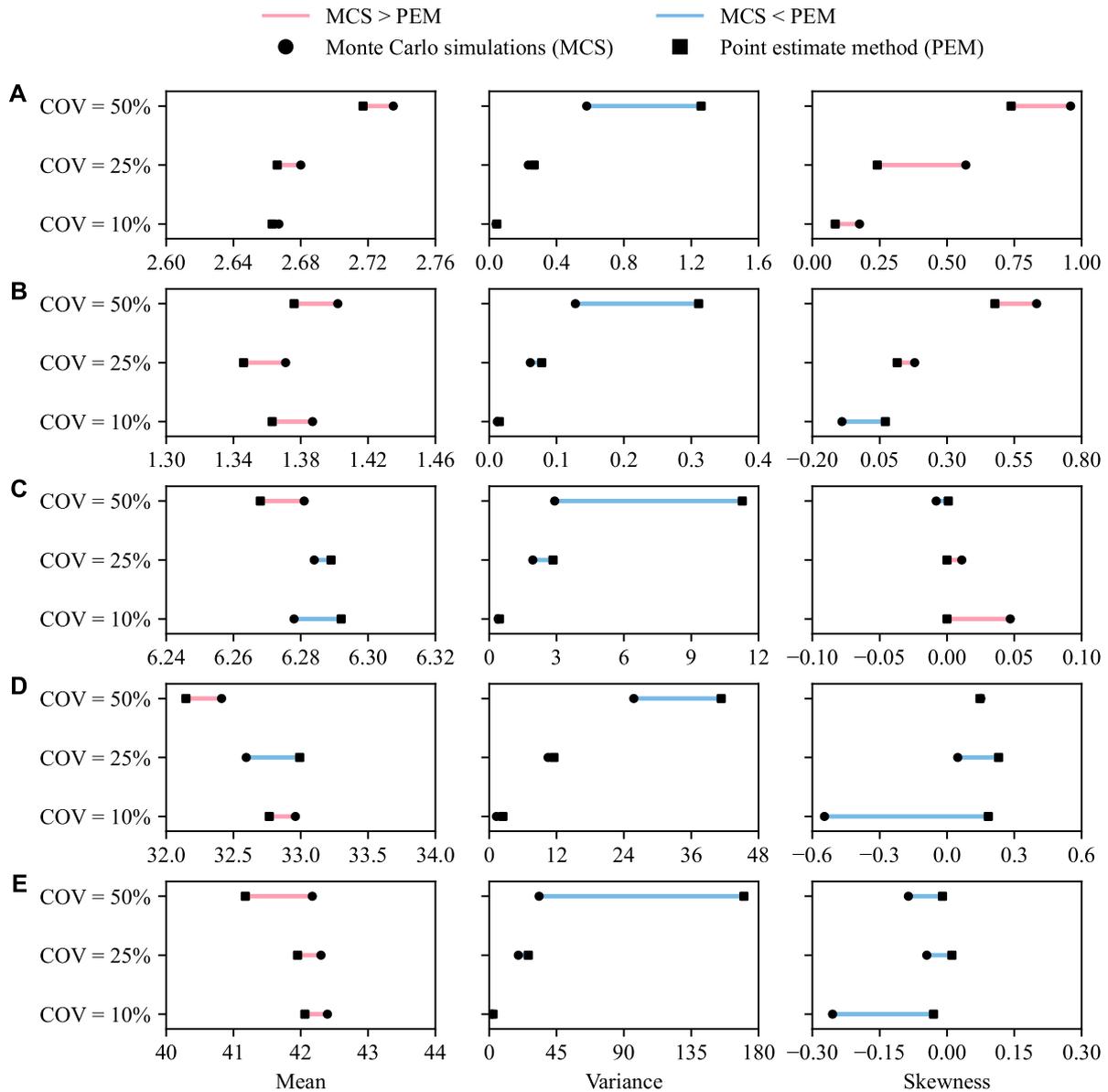
FIGURE 4
Three moments (mean, variance and skewness) of **(A)** impact area (in x10^6 m^2), **(B)** deposit area (in x10^6 m^2), **(C)** deposit volume (in x10^6 m^3), **(D)** maximum flow velocity (in m/s) and **(E)** maximum flow height (in m) at (x,y) = (1490100, 5204100) obtained from Acheron rock avalanche with Monte Carlo simulations (MCS) and point estimate method (PEM).

- the point-wise maximum flow velocity decreases with increasing dry-Coulomb friction coefficient and has little dependence on the release volume (see **Figures 5D**, **6D**),
- the point-wise maximum flow height increases with increasing release volume and has little dependence on the dry-Coulomb and turbulent friction coefficient (see **Figures 5E**, **6E**).

Differences between the two cases can be noted as (1) the deposit area has a clear negative relationship with the dry-Coulomb friction coefficient in Acheron rock avalanche, but the trend is hardly visible for the synthetic case; (2) the point-wise maximum

velocity increases with the turbulent friction coefficient in the synthetic case, but the relationship is vague in Acheron rock avalanche.

