Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_06-008-cd

ISGSR 2022

Probabilistic Slope Stability Assessment with Adaptive Monte Carlo Importance Sampling

Anton van der Meer[1], Rob Brinkman[1] and Wim Kanning[1,2]

[1]Department of Geo Engineering, Deltares, Delft, The Netherlands.
[2]Department Hydraulic Engineering, Delft University of Technology, Delft, Netherlands.
E-mail: anton.vandermeer@deltares.nl, rob.brinkman@deltares.nl, wim.kanning@deltares.nl

Abstract: In the Netherlands, levee segments are assigned a maximum allowable annual probability of flooding. One of the failure mechanisms leading to flooding is instability of the inner slope of the levee. The stability of a levee is generally assessed with a Limit Equilibrium Model (LEM), with calculation times ranging from a few to tens of seconds for a single model run. The LEM has to be coupled with a probabilistic method in order to assess whether target failure probabilities are complied with. Crude Monte Carlo Sampling results in unacceptably long calculations times as in practice, times up to an hour are tolerated. The First Order Reliability Method on the other is faster but often not robust enough due to convergence problems for stability analyses. In this paper, we adopt and apply an efficient and robust algorithm, the Adaptive Monte Carlo Importance Sampling. The approach consists of several loops of MC-IS calculations, in which the design point of one loop is the center of the importance distribution of the next loop. The approach is suited for unsupervised execution and parallelization. We develop decision and convergence criteria to guide when to break a loop and start a new loop or stop the calculation. This approach has now been successfully applied in several levee reinforcement projects. We highlight one of these applications for a selected case study and discuss the lessons learned.

Keywords: Reliability; Importance Sampling; Slope Stability; Levees; Application.

1 Introduction

In the Netherlands, levee segments are assigned a maximum allowable annual probability of flooding (Kind 2014; Van der Most et al. 2014). One of the failure mechanisms leading to flooding is instability of the inner slope of the levee. The maximum allowable annual probability of failure for stability at a specific location is roughly 10^{-5} to 10^{-7}. The stability of a levee is generally assessed with a Limit Equilibrium Model (LEM), with calculation times ranging from a few to tens of seconds for a single model run. Whether a levee complies with the target should be assessed using probabilistic methods with LEM computations as input. Using Crude Monte Carlo (CMC) leads to unacceptably long calculations times as millions of computations, and thus model runs, are needed. In practice calculation times up to an hour are tolerated but aimed is for five to ten minutes (the time for a cup of coffee). An alternative and faster approach is the gradient based First Order Reliability Method (FORM), but it is often not robust enough due to convergence problems for stability analyses, especially in case of different relevant slip planes. The aim of this paper is to develop and apply an efficient and robust probabilistic method, the Adaptive Monte Carlo Importance Sampling, to obtain the probability of an instability.

2 Slope Stability

In this paper, slope stability is calculated with the Lift-Van LEM, Van (2001). The model computes a stability factor (SF), defined as strength (R) divided by load (S), and is incorporated in the software D-Stability, Deltares (2020). The governing slip plane (i.e. the one with the lowest stability factor, $SF = R/S$) is searched with a particle swarm algorithm, Deltares (2020). The model uncertainty is accounted for with m_d. The limit state function ($Z = R - S$) is given in Eq. (1). A slope instability occurs when $Z < 0$.

$$Z = SF(\boldsymbol{X}) \cdot m_d - 1 \tag{1}$$

$\boldsymbol{X}$ are the input parameters for the LEM. In this study, the soil strength parameters (the shear strength ratio, pre-overburden pressure and friction angle) are based on undrained (impermeable soils) and drained (permeable soils) soil models. Together with water level and model uncertainty, these are incorporated as the relevant random variables. The soil strength parameters are values averaged over the soil layer. Higher water levels, lower strength parameters and a smaller model factor lead to a lower limit state value. The limit state function is monotonous. Typically, around 10 to 20 stochastic variables are used for a calculation.

