Probabilistic hazard assessment using the computer model of a volcano pyroclastic flow whose output is zero-inflated

Ksenia N. Kyzyurova
http://www.kseniak.ucoz.net, ksenia.kyzyurova@gmail.com

November 1, 2020

Abstract

Stochastic approximation to a computer model generally considers the smooth output of a model. Sometimes a computer model produces a non-smooth, zero-inflated output. Computer model TITAN2D, given a set of initial conditions, produces an output, height of a volcano pyroclastic flow at several thousands of spatial locations in a geographical region of interest. This output is non-negative and often results in exact zero (indicating the absence of a flow). In order to account for a large number of zero values in the non-negative output, a censored Gaussian stochastic process approximation to a computationally expensive computer model is proposed. Subsequent probabilistic assessment of a hazard using the proposed methodology is given in comparison to probabilistic assessment by other methods found in the literature. The corresponding difference in hazard estimates appears to be dramatic, while censoring is found to be adequate in its assessment of a probability of a hazard.

Elements of the mathematical methodology of a truncated projected GASP of a truncated output of a computer model are presented in the appendix. Mathematical foemulae of linking a Gaussian process emulator with either censored or truncated emulator in a sequence of two emulators are provided.

1 Introduction

Computer model TITAN2D of a volcano pyroclastic flow (Patra et al. 2005) produces output, the height of the flow, at a set of geographical locations within a single run of the model. A single run of the model is determined by a set of initial values of parameters supplied to a computer model for this run. The goal in this work is to emulate the maximum height of a pyroclastic flow, a scalar output, at a set of spatial locations over a three-dimensional input space to a computer model. Three inputs to TITAN2D simulator are the volume of a pyroclastic flow, basal friction angle and initial direction angle.

Post-processed data output for a single run of the model is comprised of a total of 24,576 locations associated with the island of Montserrat. These locations are given on a grid 128×192. Computer model has been run 500 times at various initial values.

Convenient is to consider that a computer model produces a smooth output of the model (Loeppky et al. 2009). In current work, however, the output of the model is recorded as exact zero values of the pile height of the flow for all 500 runs at 8,491 locations, which constitutes about one third of all the locations on the grid.

For the rest of 16,085 locations within all 500 computer model runs, two thirds of the outputs resulted in exact zero height values. The distribution of other, non-zero height values of the pyroclastic flow, is shown in Figure 1.

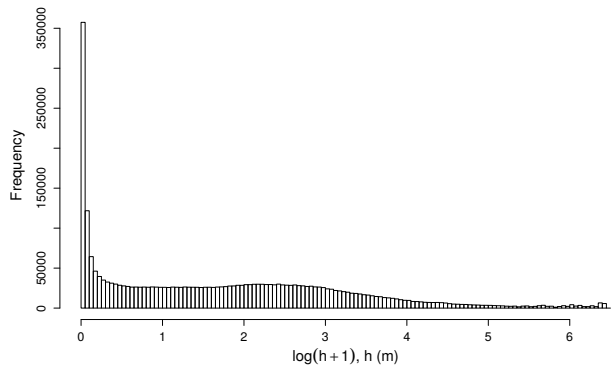

Figure 1: Distribution of non-zero height values of the pyroclastic flow on a transformed $\log(h+1)$ scale, where h stands for the maximum height of the pyroclastic flow (in meters).

The figure reveals that a problem of emulation of non-negative output together with dominated zero-value output accompanied by a large number of small-height values emerges in this framework.

Exact zeros of the height of a pyroclastic flow are not simply artefacts of a computer model. Zero-height inflation appears due to the computational handling of wet-dry conditions of the flow (Aghakhani et al. 2016). For TITAN2D unrealistically small height, that is, spurious small non-zero numerical values of the flow produced by the model, are converted to zeros.

The problem of large volume of zero output appears in other areas related to, for example, inundation of flows. Gopinathan et al. (2017) considered a computer model of tsunami to solve a Bayesian inverse problem related to this model. The model produces tsunami wave height outputs which vary in their pattern depending on a geographical location (Gopinathan 2018). In certain regions there is no flow under a particular choice of initial conditions, thus creating large zero-flow output of a computer model. Similar output structure happens with the model ADCIRC (Westerink et al. 1993, Asher 2018).

Several possibilities exist on how to proceed with emulation of a zero-inflated output of a simulator:

(a) *Projected traditional GASP without (w/o) zeros.* First possibility is to simply ignore zero-output information. Traditional GASP emulator (Bayarri et al. 2009) is trained only at points which result in positive height of the flow. The information carried by points with zero-height output is thrown away. This approach is considered for illustrative purposes, but this approach is usually unappealing because important data is ignored.

(b) *Projected traditional GASP with (w/) zeros.* Second possibility is to treat zero-inflated output data as regular points of a smooth function, so that this data is not discerned from positive-height data points and are included in construction of a traditional GASP emulator of a computer model.

The zero-problem has been acknowledged before (Spiller et al. 2014) and treated through a clever use of the two described possibilities. Authors propose the algorithmic approach to construction of a GASP emulator via eliminating "non-important" zeros which are far away from non-zero outputs and leaving only a few "important" zero values which surround computer model data points with positive height of the flow according to a chosen metric between inputs of data points. While the proposed approach is computationally and algorithmically appealing, the result is also necessarily dependent on a chosen metric.

(c) *Censored GASP.* The third possibility — to account properly for zero-inflated computer model output by constructing a censored GASP emulator of a computer model — is proposed in this work. The mathematical methodology for construction of the censored emulator has been outlined in (Kyzyurova 2017, 2019*a*).

Traditional GASP model assumes that the underlying function from a simulator is smooth. However, this assumption is inadequate for modeling zero-inflated output of the TITAN2D computer model for two reasons. First, *a priori* flow height values are non-negative, causing inherent restriction on the range of values of the output of a computer model. Second, zero-height value of an output has a *non-zero* probability to occur, as opposed to all other output values of this simulator. The proposed method is to model height of a pyroclastic flow as censored at zero non-observed (hidden) output of a computer model, with the traditional GASP of a computer model being an approximation to an imaginary latent model of an output. Censoring allows to account for both, non-negative nature of the output of the model and positive probability of occurrence of a zero-height.

The methodology proposed here is in line with methods employed for emulation of computer models with different types of constraints on the behaviour of an underlying function or its derivatives, e.g. (Wang & Berger 2016, Maatouk & Bay 2017).

(d) Last possibility which has been thought over is transformation of the output. While generally this is a useful tool employed in statistical modeling, the limitation of this approach (when employed together with GASP emulation) is very unattractive. Undesirable properties of the often used log-transformation (suggested (O'Hagan 2006) and used (Bayarri et al. 2007)) have been demonstrated (Kyzyurova 2019*c*).

The rest of the manuscript is concerned with the demonstration of the mathematical methodology and comparison of the three potential approaches: projected traditional GASP w/o zeros — method (a), projected traditional GASP w/ zeros — method (b), and censored GASP — method (c). The last method is concluded to be the only one that provides reasonable estimates of the probability of a hazard while either of the projected emulators do not. TITAN2D case study concludes the main part of the manuscript.

2 Illustrative example

Function $f(x) = 3x + \cos(5x)$ is considered in the domain $x \in [-1, 1]$ censored to the output range $[a = 0, \infty)$, that is

$$f_a(x) = \max\{a, f(x)\}\,,$$

and is used as a simulator. For the construction of any Gaussian stochastic process emulators (GASPs), they are parametrically specified with *a priori* linear mean function $\mu(\cdot)$ and squared-exponential correlation function $c(\cdot, \cdot)$. Parameters of GASPs are estimated following partially Bayesian approach presented in (Kyzyurova et al. 2018) which follows the procedure given in (Gu et al. 2018).