Figures 7, **8** show the comparison of spatial distribution of maximum flow height from synthetic case and Acheron rock avalanche, respectively. Results given in the figures, mean and standard deviation of h_{max} in each cell as well as their differences, are from the PEM and MCS analyses conducted at COV = 50%. A visual comparison of mean values (See **Figures 7A**, **B**, **8A**, **B**) look nearly identical, but the difference map shows that flow heights at the central section of the flow path in both cases are higher in MCS analysis, while
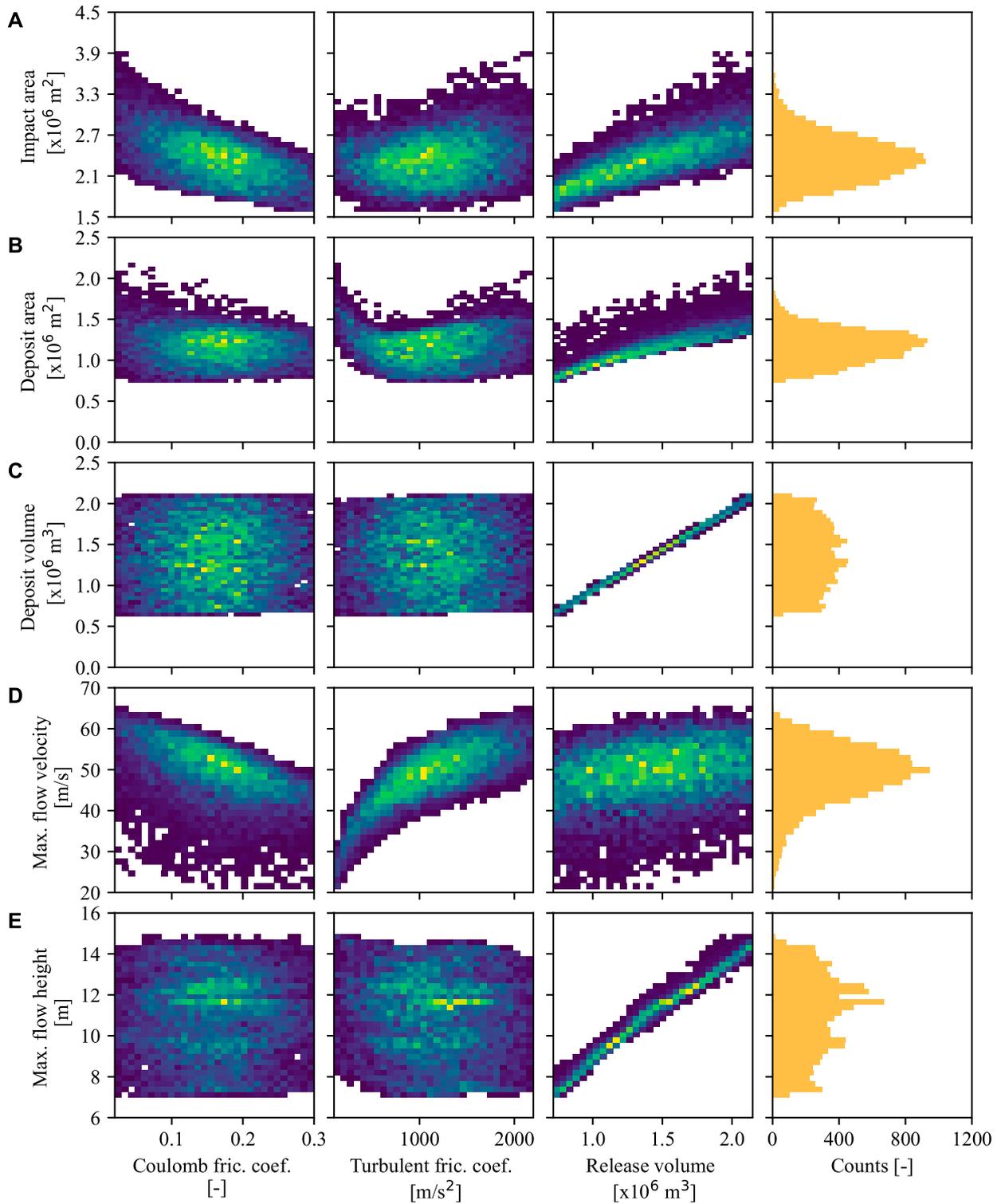
FIGURE 5
Relationships and histograms of **(A)** impact area, **(B)** deposit area, **(C)** deposit volume, **(D)** maximum flow velocity and **(E)** maximum flow height at $(x,y) =$ (1000, 2000) from synthetic case with uncertain parameters used in the Monte Carlo analysis with coefficient of variation of 50%.
