

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.



National
Laboratories

Computational Efficient Estimation of the Extreme Event Probability of the Mass Loss of Antarctic Glaciers



Presented by:

K. Liegeois, M. Perego & G. Stadler



Sandia National Laboratories



11th National Congress on Computational Mechanics

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract number DE-NA0003525.



Sandia National Laboratories is a multimission



- ▶ Introduction:
 - ▶ Land ice simulation and UQ big picture;
 - ▶ Extreme event probabilities;
- ▶ Optimization approach;
 - ▶ Computation of the most likely point;
 - ▶ Importance sampling;
- ▶ Numerical results;
 - ▶ Implementation and test problem;
 - ▶ Probabilities and performance;
- ▶ Conclusions and future work.

Unlikely events with high impacts:

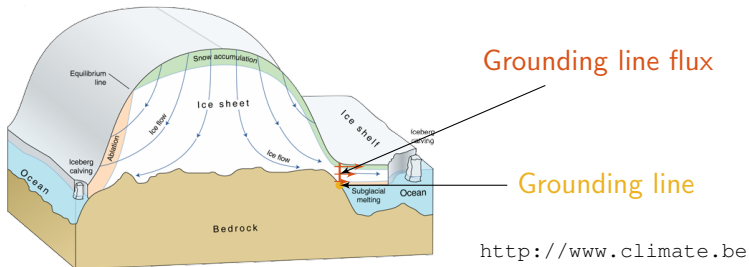
- ▶ Are difficult to quantify with standard approaches;
events with small probability require a large number of samples to be evaluated precisely;
- ▶ Can have severe consequences (even if they are rare);
large tsunamis are very rare but can lead to high losses.

When unlikely events have high impacts, it is important to have a precise estimation of their probabilities.

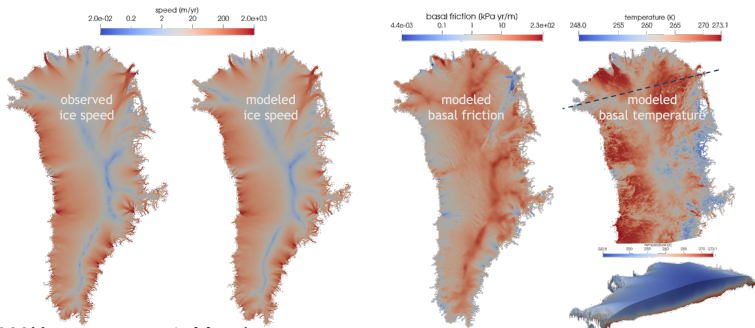
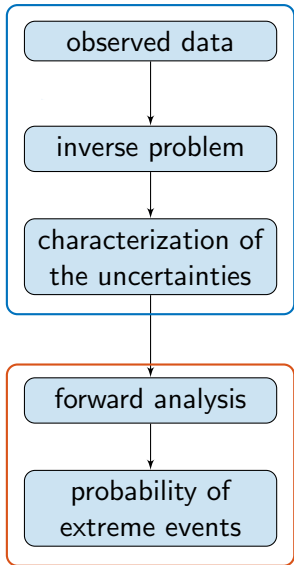
Unlikely events with high impacts for land ice simulation



In the case of land ice simulation, this will allow us to estimate smaller probabilities for different levels of sea level rise due to land ice mass loss.



Random parameter: the basal friction between the ice sheet and the bedrock (random field),
Unlikely event of interest: to have a flux at the grounding line above a certain threshold.



300K parameters, 14M unknowns.

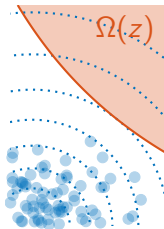
Initialization: 10 hours on 2k cores on NERSC Cori (Haswell),

The optimization is constrained by the coupled velocity-temperature solvers.

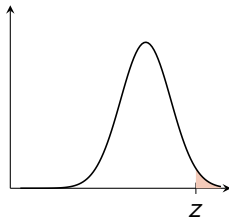
The work presented in this talk.