3 Probabilistic Method

The probabilistic method proposed is Adaptive Monte Carlo Importance Sampling. The approach consists of several loops of Monte Carlo Importance Sampling calculations, in which the design point of one loop is the center of the importance distribution of the next loop. In this section is firstly focused on Monte Carlo Importance Sampling in Section 3.1 and secondly on the 'Adaptive' part, the loops, in Section 3.2.

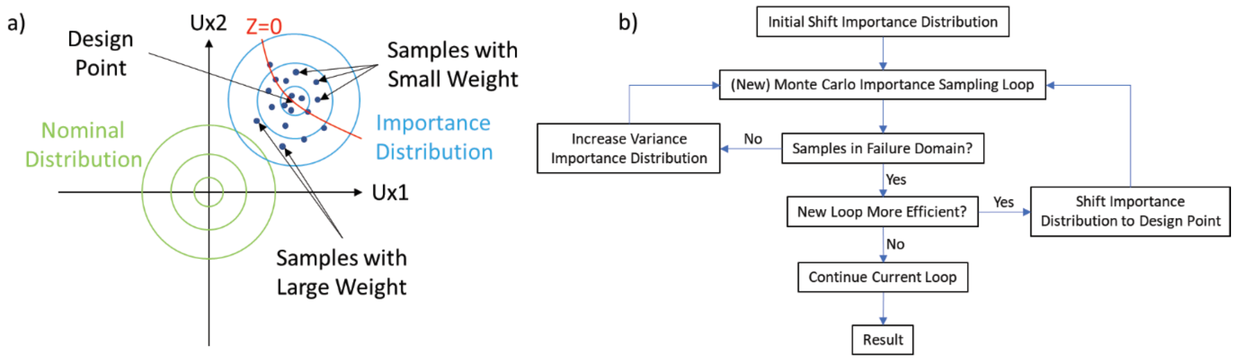

Figure 1. a) Monte Carlo Importance Sampling and **b)** The Adaptive Monte Carlo Importance Sampling approach proposed.

3.1 Monte Carlo Importance Sampling

3.1.1 Basic Monte Carlo Importance Sampling

In our approach we use nominal (standard normal) distributions for all stochastic variables. There will be a transformation from these values (the *U*-space) to the actual distribution of a variable (e.g. a log normal distribution) and from there to the physical values (the *X*-space).

While Crude Monte Carlo Sampling draws samples from the nominal distributions of the stochastic variables, Monte Carlo Importance Sampling draws samples from an importance distribution, Figure 1a. By using the importance distribution more samples are drawn in the region of interest and consequently less samples are required to get an accurate estimate of the probability of failure. To allow for this, in the (unbiased) estimate of the probability of failure is corrected for the fact that the importance distribution instead of the nominal distribution is used (Kahn and Harris 1951; Owen 2013), as follow:

$$\hat{P}_{f,MCIS} = \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)p(X_i)}{q(X_i)} = \frac{1}{n}\sum_{i=1}^{n} f(X_i)w(X_i) \qquad (2)$$

In which n denotes the number of samples, $p(X_i)$ the probability density function of the nominal distribution, $q(X_i)$ the probability density function of the importance distribution and X_i are the samples drawn from the importance distribution. The ratio $p(X_i)/q(X_i)$ is also referred to as $w(X_i)$, the weight of sample X_i. The function $f(X_i)$ represents the model result. $f(X_i)$ equals 1 if the levee has failed for sample X_i, in case the limit state function is smaller than zero ($Z(X_i) < 0$). The function $f(X_i)$ equals 0 if the levee has not failed for sample X_i and thus $Z(X_i) \geq 0$). The procedure is shown in Figure 1.(a) for two normalized variables, X_1 and X_2, both in standard normal space (noted by U).

In theory the optimal importance distribution, $q(x)$, is proportional to $f(x)p(x)$, see Kahn and Marshall (1953). In practice, it is often aimed to shift the center of the importance distribution to the most likely point of failure, the design point. Since this point is not known *a priori*, this has to be chosen smartly.