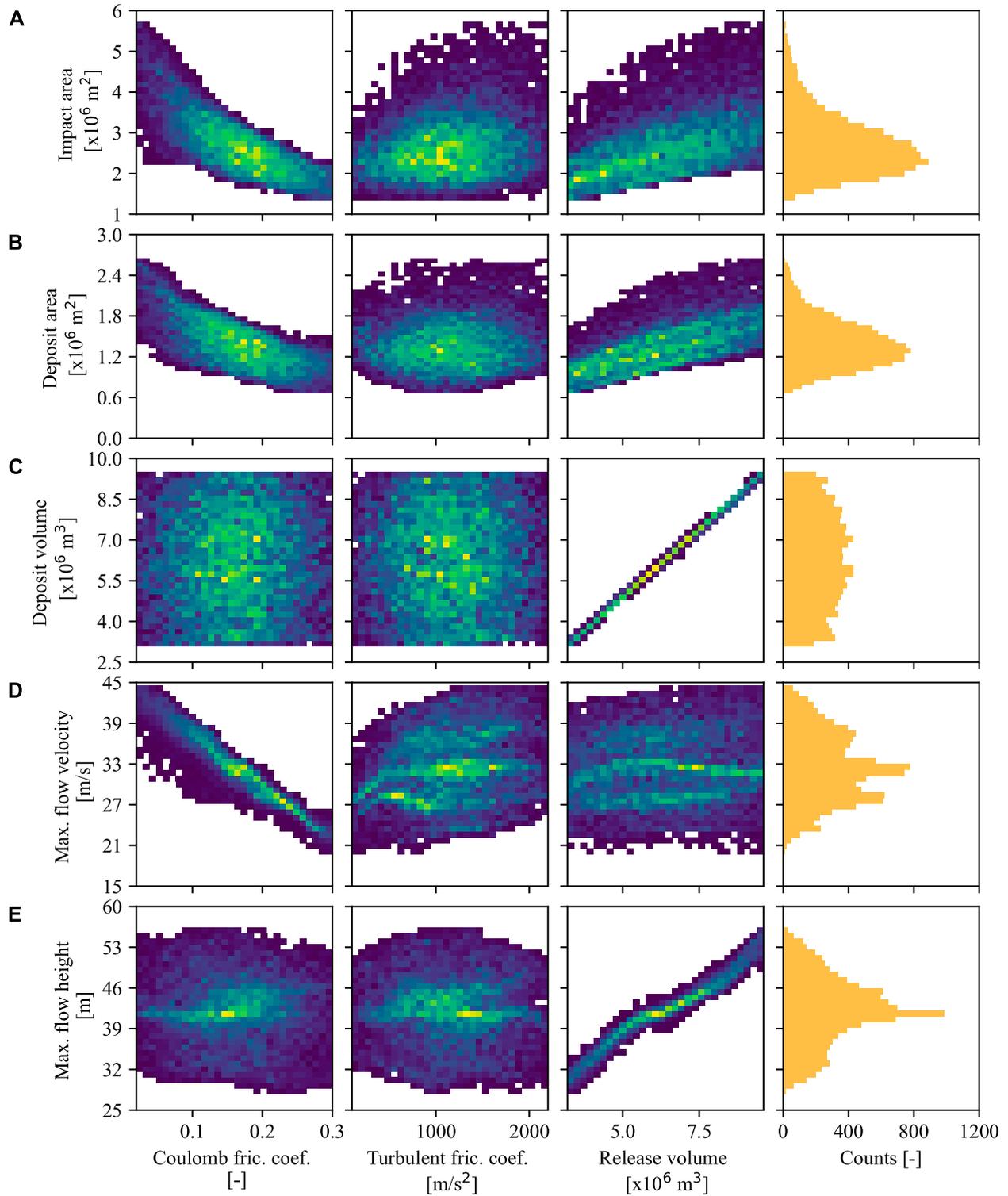
FIGURE 6
Relationships and histograms of **(A)** impact area, **(B)** deposit area, **(C)** deposit volume, **(D)** maximum flow velocity and **(E)** maximum flow height at (x,y) = (1490100, 5204100) from Acheron rock avalanche with uncertain parameters used in the Monte Carlo analysis with coefficient of variation of 50%.
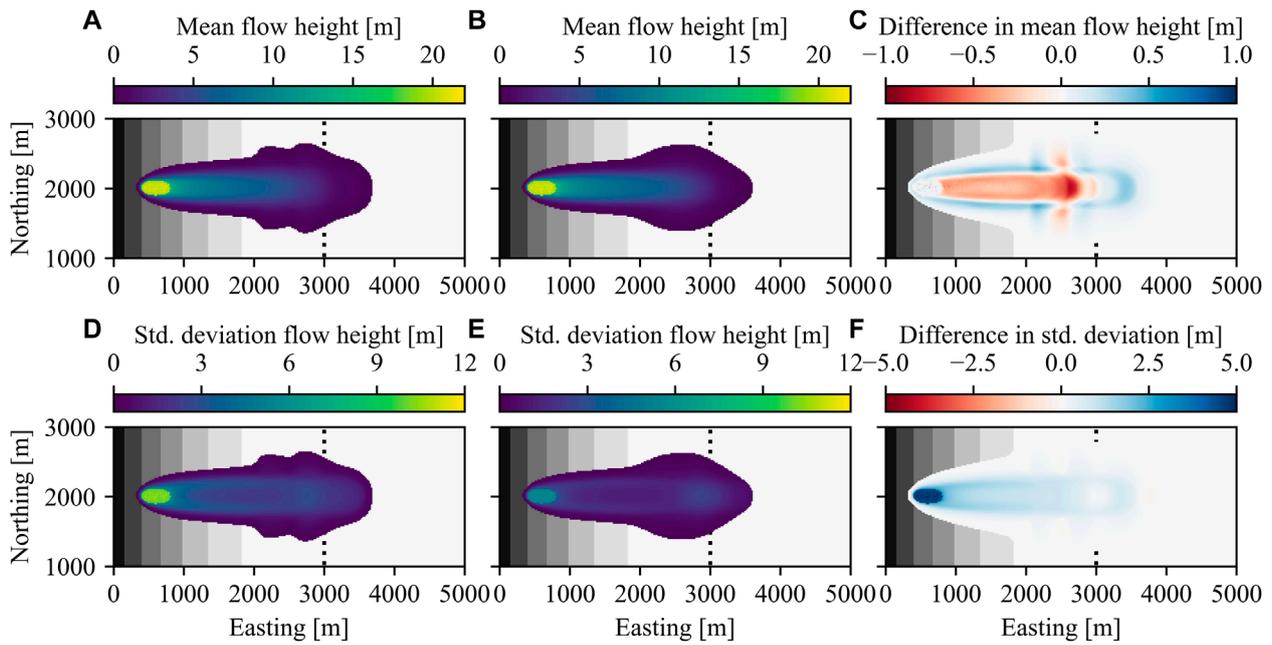
FIGURE 7
Spatial distribution of mean and standard deviations of maximum flow height (h_{max}) from synthetic case obtained with **(A, D)** Point Estimate Method (PEM) and **(B, E)** Monte Carlo Simulations (MCS) at 50% coefficient of variation. Differences of **(C)** mean value and **(F)** standard deviation are plotted by subtracting MCS results from PEM results.