Two sets of input-output data points, uncensored and censored, are $\mathbf{x}^O = (0.2, 0.4, 0.6, 0.8)$, $f_a(\mathbf{x}^O) = \{f(x_i^O)\}_{i=1}^4$ and $\mathbf{x}^C = (-1.0, -0.8, -0.6, -0.4, -0.2)$, $f_a(\mathbf{x}^C) = \mathbf{0}$ (a 5-dimensional vector of zeros). Method (a) relies on construction of the approximation to the function $f_a(x)$, namely, traditional GASP w/o zeros,

$$f^{M_a}(\cdot) \mid \cdot, \mathbf{x}^O, f(\mathbf{x}^O) \sim \mathcal{GASP}(\mu^{M_a}(\cdot), \sigma^{2^{M_a}}(\cdot, \cdot))\,, \tag{1}$$

the GASP which provides the estimate of the probability at any new point of interest conditional on the uncensored observations only.

Method (b) relies on the approximation

$$f^{M_b}(\cdot) \mid \cdot, (\mathbf{x}^O, \mathbf{x}^C), (f(\mathbf{x}^O), f_a(\mathbf{x}^C)) \sim \mathcal{GASP}(\mu^{M_b}(\cdot), \sigma^{2^{M_b}}(\cdot, \cdot))\,, \tag{2}$$

the GASP which provides the estimate of the probability at any new point conditional on both, the uncensored and censored output from a computer model.

The GASP (1) is shown in the left panel of Figure 2. This GASP does not account for censored data anyhow. The right panel of Figure 2 shows the GASP (2), for which inputs $\mathbf{x}^C = (-1.0, -0.8, -0.6, -0.4, -0.2)$ with censored outputs $f_a(\mathbf{x}^C) = \mathbf{0}$ are included as if these are true data points from the simulator f. Figure 2

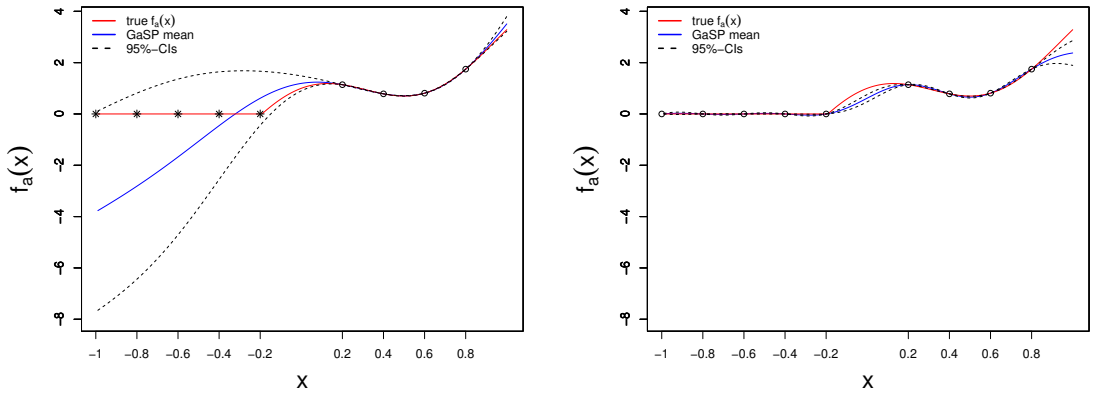

Figure 2: Left: Method (a). Latent GASP of a function $f_a(x)$. Right: Method (b). GASP with additional $x^C = \mathbf{0}$ treated as actual data. Circled points correspond to training inputs which were used to fit the emulator. Star points correspond to censored observations.

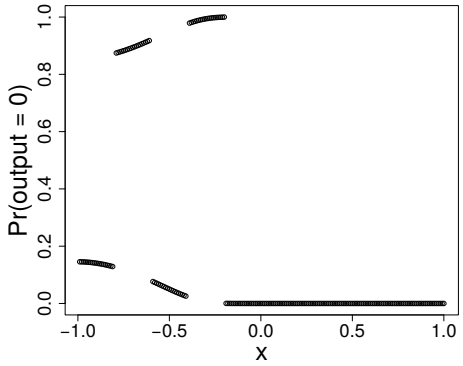

Figure 3: Unreasonable probabilities of a zero at testing points assigned by the projected traditional GASP which has zero-output points included in its training set.

shows that this GASP provides very tight credible intervals in the area where censored observations occurred. Not surprising, this GASP mean follows the output of the simulator well in the censored area, since it follows the zero region well. In the noncensored area emulator and its 95% credible area misses true function for $x \in (-0.2, 0.2)$ with large discrepancy between the mean and the true function, and for $x > 0.2$ emulator also misses original function, though it follows along it very closely.

As soon as Gaussian stochastic approximations to the computer model are constructed, probabilistic hazard estimate may be obtained as the mass of the corresponding predictive (normal) distribution $p(\cdot \mid \mu^*(\cdot), \sigma^{*2}(\cdot))$ parameterized by its mean and variance being greater than zero

$$P(f_{a=0}(\cdot) = 0) = P(f(\cdot) < 0) = \int_{-\infty}^{0} p(\cdot \mid \mu^*(\cdot), \sigma^{*2}(\cdot)) d\cdot \ . \tag{3}$$

For comparison of the method (a) and method (b) probabilities of a zero-height are given for an arbitrary chosen point $x_1^O = -0.2$. Using the posterior predictive distribution of the GASP and relying on the estimation (3), method (a) gives the probability of a simulator output being positive at the point $x_1^O = -0.2$ as 87%; and probability of a simulator output being exact zero at this point — only 13%. This unappealing estimation is illustrated in the left panel of Figure 2: instead, what we would like to see is that probability of a zero at the point $x_1^O = -0.2$ is close to 1.

Likewise, method (b) allows for calculation of a zero-height probability using (3) based on the *projected*

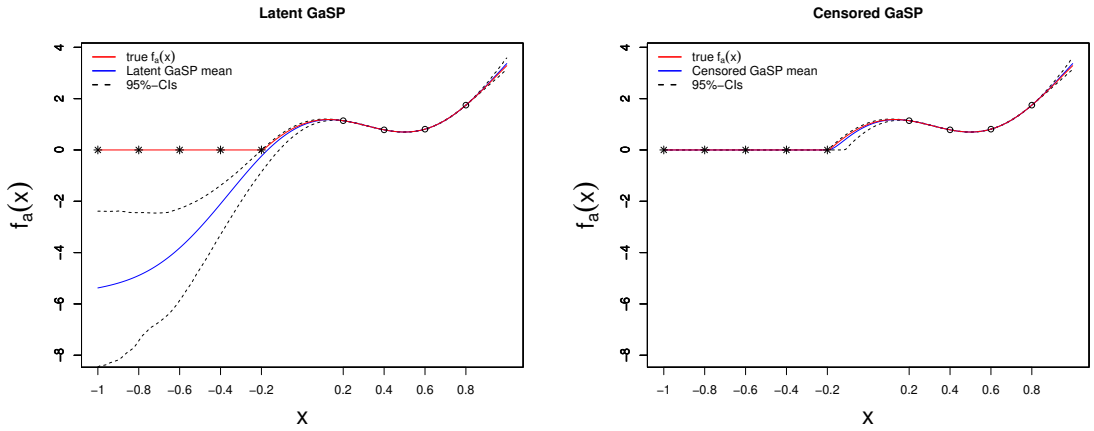

Figure 4: Method (c). Left: latent GASP. Right: the resulting censored GASP.

posterior of the GASP obtained by this method. Figure 2 shows that the probabilities of a zero-output do not correspond to adequate beliefs: four regions in the censored area are alternating in their probability estimates. Regions of input points in censored area have probabilities of a zero-output around 80-99%, while immediate nearby regions may have probability of a zero-output around 0-20%. The unappealing nature of this estimate is in that, instead, we would like to see a rather *smooth* function of the probability of a hazard.