Given a n -variate Gaussian variable $\theta \sim \mathcal{N}_n(0, \mathbf{I})$ and a F parameter-to-event map (involved PDE solve), quantity of interest:

$$F : \theta \sim \mathcal{N}_n(0, \mathbf{I}) \rightarrow \mathbb{R}$$



$$\xrightarrow{F(\theta)}$$



Target: Estimate the measure of extreme event sets for $z \gg 0$
 $\Omega(z) := \{\theta : F(\theta) \geq z\}$ i.e., compute $\mathbb{P}(F(\theta) \geq z)$ when $\mathbb{P}(F(\theta) \geq z) \ll 1$.

The used strategy relies on finding $\theta^*(z)$ the most likely point above the threshold which can be computed by solving the PDE-constrained optimization problem:

$$\theta^*(z) = \arg \min_{\theta \in \Omega(z)} I(\theta),$$

where $I(\theta) = \frac{1}{2} \|\theta\|^2$ for $\theta \sim \mathcal{N}_n(0, I)$ as $pdf(\theta) = c \exp(-I(\theta))$.

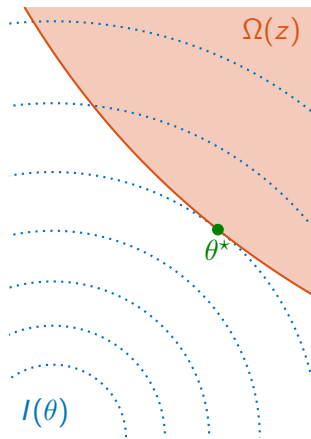
Then, the probability can be approximated as follows:

$$\mathbb{P}(F(\theta) \geq z) \approx C_0(z) \exp(-I(\theta^*(z))), \quad \text{as } z \rightarrow \infty,$$

where $C_0(z)$ is a sub-exponential prefactor.

The method relies on 2 steps:

- ▶ Compute the most likely point θ^* ,
- ▶ Compute the prefactor $C_0(F(\theta^*))$.



Under some assumptions, the minimizer over $\Omega(z)$ is reached on $\partial\Omega(z)$ and the inequality constraint is now active:

$$\theta^*(z) = \arg \min_{\theta \in \partial\Omega(z)} I(\theta).$$

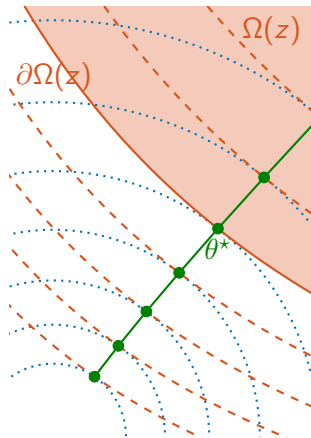
We used a quadratic penalty method as follows:

$$\theta^*(z) = \arg \min_{\theta} I(\theta) + \alpha (F(\theta) - z)^2,$$

where α is a penalty weight that should be large enough such that $F(\theta^*) \approx z$.

The used strategy is the following:

- ▶ Select an increasing sequence $z_1, \dots, z_m$ of quantity of interest,
- ▶ For a given z_i , solve the corresponding optimization problem using $\theta^*(z_{i-1})$ as the initial guess,
- ▶ Deduce the sequence $\theta^*(z_1), \dots, \theta^*(z_m)$.