3.1.2 Convergence Criteria

To get an accurate estimate of the probability of failure, the region of interest is the region near the design point in the failure domain, because they have a relevant contribution to the probability of failure. In the failure domain, the larger the weight of a sample, the closer the sample is to the design point. The sample in the failure domain the closest to the design point is thus: $\max(f(X_i)w(X_i))$. If this sample weight is significantly larger than all the weights of all the other samples in the failure domain, effectively there is only one relevant observation in the failure domain. In this case: $\max(f(X_i)w(X_i)) \approx \sum(f(X_i)w(X_i))$. If the ratio between the maximum and the sum of the sample weights in the failure domain is small, it suggests that there are multiple samples in the domain of interest. We strive to this situation, because it reduces the influence of the randomness in selecting samples. In Eq. (3) the convergence relation is defined as the ratio between the maximum and the sum of the sample weights in the failure domain. This ratio should become smaller than ε for the algorithm to convergence.

$$\frac{\max(f(X_i)w(X_i))}{\sum(f(X_i)w(X_i))} < \varepsilon \tag{3}$$

3.2 Adaptive Monte Carlo Importance Sampling

3.2.1 Basic Adaptive Monte Carlo Importance Sampling

As stated in Section 3.1.1, the aim of this paper is to locate the importance distribution at the *a priori* unknown design point. With knowledge of the physics it is possible to make an educated guess of the design point. Alternatively, one could use the design point of a previously performed calculation. This latter approach is used in an automated procedure resulting in Adaptive Monte Carlo Importance Sampling, see Figure 1b for a scheme. Several successive MC-IS calculations, which we refer to as loops, are performed. For each loop, the center of the importance distribution is the design point from the previous loop. In case no failing samples were found in the previous loop, the importance distribution is widened by increasing its variance. By default, the importance distribution of the initial loop is the nominal distribution, but this can be modified or more efficiently be determined with a preliminary sensitivity study.

3.2.2 Decision Criteria

The decision to interrupt a Monte Carlo Importance Sampling loop and start a new loop is based on the expected number of additional samples $n_{additional}$ needed to meet the convergence criteria. To estimate the expected number of additional samples the factor with which the denominator in Eq. (3) must be multiplied to meet the convergence criteria is calculated. Knowing the number of samples taken already, the estimate of the additional samples needed is obtained using Eq. (4). The assumption in this approach is that the numerator in Eq. (3) does not change. If the numerator does increase (significantly), which is especially the case in the beginning of a loop, the additional number of samples required is underestimated. The latter is not a problem, since we do want the loop to continue if there is not at least one sample close to the design point.

$$n_{additional,estimate} = n\left(\left(\left(\frac{\max\left(f(X_i)w(X_i)\right)}{\sum\left(f(X_i)w(X_i)\right)}\right)\Big/\varepsilon\right) - 1\right) \tag{4}$$

If the expected number of additional samples required is larger than the number of samples we expect to need if we would have placed the importance distribution at the ‘ideal’ location, the design point, we choose to start a new loop. The required number of samples, in case the importance distribution is shifted to the design point, depends on the probability of the instability and is empirically derived, Eq. (5). The derivation is based on a linear limit state line ($Z = 0$). For $\beta = 0$, the ideal importance distribution equals the nominal distribution and consequently all the sample weight equal one. To meet the convergence criterium, ε, $1/\varepsilon$ samples are required in the failure domain. Since around half of the samples are in the failure domain, in total around $2/\varepsilon$ are required. For $\beta > 0$, the weights in the failure domain are decreasing if the distance of a sample to the design point increases, consequently more samples are required to meet the convergence criterium. This effect is stronger for a larger β.

$$n_{required,ideal,estimate} = \frac{2(\beta + 1)}{\varepsilon} \tag{5}$$

If $n_{additional,estimate} > n_{required,ideal,estimate}$, the decision is made to start a new loop. If $n_{additional,estimate} \leq n_{required,ideal,estimate}$, the decision is made to continue with the current loop. A minimum number of samples per loop can be assigned and a maximum number of samples per loop without any failing samples can be assigned.