Method (c) constructs a latent GASP and a censored GASP $f_a^{M_c}(\cdot)$ for function $f_a(x)$ which is censored at $a = 0$ and allows for positive values

$$f_a^{M_c}(\cdot) \mid \cdot, (\mathbf{x}^O, \mathbf{x}^C), f(\mathbf{x}^O), f(\mathbf{x}^C) < \mathbf{0}\,,$$

whose predictive distribution at point x' is

$$f_a^{M_c}(x') \mid x', (\mathbf{x}^O, \mathbf{x}^C), f(\mathbf{x}^O), f(\mathbf{x}^C) < \mathbf{0} = \begin{cases} f^{M_c}(x') \mid x', f(\boldsymbol{x}^O), f(\boldsymbol{x}^C) < \mathbf{0}, & f_a(x') > 0 \\ \int_{-\infty}^{0} p(f(x') \mid f(\boldsymbol{x}^O), f(\boldsymbol{x}^C) < \mathbf{0})\, \mathrm{d}f(\boldsymbol{x}'), & f_a(x') = 0\,. \end{cases}$$

The resulting latent GASP is shown in the left panel of Figure 4, and the corresponding censored GASP is shown in the right panel of Figure 4. Here, latent GASP adequately assigns almost all of its probability mass to the negative range of values of a latent simulator output $f(x)$. The probability of occurrence of a zero at the censored input $x = -0.2$ is 87%. The censored GASP captures entirely function f_a in the region, including region of positive output.

Three methods (a), (b) and (c) posess their corresponding emulators. Performance of these emulators is assessed numerically. $d = 201$ equidistant testing points $x_i \in [-1, 1]$ for all $i = 1, \ldots, d$ are taken. Numerically, performance of three emulators of the simulator $f_a(x)$ is summarized in Tables 1, 2 and 3 using predictive checks, namely root mean square predictive error (RMSPE), empirical frequency coverage (EFC) and length of credible intervals ($\mathrm{L_{CI}}$). Let $\mathbf{x} = \{x_1, \ldots, x_d\}$ be a set of d inputs with corresponding output of a simulator $f_a(x_1), \ldots, f_a(x_d)$, and let the predictive distribution at each testing point be p_i. This distribution is assumed to have mean μ_i and 2.5% and 97.5% quantiles $q_i^{0.025}$ and $q_i^{0.975}$, then

$$\mathrm{RMSPE} = \left(\sum_{i=1}^{d} (f_a(x_i) - \mu_i)^2 / d \right)^{1/2},$$

$$\mathrm{EFC} = \left(\sum_{i=1}^{d} \mathbb{I}_{f_a(x_i) \in \mathrm{CI}_i} \right) / d\,,$$

$$\mathrm{L_{CI}} = \sum_{i=1}^{d} (q_i^{0.975} - q_i^{0.025}) / d\,,$$

Table 1: Comparison of three emulators: projected traditional GASP without zeros, projected traditional GASP with zeros and censored GASP. Comparison is made on all testing points, including both, zero-output points and positive-output points.

Method	GASP	RMSPE	EFC	L_{CI}
(c)	Censored	0.032	1.000	0.067
(b)	Projected traditional w/ zeros	0.197	0.413	0.104
(a)	Projected traditional w/o zeros	0.204	1.000	0.729

Table 2: Comparison of three emulators: projected traditional GASP without zeros, projected traditional GASP with zeros and censored GASP. Comparison is made on zero-output testing points.

Method	GASP	RMSPE	EFC	L_{CI}
(c)	Censored	0.000	1.000	0.000
(b)	Projected traditional w/ zeros	0.012	0.963	0.023
(a)	Projected traditional w/o zeros	0.235	1.000	0.412

where $CI_i = (q_i^{0.025}, q_i^{0.975})$ is the 95% credible interval.

Table 1 provides comparison of emulators in the area of all output testing points, including zero-output testing points of a simulator and including positive-output testing points of a simulator. The second table 2 provides comparison of emulators in the zero-output region only, and the third one (Table 3) provides comparison in the positive region output only.

Table 1 reveals that projected emulator (a) (which does not have zeros in its training points) performs worse than the other two emulators (b) or (c). E.g., emulator (a) and censored emulator (c) both have empirical frequency coverage equal to 100%, however, RMSPE is at least 4 times larger (Table 3) for projected emulator (a) than for censored emulator (c), and length of credible intervals is at least twice as large for projected emulator (a) (Table 3). This is anticipated, however, since last emulator does not make use of zero training points.

We turn then to comparison of a *censored* emulator (c) with *projected* emulator (b) (which has zero-training points included in its construction). Visually, the censored emulator (c) performs similar to the projected emulator (b). Emulator (b) in the area of positive output has the mean of the emulator closely following the true function. In the area of positive output, the predicted mean also closely follows the true zero-valued output. However, numerical comparison reveals clear differences in the behaviour of two emulators. Table 1 shows that RMSPE of censored emulator (c) is 6 times smaller than that of the projected emulator (b), length of credible intervals is 1.5 times smaller for censored emulator (c), while empirical coverage is only 41% for projected emulator (b) and 100% for censored emulator (c).

Numerical summaries for two emulators in zero-output testing region (Table 2) show that censored GASP perfectly predicts zero-output in this region, producing RMSPE equal to zero and also essentially zero-length credible intervals, while empirical coverage is 100%. Projected emulator (b) is a bit worse, with non-zero, but still small RMSPE and L_{CI}.

Considering the probability of a censored zero-valued output to occur in this region, the projected emulator

Table 3: Comparison of three emulators: projected traditional GASP without zeros, projected traditional GASP with zeros and censored GASP. Comparison is made on positive output testing points.

Method	GASP	RMSPE	EFC	L_{CI}
(c)	Censored	0.042	1.000	0.112
(b)	Projected traditional w/ zeros	0.254	0.042	0.157
(a)	Projected traditional w/o zeros	0.184	1.000	0.283

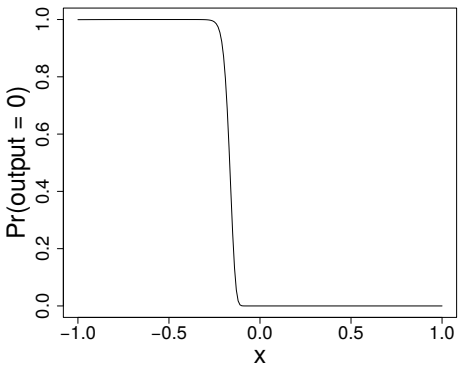

Figure 5: Probabilities of a zero at testing points assigned by the censored GASP.

(b) assigns the probability to be about 1/2 on average, while the censored emulator (c) assigns probability that a zero-value occurs there to almost 1 which is demonstrated in Figure 5. The function luckily appears to be smooth in its estimation of the hazard.

Comparison of emulators in positive output testing region (Table 3) shows different qualitative performance of the two emulators, (b) and (c). RMSPE for censored emulator (c) is 6 times smaller than that for projected emulator (b), length of credible intervals of two emulators is similar, but censored emulator (c) has length of credible intervals L_{CI} on average 1.4 times smaller. At the same time, empirical frequency coverage is only 4% for projected emulator (b), while censored emulator (c) has 100% empirical coverage.

Along with the inferior performance of projected emulator (b) to censored emulator (c), the former does not allow for the true distinction between small-height values and zero-height of the flow, while censored emulator (c) captures this difference.

3 Case study

In this section the results of three methods of construction of stochastic approximations to the TITAN2D computer model data are described. Implications on probabilistic assessment of a hazard from a volcano pyroclastic flow are demonstrated.

15,122 locations were considered at which the GASPs were trained and tested. All these locations have both, exact zero-height output points and positive height output points. Original number of 16,085 locations with non-zero output had to be reduced because of computational reasons: there need to be a minimal number of training points for constructing a traditional GASP. Only locations in which emulators have at least three non-zero output points in a training set are considered. Analogously to the illustrative example the same training points for construction of projected traditional GASP with zero-height output points included (method (b)) and for censored GASP (method (c)). Only positive height training points are used for construction of projected traditional GASP with zero-height output points excluded from its training set (method (a)). All the same testing points are used for comparison of the performance of three emulators.

The output of a computer model has not been transformed for this case study, so the output is the maximum height of a pyroclastic flow. This is due to the harmful mathematical effects anticipated from the results of such a transformation as have been demonstrated (Kyzyurova 2017, 2019*c*).[1]

Table 4 summarizes the performance of three developed emulators according to methods (a), (b) and (c).

[1]Some commonly used variants of log-transformations are

- $\log(f(\cdot))$, where $f(\cdot) \in (0, \infty)$,
- $\log \frac{f(\cdot)}{1-f(\cdot)}$ where $f(\cdot) \in (0, 1)$, and
- $\log(f(\cdot) + 1)$, where $f(\cdot) \geq 0$ and this function may possibly take large values.