TABLE 3 Correlation of input variables with location (x_L) and the length (L_{max}) of maximum lateral spread in synthetic case.

	x_L	L_{max}
Dry-Coulomb friction coefficient	.99***	.33***
Turbulent friction coefficient	No correlation	.50***
Release volume	No correlation	.80***

*** $p < .001$

the edges of the flow path have higher flow heights in the PEM analysis. Similar to the results in **Figures 3, 4**, MCS produced lower standard deviation than PEM at the majority of cells, i. e approximately 200 cells out of 6600 for synthetic case, and 830 cells out of 9800 cells for Acheron rock avalanche.

4 Discussion

Risk for a single landslide scenario has three main components: the landslide hazard, the exposure of the elements at risk, and their vulnerability. Conducting a quantitative risk analysis enables the researchers or practitioners to obtain the probability of a given level of loss and the corresponding uncertainties of these components (Corominas et al., 2014; Eidsvig et al., 2014). The hazard in a model-based landslide risk assessment is generally evaluated by simulating various scenarios with an underlying physics-based computational model. A probabilistic approach to

this assessment and the quantification of associated uncertainty inevitably create a computational bottleneck—especially for large-scale applications (Strauch et al., 2018; Jiang et al., 2022). One can opt for a simple technique to quantify uncertainty, i.e. PEM in this study, to reduce the number of simulations required, or use a technique necessitating high number of simulations, for instance MCS techniques.

The computational challenge of MCS has been well-recognised in the field of landslide modelling, which results from the large number of model realisations at randomly selected input values (Dalbey et al., 2008; McDougall, 2017). MCS can be very computationally intensive, since it typically requires tens of thousands of model runs to achieve reasonable accuracy (Salciarini et al., 2017). This is often not feasible in model-based landslide risk assessment, as a single model run may take minutes to hours. A solution to overcome this problem has been demonstrated in this study by utilising recent development of GP emulation (Gu and Berger, 2016; Gu et al., 2018; Gu et al., 2019). GP emulators are built for each case based on only 100 model runs. Then, MCS with 10,000 randomly generated inputs is conducted using the emulators, which means no further model runs are needed. Moreover, diagnostics of built GP emulators are analysed to evaluate their performance. High R^2 values and low *MAPE*s and *nRMSE*s (See **Table 2**) suggests that the scalar emulators can be used with confidence for predictions of a singular output from an input parameter combination. All $P_{CI}(95\%)$ values for the vector emulators are close to 95%, which justify their usage as a surrogate to the computational model (Gu et al., 2019). The corresponding results of GP emulation-based MCS are therefore close to results of a classical MCS, but the computational time is significantly reduced by introducing GP emulation. This demonstrates the
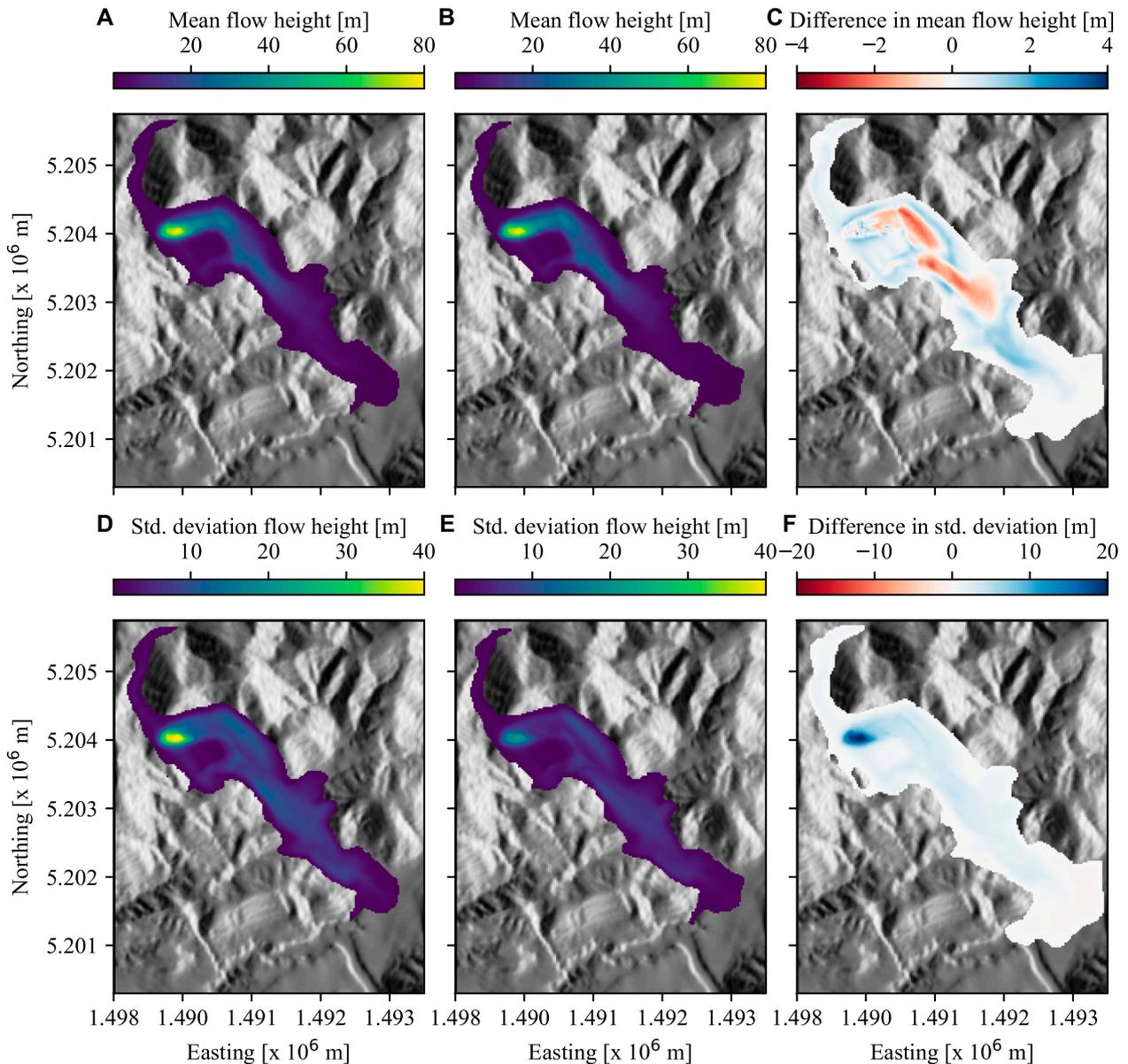
FIGURE 8
Spatial distribution of mean and standard deviations of maximum flow height (h_{max}) from Acheron rock avalanche obtained with **(A, D)** Point Estimate Method (PEM) and **(B, E)** Monte Carlo Simulations (MCS) at 50% coefficient of variation. Differences of **(C)** mean value and **(F)** standard deviation are plotted by subtracting MCS results from PEM results.