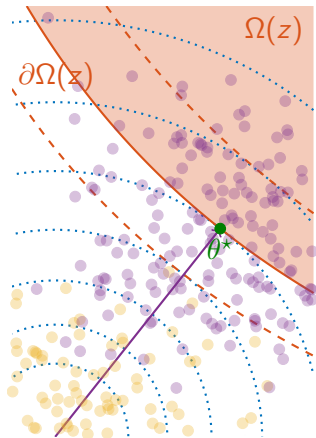
One strategy is to use an Importance Sampling (IS) strategy:

- ▶ to draw N random samples $\theta_1, \dots, \theta_N$ from the initial distribution,
- ▶ for a given value of z_i :
 - ▶ shift the samples: $\tilde{\theta}_k = \theta_k + \theta^*(z_i)$,
 - ▶ evaluate F for all the N shifted samples,
 - ▶ evaluate:

$$P_N^{IS}(z) = e^{-I(\theta^*)} \frac{1}{N} \sum_{k=1}^N \left[\mathbb{1}_{\Omega(z)}(\tilde{\theta}_k) \exp \left(- \left(\tilde{\theta}_k - \theta^* \right)^\top \theta^* \right) \right],$$

where z can be different from $F(\theta^*)$.

- ▶ Advantages: $\mathbb{E} [P_N^{IS}(z)] = P(z)$ and variance in $1/N$,
- ▶ Challenges: this approach requires $N \times m$ evaluations of F where m is the number of z values.



Monte Carlo strategy



Optimization strategy with shifted Monte Carlo sampling



Optimization strategy with shifted Monte Carlo sampling



- ▶ FE software: Albany,
- ▶ PDE constrained optimizer: algorithm: trust region, software: ROL,
- ▶ Non-linear solver: algorithm: Newton solver, software: NOX,
- ▶ Linear solver: algorithm: GMRES, software: Belos,
- ▶ Preconditioner: algorithm: Schwarz, software: FROSch,
- ▶ Preconditioner: algorithm: multigrid, software: MueLu,
- ▶ First and second derivative computation: algorithm: automatic differentiation (AD), software: Sacado,
- ▶ Reduced Hessian and Gradient vector product computed using ROL and AD.



RAPID OPTIMIZATION LIBRARY





PyAlbany allows to easily and quickly:

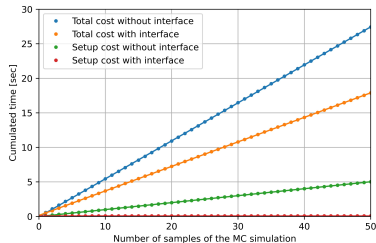
- ▶ Use Albany without C++ or bash knowledge (convenient for students),
- ▶ Prototype applications that require multiple Albany solves,
- ▶ Enable fast pre-processing and post-processing in Python,
- ▶ Use Python as a glue language to couple Albany with other software:
 - ▶ For UQ methods (PyDakota),
 - ▶ For machine learning (TensorFlow, Keras, Scikit-learn),
 - ▶ For plotting (Matplotlib, Paraview).

```

1 import numpy as np
2 from numpy.random import default_rng
3 from PyAlbany import Utils
4
5 parallelEnv = Utils.createDefaultParallelEnv()
6
7 # Create an Albany problem:
8 filename = "input_scalar.yaml"
9 paramList = Utils.createParameterList(
10     filename, parallelEnv
11 )
12
13 problem = Utils.createAlbanyProblem(paramList, parallelEnv)
14
15 parameter_map_0 = problem.getParameterMap()
16 parameter_0 = Utils.createVector(parameter_map_0)
17
18 parameter_0_view = parameter_0.getLocalView()
19
20 N = 200
21 p_min = -2.
22 p_max = 2.
23
24 # Generate N samples randomly chosen in [p_min, p_max]:
25 rng = default_rng()
26 p = rng.uniform(p_min, p_max, N)
27 QoI = np.empty(N, dtype=float)
28
29 # Loop over the N samples and evaluate the quantity of interest:
30 for i in range(0, N):
31     parameter_0_view[i] = p[i]
32     parameter_0.setLocalView(parameter_0_view)
33     problem.setParameter(0, parameter_0)
34
35     problem.performSolve()
36
37     response = problem.getResponse(0)
38     QoI[i] = response.getLocalView()[0]

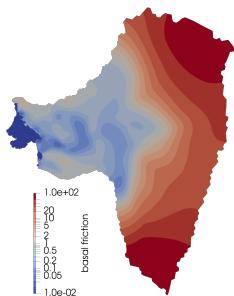
```

Listing 3: Monte Carlo example.