Table 4: Comparison of three emulators: projected traditional GASP w/o zeros, projected traditional GASP w/zeros and censored GASP. Comparison is made on all testing points, including both, zero-height output points and positive-height output points.

Method	GASP	RMSPE	EFC	L_{CI}
(c)	Censored	0.649	0.941	1.091
(b)	Projected traditional w/ zeros	0.618	0.914	1.371
(a)	Projected traditional w/o zeros	5.356	0.945	7.893

Table 5: Average probability of a zero-height of three emulators: projected traditional GASP without zeros, projected traditional GASP with zeros and censored GASP at zero-testing points.

Method	GASP	$P(h=0)$
(c)	Censored	0.968
(b)	Projected traditional w/ zeros	0.504
(a)	Projected traditional w/o zeros	0.712

Censored emulator (c) provides the best approximation while maintaining high empirical frequency coverage. Censored GASP provides RMSPE comparable to that one of a projected traditional GASP with zeros included in the training data, however the average length of credible intervals is smaller for censored GASP, than for a projected traditional one.

Interestingly, projected traditional GASP without zeros in the data performs much worse than the other two emulators: RMSPE and L_{CI} are very large. This is because the amount of positive data points is only about one third of the amount of all the data points (including zero-height output points) available. Also zero-height testing points, similar to the simulation example, occur in "clusters". This is intentionally so, because the application with computer model TITAN2D has such a pattern of occurrence of zeros. Thus the latent emulator from method (a) without zeros has to do extrapolation at such points. This becomes the major drawback of method (a). As shown below, this is not so for methods (b) and (c).

Table 5 shows the average probability of a zero-height flow over all zero-output testing points. Censored GASP results in such a flow with probability close to 1, which is what we would like to see from the performance of an emulator. Two other traditional emulators result in average probabilities of zero being too far away from the expected probability of 1. These are the artefacts of these GASPs, which make these GASPs inappropriate for direct emulation of a zero-inflated output of a computer model.

Table 6 summarizes the performance of the censored GASP and projected traditional GASP with zeros in the training data on zero-height testing points only. The RMSPE for censored GASP is about two thirds of that of a projected traditional GASP and L_{CI} is more than 6 times smaller for censored GASP. Empirical frequency coverage of both emulators is more than 99%.

Table 7 summarizes the performance of the censored GASP and projected traditional GASP with zeros in the training data on positive height testing points only. The RMSPE for censored GASP and traditional GASP with zeros are close. Average length of predictive intervals L_{CI} is about 25% greater for censored GASP than for the traditional GASP, though the empirical frequency coverage (EFC) of projected traditional GASP is only 80%, while the censored GASP has greater EFC of 86%. Projected traditional with zeros GASP

Table 6: Comparison of two emulators: traditional with zeros and censored GASP on zero-testing points.

Method	GASP	RMSPE	EFC	L_{CI}
(c)	Censored	0.111	0.999	0.175
(b)	Projected traditional w/ zeros	0.160	0.998	1.165

Table 7: Comparison of two emulators: traditional with zeros and censored GASP on non-zero testing points.

Method	GASP	RMSPE	EFC	L_{CI}
(c)	Censored	0.833	0.864	1.930
(b)	Projected traditional w/ zeros	0.853	0.806	1.543

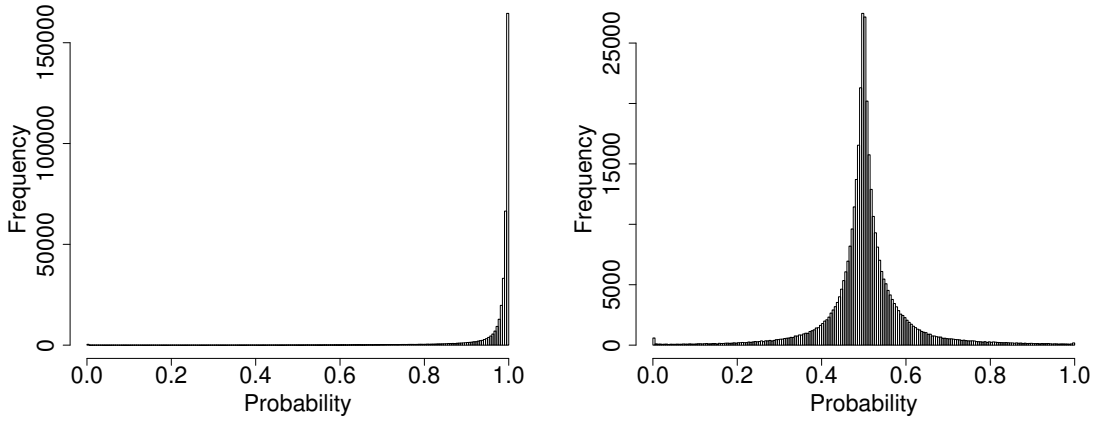

Figure 6: Left: histogram of probability of zero by the censored GASP. These are probabilities at zero-height testing points. Right: histogram of probability of a zero by the projected traditional GASP w/ zeros if posterior is projected to be censored at zero.

underestimates uncertainty in a region of positive height of the flow, because it has two thirds of its training data being zero-output data points included in construction of the GASP.

The distributions of resulting probabilities over all zero-testing points are presented in Figure 6. The left panel of the figure provides a histogram of censored GASP probabilities of a zero to occur. This distribution is highly concentrated near probability value being 1. The right panel of the figure demonstrates that projected traditional GASP with zeros included as data points provides probability of a zero at zero-testing points concentrated around 50%, indicating bad performance of an emulator.

These distributions induce the following distributions of a zero-height to occur among all testing points: zero-height testing points and positive height testing points, shown in Figure 7. The left histogram of values, which correspond to the ones given by censored GASP emulator shows that probability of a zero-height flow concentrates near zero-probability or near 100% probability for a zero-height to occur. The histogram on the right indicates that traditional GASP with zeros provides unattractive strong peak at probability of a zero-height being around 50% in addition to extremes of near 0% and 100% probabilities.

In this work stochastic approximation using individual emulators at each spatial location is performed. Alternative solution would be to utilize the parallel partial emulator (Gu & Berger 2016). However, the latter and Kyzyurova (2017, 2019*b*) show that spatial structure on outputs is not relevant for the purpose of emulation. To add, proper emulation of the zero output, employing censoring methodology, in this case becomes seemingly a prohibitive task due to large dimensionality of the output over the product of the number of spatial locations and the number of input parameters. Many independent emulators obtained at each spatial location is a reasonable and computationally cheap alternative to parallel partial emulator. If the aim is to post-process the resulting emulators, for instance, to provide probabilistic maps of volcano pyroclastic flow hazards, in theory we might need to smooth out the results from such a post-processing, but in practice we receive rather smooth performance across all emulators over the spatial locations.

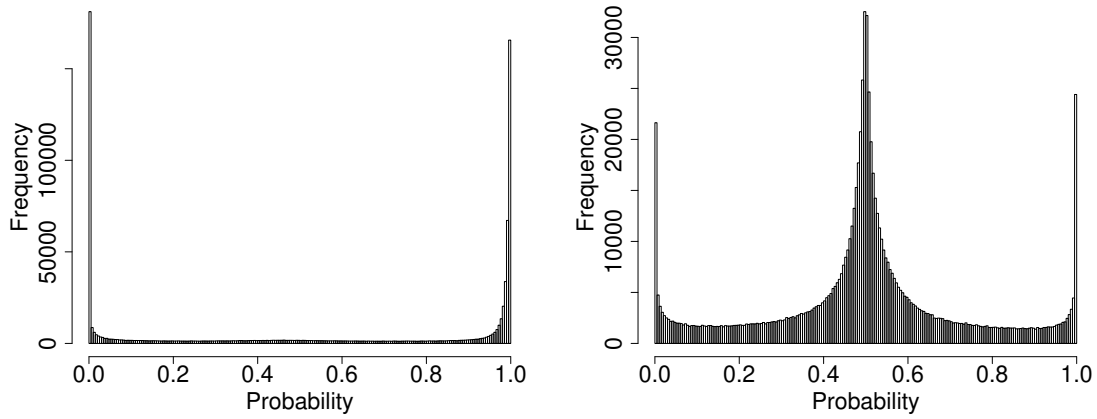

Figure 7: Left: histogram of probability of zero by the censored GASP. These are probabilities for all testing points. Right: histogram of probability of zero by the traditional GASP w/ zeros if posterior is projected to be censored at zero.