applicability of GP emulation for uncertainty quantification of landslide run-out models. A similar methodology applied to landslide generated waves was also found promising to perform probabilistic hazard analysis based on computationally intensive models (Snelling et al., 2020).

Comparative studies in landslide research showed that similar mean values of the QoI can obtained with PEM or MCS (Tsai et al., 2015; Przewlocki et al., 2019). As shown in **Figures 3**, **4**, when an aggregated (e.g. impact area, deposit area, deposit volume) or point-wise (velocity and height at a predetermined coordinate) output is calculated, PEM and MCS yielded similar expectancies (mean values) in this study. In addition, both PEM and MCS lead to

similar variance values for relatively low COV, i.e. 10% and 25%. This implies that if one only aims at computing low-order moments in comparative topographic settings, PEM can achieve reasonable results (Fanelli et al., 2018). PEM is particularly computationally appealing for low-dimensional problems due to the requirement of 2^n realisations, where n is the dimension of the input parameter space. However, as pointed out by Christian and Baecher (2002), caution should be used in approximating skewness or other higher order moments based on PEM. This is supported by the large difference between PEM- and MCS-based skewness results as shown in **Figures 3**, **4**. Christian and Baecher (2002) also pointed out that the results deviate significantly, when the *COV* of uncertain inputs is

large, which is confirmed by the large difference between the variance computed by PEM and MCS in the cases of $COV = 50\%$ as shown in **Figures 3**, **4**. Comparing **Figures 3**, **4**, it can be seen that the complexity of topography, which refers to the real topography of Acheron rock avalanche in comparison to the parabolic slope of the synthetic case, seems to have limited impact on the general trend of moments based on PEM and MCS, especially mean and variance. It may imply the transferability of the trends described above to different topographies.

When input uncertainty is high and the input parameter space is high-dimensional, MCS is clearly favourable over PEM to compute desired statistics of the output of interest. The benefit of GP-integrated MCS is that it can not only compute the desired statistics, but it also provides the PDF (Marin and Mattos, 2020). Cepeda et al. (2013) recommends a stochastic approach for uncertainty quantification such as MCS to be part of the routine of any landslide hazard risk assessment, and Hussin et al. (2012) denotes the frequency distributions of model outputs important as a first step to assess the spatial probability in future debris flow hazard assessments. The workflow of GP emulation-based uncertainty quantification in this study is therefore expected to improve landslide risk assessment. It should be noted that the number of model runs to train GP emulators are effected by the dimension of input parameter space even though the number of model runs of MCS is independent. Due to this limitation, the gain of computational efficiency by using GP emulation decreases with increasing dimension of input parameter space. If the dimension is too high, other emulation techniques or dimension reduction can be considered (Liu and Guillas, 2017).