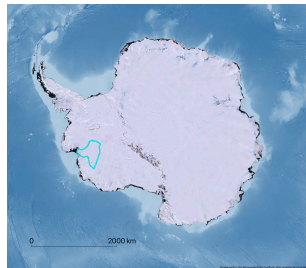
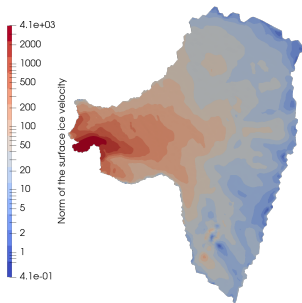


Mean case:

Basal friction:



Surface ice velocity:



courtesy of T. Hillebrand,
modified from Quantarctica

The velocity is faster if the friction is smaller, the quantity of interest is the flux at the grounding line; we expect the extreme events to be associated to smaller basal friction values.

- ▶ Random parameter: the basal friction represented using a log-normal random field and a KL expansion with 50 modes:

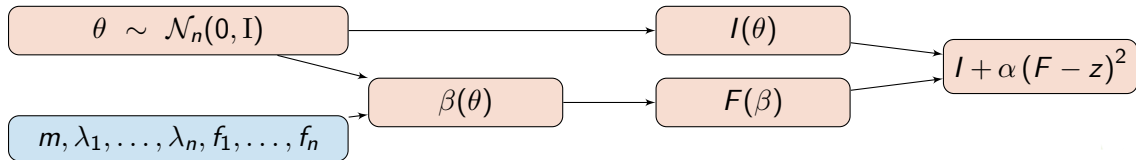
$$\beta(\mathbf{x}, \theta) = e^{\log(m(\mathbf{x})) + \sum_{i=0}^n \sqrt{\lambda_i} \varepsilon_i(\theta) f_i(\mathbf{x})},$$

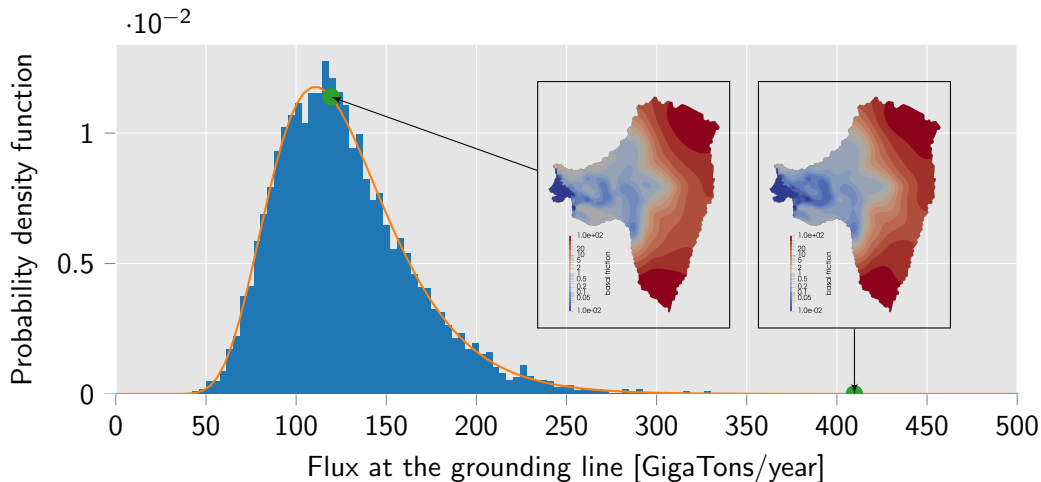
$$\int_{\Omega} \text{Cov}(\mathbf{x}, \mathbf{x}') f_i(\mathbf{x}') d\mathbf{x}' = \lambda_i f_i(\mathbf{x}),$$

$$\text{Cov}(\mathbf{x}, \mathbf{x}') = \sigma e^{-\frac{\|\mathbf{x} - \mathbf{x}'\|}{\ell}},$$