4 Probability of a hazard

Once a valid emulator of a pyroclastic flow height is constructed, (assuming this is done so employing the methodology of the linked emulator (Kyzyurova et al. 2018) for coupled models which allows to submerge the basal friction parameter of the computer model TITAN2D), and the distributions $p(V, \theta)$ of the initial volume V and initiation angle θ are available, one may generate values of V and θ according to the distributions. Conditional probability of a hazard (the maximum height of the flow exceeding some critical value $h_{\text{crit.}}$) given the occurrence of the pyroclastic flow event (PF) may then be estimated as a Monte-Carlo integration

$$P(h \ > \ h_{\text{crit.}} \ | \ \text{PF}) \ = \ \int\int_{[0,2\pi)\times[\epsilon,\infty)} P(h \ > \ h_{crit.} \ | \ V, \theta) p(V, \theta) dV d\theta \ \approx \ \frac{1}{N} \sum_{i=1}^{N} P(h \ > \ h_{\text{crit.}} \ | \ V_i, \theta_i) \, ,$$

where ϵ is some small fixed value for a volume of the flow V_i, $V_i \sim$ Pareto(scale $= 5 \cdot 10^4$, shape $= .64$) and $\theta_i \sim$ Unif$(0, 2\pi)$ — based on the exploratory analysis provided in (Bayarri et al. 2009).

The intuition says that results obtained in this work show that, if calculated, such an estimate of a catastrophe would be very small, since for the most locations the height of the flow is zero. This means that, according to the computer model, the catastrophe is indeed a rare event. Yet, locations most prone to the catastrophe may be of interest to identify to enable reasonable policy making, perhaps, in order to avoid the choice of such locations for subsequent use, such as, e.g., for a living, logistics or transportation use.

5 Appendix

This section builds on the idea to employ truncated versions of the GASP emulator motivated by an example when such a mathematical construction is useful. Elements of construction of linked emulators by linking traditional GASP with either censored or truncated GASP conclude this manuscript.

Sometimes another closely related to censoring but different problem — truncation of the output — arises. One of the inputs to computer model TITAN2D is basal friction of a pyroclastic flow, friction of the flow with the surface. The relationship between basal friction of a pyroclastic flow and volume of a flow has been modelled (Ogburn et al. 2016), and stochastic representation has been obtained

$$y \mid x \sim \mathcal{N}(-0.5336 + -0.1971x, 0.00037205 - 0.00054567x + 0.00044155x^2) \, , \tag{4}$$

where y is a basal friction, and $x = V - 5.5$ with V being volume of a flow on a $\log_{10}$ scale.

For large volumes of a pyroclastic flow there are only small number of historical data points, and an inferential relationship (4) is less reliable. Geological reasoning suggests that basal friction does not decrease that much for large volumes, as the inferential relationship states, and this relationship should adopt some threshold c below which basal friction does not decrease. Thus, one appealing solution is to model the relationship between basal friction and volume of a distribution, not directly using relationship (4), but having this relationship truncated at c, i.e.

$$y \mid x \sim \mathcal{TN}_{(c,\infty)}(-0.5336 + -0.1971x, 0.00037205 - 0.00054567x + 0.00044155x^2)\,.$$

5.1 Truncated projected GASP

5.1.1 Methodology

General setting for which a truncated version of the GASP may be applicable is introduced. Suppose that a smooth function $g(\cdot)$ represents a simulator. The simulator data $g(\mathbf{z})$ are evaluations of the simulator at several input points $\mathbf{z} = (z_1, \ldots, z_\ell)$. Suppose that additionally it is known that a simulator takes values in the range (a, b). That is, $a < g(z_i) < b$ for each $i = 1, \ldots, \ell$. The difference in this framework from censoring framework (discussed in the main paper) constitutes in that there is no positive point mass on any value of the output variable.

Thus an emulator which takes values in a specific region and has representation via an absolutely continuous random variable is desirable to be constructed. Unfortunately, the following stochastic process

$$g^M_{tr}(\cdot) \sim \mathcal{TN}_{(a,b)}(\mu(\cdot), \sigma^2 c(\cdot,\cdot))\,,$$

such that all finite-dimensional multivariate distributions are joint truncated multivariate normal, does not exist, which may be demonstrated using results of (Horrace 2005).

An *ad hoc* method, (for instance, Lin & Dunson (2014), — for constraining GASPs of monotonic functions), is to construct a traditional GASP, and then truncate posterior predictive distribution to a region (a, b).

Let $g^M(\mathbf{z}') \mid g^M(\mathbf{z})$ be a posterior predictive distribution at a new input point $\mathbf{z}'$, approximation to simulator output $g(\mathbf{z}')$, then the proposed approximation is

$$g^M_{tr}(\mathbf{z}') = g^M(\mathbf{z}') \mid g^M(\mathbf{z}), a < g^M(\mathbf{z}') < b\,.$$

Joint distribution of emulator output at points $\mathbf{z}$ and $\mathbf{z}'$ follows multivariate normal distribution

$$\begin{pmatrix} g^M(\mathbf{z}) \\ g^M(\mathbf{z}') \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu(\mathbf{z}) \\ \mu(\mathbf{z}') \end{pmatrix}, \sigma^2 \begin{pmatrix} \mathbf{C}_{\mathbf{z}} & c(\mathbf{z},\mathbf{z}') \\ c(\mathbf{z}',\mathbf{z}) & \mathbf{C}_{\mathbf{z}'} \end{pmatrix} \right)\,,$$

where $\mathbf{C}_{\mathbf{z}}$ and $\mathbf{C}_{\mathbf{z}'}$ are correlation matrices whose (k,l)th elements are given by a correlation function $c(\cdot,\cdot)$.

Conditional on the observed computer model evaluations $g^M(\mathbf{z})$ the truncated posterior predictive at any input $\mathbf{z}'$ follows a truncated to (a, b) normal distribution with mean $\mu^*(\mathbf{z}')$ and variance $\sigma^{*2}(\mathbf{z}'))$, i.e.

$$g^M_{tr}(\mathbf{z}') \mid g^M(\mathbf{z}) \sim \mathcal{TN}_{(a,b)}(\mu^*(\mathbf{z}'), \sigma^{*2}(\mathbf{z}'))\,,$$

where

$$\begin{aligned} \mu^*(\mathbf{z}') &= \mu(\mathbf{z}') + c(\mathbf{z}',\mathbf{z})\mathbf{C}_{\mathbf{z}}{}^{-1}(g^M(\mathbf{z}) - \mu(\mathbf{z}))\,, \\ \sigma^{*2}(\mathbf{z}') &= \sigma^2(\mathbf{C}_{\mathbf{z}'} - c(\mathbf{z}',\mathbf{z})\mathbf{C}_{\mathbf{z}}{}^{-1}c(\mathbf{z},\mathbf{z}'))\,. \end{aligned}$$

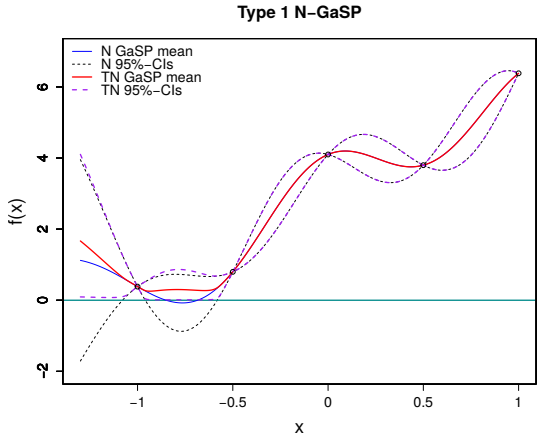

Figure 8: GASP and truncated GASP of the test function $f(x)$. Circled points correspond to training inputs which were used to fit the GASP emulator.