Histograms plotted in **Figures 5**, **6** can be used to infer on the complexity of the output functions, see Eqs. **9**. Input parameters were assumed to be normally distributed in MCS analysis. It can be seen that the deposit volume (see **Figures 5C**, **6C**) is the only parameter that has the shape of the truncated normal distribution with nearly no skewness. Other scalar outputs result in skewed and even bi-modal distributions that clearly deviate from the initial Gaussian distribution of the corresponding input parameter. Both cases show a nearly perfect correlation between deposit and release volume with no effects of the friction coefficient indicating parameter-independent mass conservation. It should be noted that the deposit volume was calculated considering cells in which height exceeded .1 m at the last simulation time step. As there was no stopping criteria defined and no entrainment was considered, the deposit volume was nearly equal to the release volume. Highly linear function of deposit volume and release volume translates into the deposit volume having a Gaussian distribution, as the linear transformation of a Gaussian distribution is also a Gaussian distribution. However, the non-linearity of the other functions that can be used the express the scalar outputs—except deposit volume—results in distributions different than those of the input parameters. For example, Cepeda et al. (2013) fitted Gamma distributions for flow height and velocity in two different cases.

Patterns observed in **Figures 7C**, **8C** can be explained by analysing the effects of input variables on the general shape of the flow path. Hence, the simulation in a synthetic topography is an ideal example due to its simplicity. As explained in **Section 3.1** and shown in **Figure 1A**, the flow path of the synthetic case can be described as a concentrated central flow superposed by lateral spread at the toe of the slope. It can be seen in **Figure 7A** that the mean flow height has

two pronounced *dents* in its spatial distribution obtained with PEM. These correspond to the initiation of the lateral spread at different configurations of the input variables are considered. More specifically, PEM analysis is run only at few discrete values in parameter space chosen at a distance of one standard deviation away from the mean. When the friction coefficient is chosen at a $COV = 50\%$ with a mean of .16, PEM simulations are characterised by a lateral spread very early in the flow path, or a lateral spread that kicks in much further downstream. Therefore, the *dents* in the maximum flow height map (See **Figure 7A**) is a direct consequence of the coarse discretisation of the parameter space in the PEM approach. In contrast, MCS yields a homogeneously distributed maximum flow height map at the toe of the slope as expected in this almost linear setting. As a consequence, flow heights at the upper sections of the lateral spread are higher in PEM, whereas MCS yields higher values at the mid-section of the lateral spread (See **Figure 7C**).

To recognise patterns between the location (x_L) and the magnitude (L_{max}) of the maximum lateral spread with the input variables, linear regression analysis was conducted. **Table 3** shows that the location is controlled dominantly by the dry-Coulomb friction coefficient with a negative correlation, i.e. higher the friction coefficient earlier the lateral spread starts, and how much the flow spreads in y-axis is mostly controlled by the release volume, even though the friction coefficients affect to a certain extent.

5 Conclusion

Uncertainty quantification is a computationally demanding task for designing and developing a model-based landslide risk assessment. Classical MCS is often computationally infeasible due to the large number of required forward evaluations of the computational model. It has been demonstrated that GP emulation-based MCS can greatly improve the computational efficiency which makes GP-integrated MCS applicable for landslide run-out modelling. One clear advantage of using GP emulation-based MCS is the ability to sample parameter uncertainty in a dense way, as evaluation time of the forward simulation is no longer a computational bottleneck. As a consequence, the output's probability distribution reflecting the propagated uncertainty is captured at high accuracy and provides additional information about skewness and possible multi-modality. In contrast to this PEM provides only limited information on the output's probability distribution. A comparative study between PEM and GP emulation-based MCS has been conducted based on the three moments of the probability distribution, i.e. mean, variance, and skewness. The simpler approach, PEM, yielded a similar expectancy values to GP emulation-based MCS. However, PEM and MCS differed in higher order moments, such as variances and skewness, hence also in the respective spatial distribution of the flow path, and the subsequent hazard map. This finding is of high practical relevance: While a computationally cheap PEM based workflow predicts the mean of a probabilistic landslide risk assessment well, it is in general cases not suitable to assess the reliability of the prediction, for instance in the sense of a probabilistic simulation's standard deviation. The latter requires a MCS approach, which often is computationally infeasible. GP-emulated MCS overcome this limitations by introducing a surrogate model trained based on an empirical error control. It can be suggested that highly uncertain and

high-dimensional input parameter spaces, e.g. complex topographies, advanced material models, models with empirical parameters, inevitably requires an uncertainty quantification workflow that is able to account for non-Gaussian, potentially multi-modal distributions. It should be noted that the gain in compute time by using GP emulation critically depends on the computational effort needed to train the GP emulator. This means that computational resources significantly increase, as the dimension of the input parameter space increases. Alternative techniques will have to be incorporated if the input dimension is too high.

Data availability statement

The datasets generated for this study can be found on Figshare as a collection titled "Uncertainty quantification in landslide runout simulations" (doi: 10.6084/m9.figshare.c.6172702). General workflow can be accessed as a Python package in a compressed format hosted on Figshare (doi: 10.6084/m9.figshare.20730871), also as a Git repository (https://github.com/yildizanil/frontiers_yildizetal).

Author contributions

Contributions are defined according to CRediT author statement. AY: Conceptualisation, Methodology, Software, Formal analysis, Investigation, Writing—Original Draft, Writing—Review and Editing, Visualization. HZ: Methodology, Software, Formal analysis, Investigation, Writing—Original Draft, Writing—Review and Editing. JK: Conceptualisation, Methodology, Writing—Review and Editing, Supervision, Funding acquisition.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aaron, J., McDougall, S., Kowalski, J., Mitchell, A., and Nolde, N. (2022). Probabilistic prediction of rock avalanche runout using a numerical model. *Landslides* 19, 2853–2869. doi:10.1007/s10346-022-01939-y

Aaron, J., McDougall, S., and Nolde, N. (2019). Two methodologies to calibrate landslide runout models. *Landslides* 16, 907–920. doi:10.1007/s10346-018-1116-8

Brezzi, L., Carraro, E., Gabrieli, F., Santa, G. D., Cola, S., and Galgaro, A. (2021). Propagation analysis and risk assessment of an active complex landslide using a Monte Carlo statistical approach. *IOP Conf. Ser. Earth Environ. Sci.* 833, 012130. doi:10.1088/1755-1315/833/1/012130

Cepeda, J., Luna, B. Q., and Nadim, F. (2013). "Assessment of landslide run-out by Monte Carlo simulations," in *Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering* (Paris, France: Presses des Ponts), 2157.