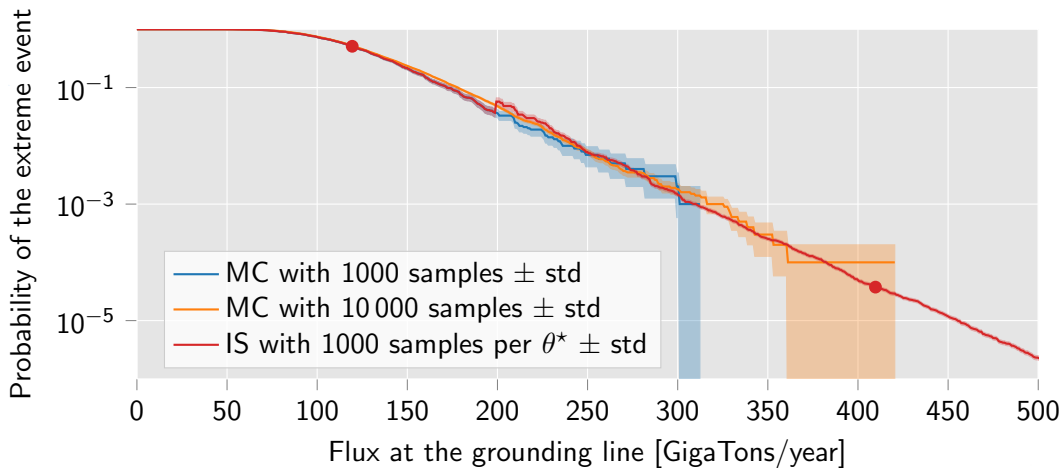
with $\sigma = 0.3$, $\ell = 50$ km, and $n = 50$. Those modes allow to capture 99% of the variance.

- ▶ Quantity of interest: flux at the grounding line.
- ▶ PDE: steady state first order Stokes equation, Blatter-Pattyn model.





Histogram of 10 000 samples, the orange curve is a log-normal distribution that fits the data the best.



For smaller values of z , all the methods are consistent. When increasing z , the MC approach needs more and more samples to be consistent with the importance sampling.

Average cost per simulation (measured on Skylake):

	Wall-clock time	Relative cost
To compute $F(\theta)$	$c_F = 39 \text{ sec}$	1
To compute $\theta^*(z)$	$c_{\theta^*} = 11842 \text{ sec}$	304

Expected cost per method:

	$\mathbb{E}[c]$	$\mathbb{E}[c/c_F]$
Monte Carlo	$N_{MC} c_F$	N_{MC}
Importance Sampling per θ^*	$c_{\theta^*} + N_{IS} c_F$	$304 + N_{IS}$

Comparison of Monte Carlo with $N_{MC} = 10\,000$ and importance sampling with $N_{IS} = 1000$:

$$\frac{N_{MC}}{304 + N_{IS}} = \frac{10\,000}{1304} = 7.66,$$

the importance sampling is 7 times faster.

The more extreme the event, the more samples will be required for the Monte Carlo method and the more efficient the importance sampling will be compared to standard Monte Carlo.



Conclusions:

- ▶ Discussion of the usage of optimization strategies to compute the probability of extreme events,
- ▶ Discussion of the implementation using open source libraries and software,
- ▶ Computation of the extreme event probabilities of high fluxes at the grounding line of the Thwaites glacier,
- ▶ Performance comparison of the proposed approach with the standard Monte Carlo method.

Current work:

- ▶ Move towards transient analysis.

Future work:

- ▶ Consider solving the constrained problem instead of using a quadratic penalty method,
- ▶ Consider larger problems,
- ▶ Use characterization of the uncertainties computed using the inverse problem,
- ▶ Deduce probability of extreme sea level rise due to land ice mass loss for the future.