Mean, any quantile of the distribution $g_{tr}^M(\mathbf{z}') \mid g^M(\mathbf{z})$ and variance of this distribution have the following closed-form expressions

$$\mathrm{E}\, g_{tr}^M(\mathbf{z}') \mid g^M(\mathbf{z}) = \mu^*(\mathbf{z}') - \frac{\phi\left(\frac{b-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right) - \phi\left(\frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right)}{\Phi\left(\frac{b-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right) - \Phi\left(\frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right)}\sigma^*(\mathbf{z}')\,,$$

$$\mathrm{Q}_{g_{tr}^M(\mathbf{z}')|g^M(\mathbf{z})}(q) = \Phi^{-1}\left(\frac{\Phi\left(\frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right) + \Phi\left(\frac{b-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right) - \Phi\left(\frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right)q - \mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right),$$

$$\mathrm{V}\, g_{tr}^M(\mathbf{z}') \mid g^M(\mathbf{z}) = \sigma^{*2}(\mathbf{z}')\left(1 - \frac{\frac{b-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\phi\left(\frac{b-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right) - \frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\phi\left(\frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right)}{\Phi\left(\frac{b-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right) - \Phi\left(\frac{a-\mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right)} - \left(\frac{\mathrm{E}\, g^M(\mathbf{z}') \mid g^M(\mathbf{z}) - \mu^*(\mathbf{z}')}{\sigma^*(\mathbf{z}')}\right)^2\right),$$

where $\mathrm{Q}_{g_{tr}^M(\mathbf{z}')|g^M(\mathbf{z})}(q)$ is a qth quantile of distribution $g_{tr}^M(\mathbf{z}') \mid g^M(\mathbf{z})$; ϕ and Φ are, respectively, probability density function and cumulative distribution function of a standard normal distribution.

5.1.2 Illustrative example

The example demonstrates the methodology. Test function $f(x) = 3x + \cos(5x) + 3.1$ in the range $x \in [-1.3, 1]$ is considered. This function $f(x) > 0$, so there is the threshold $c = 0$ below which simulator output should not exist. Truncated normal approximation accounts for truncation region with no assignment of positive probability to regions of actual zero probability, opposed to the behaviour of a usual GASP, with substantial probability for negative values.

5.2 Linking a GASP emulator with a truncated GASP

Following Kyzyurova et al. (2018), where the methodology for linked emulators have been introduced, formulae for linked traditional GASP and truncated GASP are provided below. This section is followed by the one which provides formulae for linking traditional and censored GASP emulators. Parameters of the GaSPs are assumed to be either known or having corresponding point estimates plugged in (for example, using maximum likelihood or maximum *a posteriori* estimation procedures). These frameworks are generalizable to linking GASPs with their Bayesian implementation, but the corresponding theorems are not provided.

Let denote truncated GASP emulators f_j^M, whose distributions at a new point is in the form

$$f_j^M(\cdot) \sim \mathcal{TN}_{c_j}(s_j, \sigma_j^2)\,. \tag{5}$$

Theorem. Let g^M with given parameters $\theta_{\mathbf{g}} = (\beta, \sigma^2, \{\delta_j\}_{j=1,\ldots,d}, \eta)$ be a GASP emulator of a simulator g exercized at training input points $\mathbf{z}$. Suppose the mean is linear in the qth cooordinate of an input $\mathbf{z}'$, so that the mean is $h(\mathbf{z}')\boldsymbol{\beta} = \beta_0 + \beta_1 z_q')$. Let the $g^M(\cdot)$ GASP product power correlation function smoothness parameters α_j of coordinates $j \in b, \ldots, d$ be equal to 2. For each $j \in b, \ldots, q-1$ let f_j^M be an independent GASP emulator of a simulator f_j, corresponding to the coordinate j of the input to the simulator g, i.e. $f_j^M(\cdot)$. For each $j \in q, \ldots, d$ let f_j^M be an independent truncated normal GASP emulator of a simulator f_j, corresponding to the coordinate j of the input to the simulator g, i.e. $f_j^M(\cdot) \sim \mathcal{TN}_{c_j}(s_j, \sigma_j^2)$ (5). Then the closed-form mean $\mathrm{E}\xi$ and variance $\mathrm{V}\xi$ of the linked emulator ξ of the coupled simulator $(g \circ (f_b, \ldots, f_d))(\mathbf{u})$ are

$$\begin{aligned}
\mathrm{E}\xi &= \beta_0 + \beta_1 \mathrm{E} f_q^M + \sum_{i=1}^{m} a_i \prod_{j=1}^{b-1} e^{-\left(\frac{|u_j - z_{ij}|}{\delta_j}\right)^{\alpha_j}} \prod_{j=b}^{q-1} I_j^i \prod_{j=q}^{d} {I^0}_j^i\,,\\
\mathrm{V}\xi &= \sigma^2 \left(1 + \eta - \sum_{k,l=1}^{m} \{{C_z}^{-1}\}_{k,l} \prod_{j=1}^{b-1} e^{-\left(\left(\frac{|u_j - z_{kj}|}{\delta_j}\right)^{\alpha_j} + \left(\frac{|u_j - z_{lj}|}{\delta_j}\right)^{\alpha_j}\right)} \prod_{j=b}^{q-1} {I^{1}}_j^{k,l} \prod_{j=q}^{d} {I^{2}}_j^{k,l}\right) +\\
&\quad \beta_0^2 + 2\beta_0\beta_1 \mathrm{E} f_q^M + \beta_1^2 (\mathrm{V} f_q^M + (\mathrm{E} f_q^M)^2) +\\
&\quad 2\sum_{i=1}^{m} a_i \prod_{j=1}^{b-1} \exp\left(-\left(\frac{|u_j - z_{ij}|}{\delta_j}\right)^{\alpha_j}\right) (\beta_0 I_q^i + \beta_1 {I^{+}}_q^i) \prod_{j=b}^{q-1} I_j^i \prod_{j=q+1}^{d} {I^0}_j^i +\\
&\quad \sum_{k,l=1}^{m} a_l a_k \prod_{j=1}^{b-1} e^{-\left(\left(\frac{|u_j - z_{kj}|}{\delta_j}\right)^{\alpha_j} + \left(\frac{|u_j - z_{lj}|}{\delta_j}\right)^{\alpha_j}\right)} \prod_{j=b}^{q-1} {I^{1}}_j^{k,l} \prod_{j=q}^{d} {I^{2}}_j^{k,l} - (\mathrm{E}\xi)^2,
\end{aligned}$$

where $a = (a_1, \ldots, a_m)^T = {\mathbf{C}_{\mathbf{z}}}^{-1}(g^M(\mathbf{z}) - h(\mathbf{z})\boldsymbol{\beta})$, $t_q = \frac{c_q - s_q}{\sigma^*_{f_q}}$ and

$$\begin{aligned}
\mathrm{E} f_q^M &= \left(s_q + \sigma^*_{f_q} \phi(t_q)/\Phi(-t_q)\right)\,,\\
\mathrm{V} f_q^M &= \sigma^*_{f_q} \left(1 - \frac{\phi(t_q)}{\Phi(-t_q)} \left(\frac{\phi(t_q)}{\Phi(-t_q)} - t_q\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
I_j^i &= \frac{1}{\sqrt{1 + 2\frac{\sigma^{*2}_{f_j}(\mathbf{u^j})}{\delta_j^2}}} \exp\left(-\frac{(z_{ij} - \mu^*_{f_j}(\mathbf{u^j}))^2}{\delta_j^2 + 2\sigma^{*2}_{f_j}(\mathbf{u^j})}\right),\\
{I^0}_j^i &= \frac{1}{\Phi\left(-\frac{c_j - s_j}{\sigma^*_{f_j}}\right)} \frac{1}{\sqrt{1 + 2\frac{\sigma^{*2}_{f_j}}{\delta_j^2}}} \exp\left(-\frac{(z_{ij} - s_j)^2}{\delta_j^2 + 2\sigma^{*2}_{f_j}}\right)\\
&\quad \Phi\left(\frac{2(z_{ij} - c_j)\sigma^{*2}_{f_j} + (s_j - c_j)\delta_j^2}{\sqrt{\delta_j^2 + 2\sigma^{*2}_{f_j}}}\right),\\
{I^1}_j^{k,l} &= \frac{1}{\sqrt{1 + 4\frac{\sigma^{*2}_{f_j}(\mathbf{u^j})}{\delta_j^2}}} e^{-\frac{\left(\frac{z_{kj} + z_{lj}}{2} - \mu^*_{f_j}(\mathbf{u^j})\right)^2}{\frac{\delta_j^2}{2} + 2\sigma^{*2}_{f_j}(\mathbf{u^j})}} e^{-\frac{(z_{kj} - z_{lj})^2}{2\delta_j^2}},
\end{aligned}$$