Christian, J. T., and Baecher, G. B. (1999). Point-estimate method as numerical quadrature. *J. Geotechnical Geoenvironmental Eng.* 125, 779–786. doi:10.1061/(asce)1090-0241(1999)125:9(779)

Christian, J. T., and Baecher, G. B. (2002). The point-estimate method with large numbers of variables. *Int. J. Numer. Anal. Methods Geomechanics* 26, 1515–1529. doi:10.1002/nag.256

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J.-P., Fotopoulou, S., et al. (2014). Recommendations for the quantitative analysis of landslide risk. *Bull. Eng. Geol. Environ.* 73, 209–263. doi:10.1007/s10064-013-0538-8

Dalbey, K., Patra, A. K., Pitman, E. B., Bursik, M. I., and Sheridan, M. F. (2008). Input uncertainty propagation methods and hazard mapping of geophysical mass flows. *J. Geophys. Res.* 113, B05203. doi:10.1029/2006JB004471

D'Amboise, C. J. L., Neuhauser, M., Teich, M., Huber, A., Kofler, A., Perzl, F., et al. (2022). Flow-py v1.0: A customizable, open-source simulation tool to estimate runout and intensity of gravitational mass flows. *Geosci. Model Dev.* 15, 2423–2439. doi:10.5194/gmd-15-2423-2022

Eidsvig, U., Papathoma-Köhle, M., Du, J., Glade, T., and Vangelsten, B. (2014). Quantification of model uncertainty in debris flow vulnerability assessment. *Eng. Geol.* 181, 15–26. doi:10.1016/j.enggeo.2014.08.006

Fanelli, G., Salciarini, D., and Tamagnini, C. (2018). "A comparison between probabilistic approaches for the evaluation of rainfall-induced landslide susceptibility at regional scale," in *Landslides and engineered slopes. Experience, theory and practice*. Editors S. Aversa, L. Cascini, L. Picarelli, and C. Scavia (CRC Press), 879. doi:10.1201/9781315375007-93

Fenton, G. A., and Griffiths, D. V. (2008). *Risk assessment in geotechnical engineering*. John Wiley & Sons.

Fischer, J.-T., Kofler, A., Fellin, W., Granig, M., and Kleemayr, K. (2015). Multivariate parameter optimization for computational snow avalanche simulation. *J. Glaciol.* 61, 875–888. doi:10.3189/2015JoG14J168

Gu, M. Y., and Berger, J. O. (2016). Parallel partial Gaussian process emulation for computer models with massive output. *Ann. Appl. Statistics* 10, 1317–1347. doi:10.1214/16-AOAS934

Gu, M. Y., Palomo, J., and Berger, J. O. (2019). Robustgasp: Robust Gaussian stochastic process emulation in R. *R J.* 11, 112–136. doi:10.32614/RJ-2019-011

Gu, M. Y., Wang, X. J., and Berger, J. O. (2018). Robust Gaussian stochastic process emulation. *Ann. Statistics* 46, 3038–3066. doi:10.1214/17-AOS1648

Hussin, H. Y., Luna, B. Q., van Westen, C. J., Christen, M., Malet, J.-P., and van Asch, T. W. J. (2012). Parameterization of a numerical 2-d debris flow model with entrainment: A case study of the faucon catchment, southern French alps. *Nat. Hazards Earth Syst. Sci.* 12, 3075–3090. doi:10.5194/nhess-12-3075-2012

Jiang, S.-H., Huang, J., Griffiths, D., and Deng, Z.-P. (2022). Advances in reliability and risk analyses of slopes in spatially variable soils: A state-of-the-art review. *Comput. Geotechnics* 141, 104498. doi:10.1016/j.compgeo.2021.104498

Kaynia, A., Papathoma-Köhle, M., Neuhäuser, B., Ratzinger, K., Wenzel, H., and Medina-Cetina, Z. (2008). Probabilistic assessment of vulnerability to landslide: Application to the village of lichtenstein, baden-württemberg, Germany. *Eng. Geol.* 101, 33–48. doi:10.1016/j.enggeo.2008.03.008

Liu, X., and Guillas, S. (2017). Dimension reduction for Gaussian process emulation: An application to the influence of bathymetry on tsunami heights. *SIAM/ASA J. Uncertain. Quantification* 5, 787–812. doi:10.1137/16M1090648

Liu, X., Wang, Y., and Li, D. Q. (2019). Investigation of slope failure mode evolution during large deformation in spatially variable soils by random limit equilibrium and material point methods. *Comput. Geotechnics* 111, 301–312. doi:10.1016/j.compgeo.2019.03.022

Ma, G. T., Rezania, M., Nezhad, M. M., and Hu, X. W. (2022). Uncertainty quantification of landslide runout motion considering soil interdependent anisotropy and fabric orientation. *Landslides* 19, 1231–1247. doi:10.1007/s10346-021-01795-2