$$
\begin{aligned}
I^{2}{}_{j}^{k,l} &= \frac{1}{\Phi\left(-\frac{c_j - s_j}{\sigma^*_{f_j}}\right)} \frac{1}{\sqrt{1 + 4\frac{\sigma^{*2}_{f_j}}{\delta_j^2}}} \exp - \frac{(z_{kj} - z_{lj})^2}{2\delta_j^2} \\
&\quad \exp\left(-\frac{(\frac{z_{kj}+z_{lj}}{2} - s_j)^2}{\delta_j^2/2 + 2\sigma^{*2}_{f_j}}\right) \Phi\left(\frac{2(z_{kj} + z_{lj} - 2c_j)/\delta_j^2 + (s_j - c_j)/\sigma^{*2}_{f_j}}{\sqrt{4/\delta_j^2 + 1/\sigma^{*2}_{f_j}}}\right), \\
I^{+i}{}_{q} &= \frac{1}{\Phi\left(-\frac{c_j - s_j}{\sigma^*_{f_j}}\right)} \frac{1}{\sqrt{(1 + 2\sigma^{*2}_{f_q}/\delta_q^2)}} e^{-\frac{(z_{iq} - s_q)^2}{\delta_q^2 + 2\sigma^{*2}_{f_q}}} \\
&\quad \left(\frac{2\sigma^{*2}_{f_q} z_{iq} + s_q \delta_q^2}{2\sigma^{*2}_{f_q} + \delta_q^2} \Phi\left(\frac{2\sigma^{*2}_{f_q}(z_{iq} - c_q) + \delta_q^2(s_q - c_q)}{\delta_q \sigma^*_{f_q} \sqrt{2\sigma^{*2}_{f_q} + \delta_q^2}}\right) + \right. \\
&\quad \left. \frac{\sigma^*_{f_q} \delta_q}{\sqrt{\delta_q^2 + 2\sigma^{*2}_{f_q}}} \phi\left(\frac{2\sigma^{*2}_{f_q}(z_{iq} - c_q) + \delta_q^2(s_q - c_q)}{\delta_q \sigma^*_{f_q} \sqrt{2\sigma^{*2}_{f_q} + \delta_q^2}}\right)\right).
\end{aligned}
$$

5.3 Linking a GASP emulator with a censored GASP

Let the predictive distribution of a censored GASP $f_j^M(\cdot)$ be

$$
f_j^M(x \mid s_j, \sigma_j^2) = \begin{cases} \Phi\left(\frac{c_j - s_j}{\sigma_j}\right), & x = c_j \\ \frac{1}{\sigma_j}\phi\left(\frac{x - s_j}{\sigma_j}\right), & x > c_j \end{cases}, \tag{6}
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are density and cumulative distribution functions of a normal distribution.

Theorem. Let g^M with given parameters $\theta_{\mathbf{g}} = (\beta, \sigma^2, \{\delta_j\}_{j=1,\ldots,d}, \eta)$ be an emulator of a simulator g exercized at training input points $\mathbf{z}$. Suppose the mean is linear in the qth coordinate of an input $\mathbf{z}'$, so that the mean is $h(\mathbf{z}')\boldsymbol{\beta} = \beta_0 + \beta_1 z'_q$). Let the $g^M(\cdot)$ GASP product power correlation function smoothness parameters α_j of coordinates $j \in b, \ldots, d$ be equal to 2. For each $j \in b, \ldots, q-1$ let f_j^M be an independent GASP emulator of a simulator f_j, corresponding to the coordinate j of the input to the simulator g, i.e. $f_j^M(\cdot)$. For each $j \in q, \ldots, d$ let f_j^M be an independent censored GASP emulator of a simulator f_j, corresponding to the coordinate j of the input to the simulator g, i.e. $f_j^M(\cdot)$ (6). Then the mean $\mathrm{E}\xi$ and variance $\mathrm{V}\xi$ of the linked emulator ξ of the coupled simulator $(g \circ (f_b, \ldots, f_d))(\mathbf{u})$ are

$$
\begin{aligned}
\mathrm{E}\xi &= \beta_0 + \beta_1 \mathrm{E} f_q^M + \sum_{i=1}^{m} a_i \prod_{j=1}^{b-1} e^{-\left(\frac{|u_j - z_{ij}|}{\delta_j}\right)^{\alpha_j}} \prod_{j=b}^{q-1} I_j^i \prod_{j=q}^{d} I^{0i}{}_j, \\
\mathrm{V}\xi &= \sigma^2 \left(1 + \eta - \sum_{k,l=1}^{m} \{C_z{}^{-1}\}_{k,l} \prod_{j=1}^{b-1} e^{-\left(\left(\frac{|u_j - z_{kj}|}{\delta_j}\right)^{\alpha_j} + \left(\frac{|u_j - z_{lj}|}{\delta_j}\right)^{\alpha_j}\right)} \prod_{j=b}^{q-1} I^{1}{}_j^{k,l} \prod_{j=q}^{d} I^{2}{}_j^{k,l}\right) + \\
&\quad \beta_0^2 + 2\beta_0\beta_1 \mathrm{E} f_q^M + \beta_1^2 (\mathrm{V} f_q^M + (\mathrm{E} f_q^M)^2) + \\
&\quad 2\sum_{i=1}^{m} a_i \prod_{j=1}^{b-1} \exp\left(-\left(\frac{|u_j - z_{ij}|}{\delta_j}\right)^{\alpha_j}\right) (\beta_0 I_q^i + \beta_1 I^{+i}{}_q) \prod_{j=b}^{q-1} I_j^i \prod_{j=q+1}^{d} I^{0i}{}_j + \\
&\quad \sum_{k,l=1}^{m} a_l a_k \prod_{j=1}^{b-1} e^{-\left(\left(\frac{|u_j - z_{kj}|}{\delta_j}\right)^{\alpha_j} + \left(\frac{|u_j - z_{lj}|}{\delta_j}\right)^{\alpha_j}\right)} \prod_{j=b}^{q-1} I^{1}{}_j^{k,l} \prod_{j=q}^{d} I^{2}{}_j^{k,l} - (\mathrm{E}\xi)^2,
\end{aligned}
$$

where $a = (a_1, \ldots, a_m)^T = \mathbf{C_z}^{-1}(g^M(\mathbf{z}) - h(\mathbf{z})\boldsymbol{\beta})$, $t_q = \frac{c_q - s_q}{\sigma^*_{f_q}}$ and

$$\begin{aligned}
\mathrm{E}f_q^M &= \left(s_q\Phi\left(-t_q\right) + \sigma^*_{f_q}\phi(t_q) + c_q\Phi(t_q)\right), \\
\mathrm{V}f_q^M &= c_q^2\Phi(t_q) + \sigma^{*2}_{f_q}\left(\Phi(-t_q) - \phi(t_q)\left(\frac{\phi(t_q)}{\Phi(-t_q)} - t_q\right)\right) + \\
&\qquad \frac{\left(s_q\Phi(-t_q) + \sigma^*_{f_q}\phi(t_q)\right)^2}{\Phi(-t_q)} - (\mathrm{E}f_q^M)^2,
\end{aligned}$$

and

$$\begin{aligned}
I_j^i &= \frac{1}{\sqrt{1 + 2\frac{\sigma^{*2}_{f_j}(\mathbf{u^j})}{\delta_j^2}}}\exp\left(-\frac{(z_{ij} - \mu^*_{f_j}(\mathbf{u^j}))^2}{\delta_j^2 + 2\sigma^{*2}_{f_j}(\mathbf{u^j})}\right), \\
I^{0}{}_j^i &= \exp\left(-\frac{(c_j - z_{ij})^2}{\delta_j^2}\right)\Phi\left(\frac{c_j - s_j}{\sigma^*_{f_j}}\right) + \\
&\qquad \frac{1}{\sqrt{1 + 2\frac{\sigma^{*2}_{f_j}}{\delta_j^2}}}\exp\left(-\frac{(z_{ij} - s_j)^2}{\delta_j^2 + 2\sigma^{*2}_{f_j}}\right)\Phi\left(\frac{2(z_{ij} - c_j)\sigma^{*2}_{f_j} + (s_j - c_j)\delta_j^2}{\sqrt{\delta_j^2 + 2\sigma^{*2}_{f_j}}}\right), \\
I^{1}{}_j^{k,l} &= \frac{1}{\sqrt{1 + 4\frac{\sigma^{*2}_{f_j}(\mathbf{u^j})}{\delta_j^2}}}e^{-\frac{\left(\frac{z_{kj} + z_{lj}}{2} - \mu^*_{f_j}(\mathbf{u^j})\right)^2}{\frac{\delta_j^2}{2} + 2\sigma^{*2}_{f_j}(\mathbf{u^j})}}e^{-\frac{(z_{kj} - z_{lj})^2}{2\delta_j^2}},
\end{aligned}$$

$$\begin{aligned}
I^{2}{}_j^{k,l} &= \exp\left(-\frac{(c_j - z_{kj})^2}{\delta_j^2} - \frac{(c_j - z_{lj})^2}{\delta_j^2}\right)\Phi\left(\frac{c_j - s_j}{\sigma^*_{f_j}}\right) + \frac{1}{\sqrt{1 + 4\frac{\sigma^{*2}_{f_j}}{\delta_j^2}}} \\
&\qquad \exp\left(-\frac{(z_{kj} - z_{lj})^2}{2\delta_j^2}\right)\exp\left(-\frac{(\frac{z_{kj} + z_{lj}}{2} - s_j)^2}{\delta_j^2/2 + 2\sigma^{*2}_{f_j}}\right) \\
&\qquad \Phi\left(\frac{2(z_{kj} + z_{lj} - 2c_j)/\delta_j^2 + (s_j - c_j)/\sigma^{*2}_{f_j}}{\sqrt{4/\delta_j^2 + 1/\sigma^{*2}_{f_j}}}\right), \\
I^{+}{}_q^i &= c_q e^{-\frac{(c_q - z_{iq})^2}{\delta_q^2}}\Phi\left(\frac{c_q - s_q}{\sigma^*_{f_q}}\right) + \frac{1}{\sqrt{(1 + 2\sigma^{*2}_{f_q}/\delta_q^2)}}e^{-\frac{(z_{iq} - s_q)^2}{\delta_q^2 + 2\sigma^{*2}_{f_q}}} \\
&\qquad \left(\frac{2\sigma^{*2}_{f_q}z_{iq} + s_q\delta_q^2}{2\sigma^{*2}_{f_q} + \delta_q^2}\Phi\left(\frac{2\sigma^{*2}_{f_q}(z_{iq} - c_q) + \delta_q^2(s_q - c_q)}{\delta_q\sigma^*_{f_q}\sqrt{2\sigma^{*2}_{f_q} + \delta_q^2}}\right) + \right. \\
&\qquad \left. \frac{\sigma^*_{f_q}\delta_q}{\sqrt{\delta_q^2 + 2\sigma^{*2}_{f_q}}}\phi\left(\frac{2\sigma^{*2}_{f_q}(z_{iq} - c_q) + \delta_q^2(s_q - c_q)}{\delta_q\sigma^*_{f_q}\sqrt{2\sigma^{*2}_{f_q} + \delta_q^2}}\right)\right).
\end{aligned}$$

References

Aghakhani, H., Dalbey, K., Salac, D. & Patra, A. K. (2016), 'Heuristic and Eulerian interface capturing approaches for shallow water type flow and application to granular flows', *Computer Methods in Applied Mechanics and Engineering* **304**, 243–264.

Asher, T. (2018), 'personal communication'.

Bayarri, M., Berger, J. O., Calder, E. S., Dalbey, K., Lunagomez, S., Patra, A. K., Pitman, E. B., Spiller, E. T. & Wolpert, R. L. (2009), 'Using statistical and computer models to quantify volcanic hazards', *Technometrics* **51**(4), 402–413.

Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C.-H. & Tu, J. (2007), 'A framework for validation of computer models', *Technometrics* **49**(2), 138–154.

Gopinathan, D. (2018), 'personal communication'.

Gopinathan, D., Venugopal, M., Roy, D., Rajendran, K., Guillas, S. & Dias, F. (2017), 'Uncertainties in the 2004 sumatra–andaman source through nonlinear stochastic inversion of tsunami waves', *Proc. R. Soc. A* **473**(2205), 20170353.

Gu, M. & Berger, J. O. (2016), 'Parallel partial gaussian process emulation for computer models with massive output', *The Annals of Applied Statistics* **10**(3), 1317–1347.

Gu, M., Wang, X., Berger, J. O. et al. (2018), 'Robust gaussian stochastic process emulation', *The Annals of Statistics* **46**(6A), 3038–3066.

Horrace, W. C. (2005), 'Some results on the multivariate truncated normal distribution', *Journal of Multivariate Analysis* **94**(1), 209–221.

Kyzyurova, K. N. (2017), On Uncertainty Quantification for Systems of Computer Models, PhD thesis, Duke University.

Kyzyurova, K. N. (2019*a*), *Analysis of scientific computer models. Methodology in numerical simulator data analysis*, ■.

Kyzyurova, K. N. (2019*b*), *On linear model of coregionalization*, ■.

Kyzyurova, K. N. (2019*c*), *On* log*-transformation of a computer model data. Emulation of a positive continuous simulator*, ■.

Kyzyurova, K. N., Berger, J. O. & Wolpert, R. L. (2018), 'Coupling computer models through linking their statistical emulators', *SIAM/ASA Journal on Uncertainty Quantification* **6**(3), 1151–1171.

Lin, L. & Dunson, D. B. (2014), 'Bayesian monotone regression using Gaussian process projection', *Biometrika* **101**(2).

Loeppky, J. L., Sacks, J. & Welch, W. J. (2009), 'Choosing the sample size of a computer experiment: A practical guide', *Technometrics* **51**(4), 366–376.

Maatouk, H. & Bay, X. (2017), 'Gaussian process emulators for computer experiments with inequality constraints', *Mathematical Geosciences* **49**(5), 557–582.

Ogburn, S. E., Berger, J., Calder, E. S., Lopes, D., Patra, A., Pitman, E. B., Rutarindwa, R., Spiller, E. & Wolpert, R. L. (2016), 'Pooling strength amongst limited datasets using hierarchical bayesian analysis, with application to pyroclastic density current mobility metrics', *Statistics in Volcanology* **2**(1), 1.

O'Hagan, A. (2006), 'Bayesian analysis of computer code outputs: A tutorial', *Reliability Engineering & System Safety* **91**(10-11), 1290–1300.

Patra, A. K., Bauer, A., Nichita, C., Pitman, E. B., Sheridan, M., Bursik, M., Rupp, B., Webber, A., Stinton, A., Namikawa, L. et al. (2005), 'Parallel adaptive numerical simulation of dry avalanches over natural terrain', *Journal of Volcanology and Geothermal Research* **139**(1-2), 1–21.

Spiller, E. T., Bayarri, M., Berger, J. O., Calder, E. S., Patra, A. K., Pitman, E. B. & Wolpert, R. L. (2014), 'Automating emulator construction for geophysical hazard maps', *SIAM/ASA Journal on Uncertainty Quantification* **2**(1), 126–152.

Wang, X. & Berger, J. O. (2016), 'Estimating shape constrained functions using a new class of gaussian processes', *SIAM/ASA Journal on Uncertainty Quantification* **4**(1), 1–25.

Westerink, J. J., Luettich Jr, R. & Scheffner, N. (1993), Adcirc: an advanced three-dimensional circulation model for shelves, coasts, and estuaries. report 3. development of a tidal constituent database for the western north atlantic and gulf of mexico, Technical report, COASTAL ENGINEERING RESEARCH CENTER VICKSBURG MS.