Marin, R. J., and Mattos, Á. J. (2020). Physically-based landslide susceptibility analysis using Monte Carlo simulation in a tropical mountain basin. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* 14, 192–205. doi:10.1080/17499518.2019.1633582

McDougall, S. (2017). 2014 Canadian geotechnical colloquium: Landslide runout analysis — Current practice and challenges. *Can. Geotechnical J.* 54, 605–620. doi:10.1139/cgj-2016-0104

Mergili, M., Fischer, J.-T., Krenn, J., and Pudasaini, S. P. (2017). r.avaflow v1, an advanced open-source computational framework for the propagation and interaction of two-phase mass flows. *Geosci. Model Dev.* 10, 553–569. doi:10.5194/gmd-10-553-2017

Przewlocki, J., Zabuski, L., and Winkelmann, K. (2019). Reliability analysis of sea cliff slope stability by point estimate method. *IOP Conf. Ser. Mater. Sci. Eng* 471, 042003. doi:10.1088/1757-899X/471/4/042003

Quan Luna, B., Cepeda, J., Stumpf, A., van Westen, C. J., Remaître, A., Malet, J., et al. (2013). Analysis and uncertainty quantification of dynamic run-out model parameters for landslides, in *Landslide science and practice: Volume 3: Spatial analysis and modelling*. Editors C. Margottini, and P. Canuti, doi:10.1007/978-3-642-31310-3_42

Salciarini, D., Fanelli, G., and Tamagnini, C. (2017). A probabilistic model for rainfall—Induced shallow landslide prediction at the regional scale. *Landslides* 14, 1731–1746. doi:10.1007/s10346-017-0812-0

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

Smith, G., Bell, D., and Davies, T. (2012). The acheron rock avalanche deposit, canterbury, New Zealand: Age and implications for dating landslides. *N. Z. J. Geol. Geophys.* 55, 375–391. doi:10.1080/00288306.2012.733947

Smith, G. M., Davies, T. R., McSaveney, M. J., and Bell, D. H. (2006). The acheron rock avalanche, canterbury, New Zealand—Morphology and dynamics. *Landslides* 3, 62–72. doi:10.1007/s10346-005-0012-1

Snelling, B., Neethling, S., Horsburgh, K., Collins, G., and Piggott, M. (2020). Uncertainty quantification of landslide generated waves using Gaussian process emulation and variance-based sensitivity analysis. *Water* 12, 416. doi:10.3390/w12020416

Strauch, R., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., and Tucker, G. E. (2018). A hydroclimatological approach to predicting regional landslide probability using landlab. *Earth Surf. Dyn.* 6, 49–75. doi:10.5194/esurf-6-49-2018

Sun, X. P., Zeng, P., Li, T. B., Wang, S., Jimenez, R., Feng, X. D., et al. (2021b). From probabilistic back analyses to probabilistic run-out predictions of landslides: A case study of heifangtai terrace, gansu province, China. *Eng. Geol.* 280, 105950. doi:10.1016/j.enggeo.2020.105950

Sun, X., Zeng, P., Li, T., Zhang, T., Feng, X., and Jimenez, R. (2021a). Run-out distance exceedance probability evaluation and hazard zoning of an individual landslide. *Landslides* 18, 1295–1308. doi:10.1007/s10346-020-01545-w

Svalova, A., Helm, P., Prangle, D., Rouainia, M., Glendinning, S., and Wilkinson, D. J. (2021). Emulating computer experiments of transport infrastructure slope stability using Gaussian processes and bayesian inference. *Data-Centric Eng.* 2, e12. doi:10.1017/dce.2021.14

Tsai, T.-L., Tsai, P.-Y., and Yang, P.-J. (2015). Probabilistic modeling of rainfall-induced shallow landslide using a point-estimate method. *Environ. Earth Sci.* 73, 4109–4117. doi:10.1007/s12665-014-3696-5

Wilhelm, S., and Manjunath, B. G. (2010). Tmvtnorm: A package for the truncated multivariate normal distribution. *R J.* 2, 25–29. doi:10.32614/RJ-2010-005

Yildiz, A., Zhao, H., and Kowalski, J. (2022a). Uncertainty quantification in landslide run-out simulations. doi:10.6084/m9.figshare.c.6172702

Yildiz, A., Zhao, H., and Kowalski, J. (2022b). Uncertainty quantification workflow in landslide risk. doi:10.6084/m9.figshare.20730871

Zeng, P., Sun, X. P., Xu, Q., Li, T. B., and Zhang, T. L. (2021). 3D probabilistic landslide run-out hazard evaluation for quantitative risk assessment purposes. *Eng. Geol.* 293, 106303. doi:10.1016/j.enggeo.2021.106303

Zhao, H., Amann, F., and Kowalski, J. (2021). Emulator-based global sensitivity analysis for flow-like landslide run-out models. *Landslides* 18, 3299–3314. doi:10.1007/s10346-021-01690-w

Zhao, H., and Kowalski, J. (2022). Bayesian active learning for parameter calibration of landslide run-out models. *Landslides* 19, 2033–2045. doi:10.1007/s10346-022-01857-z

Zhao, H., and Kowalski, J. (2020). Topographic uncertainty quantification for flow-like landslide models via stochastic simulations. *Nat. Hazards Earth Syst. Sci.* 20, 1441–1461. doi:10.5194/nhess-20-1441-2020