4 Application Example

We apply Adaptive Monte Carlo Importance Sampling for a simplified example, as explained next. Consider the levee in Figure 2. The subsoil consists of three Holocene soil layers on top of the Pleistocene Sand Layer. The stochastic soil parameters of the layer-averaged strength parameters are given in Table 1. The model uncertainty has a lognormal distribution with a mean of 1.005 and a standard deviation of 0.033. The calculation is performed conditional on a river water level of 4 meters relative to the datum (NAP). For the given location, this water level has an exceedance frequency of 1/10 000 per year. It is common in The Netherlands to first compute the conditional probability of flooding for different water levels and subsequently integrate the fragility curve with the probability density function of the water level (Kanning 2016).

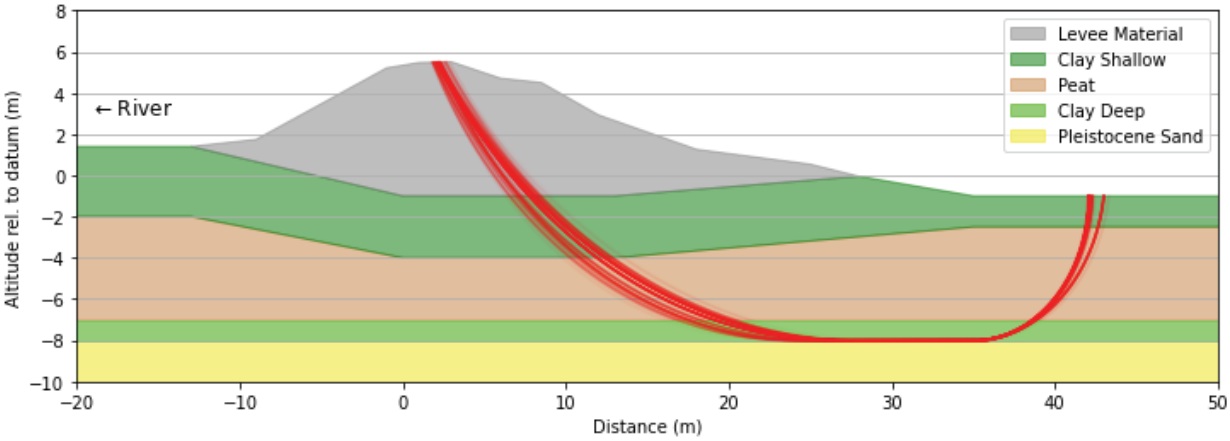

Figure 2. The considered cross-section. The slip planes are the samples with limit state value lower than zero.

Table 1. The soil strength parameters. All distributions are lognormal.

Soil Layer	Strength Model	Friction Angle	Shear Strength Ratio	Pre-Overburden Pressure (kPa)*
Levee Material	Drained	μ=34; σ=2.5	-	-
Clay Shallow	Undrained	-	μ=0.34; σ=0.025	μ=27; σ=09
Peat	Undrained	-	μ=0.38; σ=0.025	μ=11; σ=16
Clay Deep	Undrained	-	μ=0.27; σ=0.020	μ=24; σ=15

* The Pre-Overburden Pressures of the Holocene layers (Clay Shallow, Peat and Clay Deep) are fully correlated.

The initial importance distribution is (for this example) chosen to be identical to the nominal distribution. In the first loop after 100 samples, still no failing samples are found. Consequently, the choice is made to start a new loop (loop 2) with an increased variance ($\sigma_{importance} = 2 \cdot \sigma_{nominal}$). In Figure 3a it can be seen that after 48 samples a failing sample is found. After 100 samples, the failure probability is 10^{-5}, the convergence (Figure 3b) is about 1 (meaning that only one failing sample is relevant) and the estimated number of additional samples needed is almost 891 while the expected number of samples for an ideal shifted distribution is with 105 significantly lower. The decision is made to start a new loop (loop 3) and shift the distribution to the design point of loop 2. In loop 5 after 100 samples the estimation is that less additional samples are needed (21 samples) than the expected number of samples for an ideal shifted distribution (80 samples) and the decision is made to continue the loop. Finally, after 125 samples the Monte Carlo analysis has sufficiently converged. The probability of failure is $1.4 \cdot 10^{-3}$.

Note that the first two or three loops can be prevented if the initial importance distribution is already chosen wisely. In multiple of the application projects the initial importance distribution is automatically based on a sensitivity study.

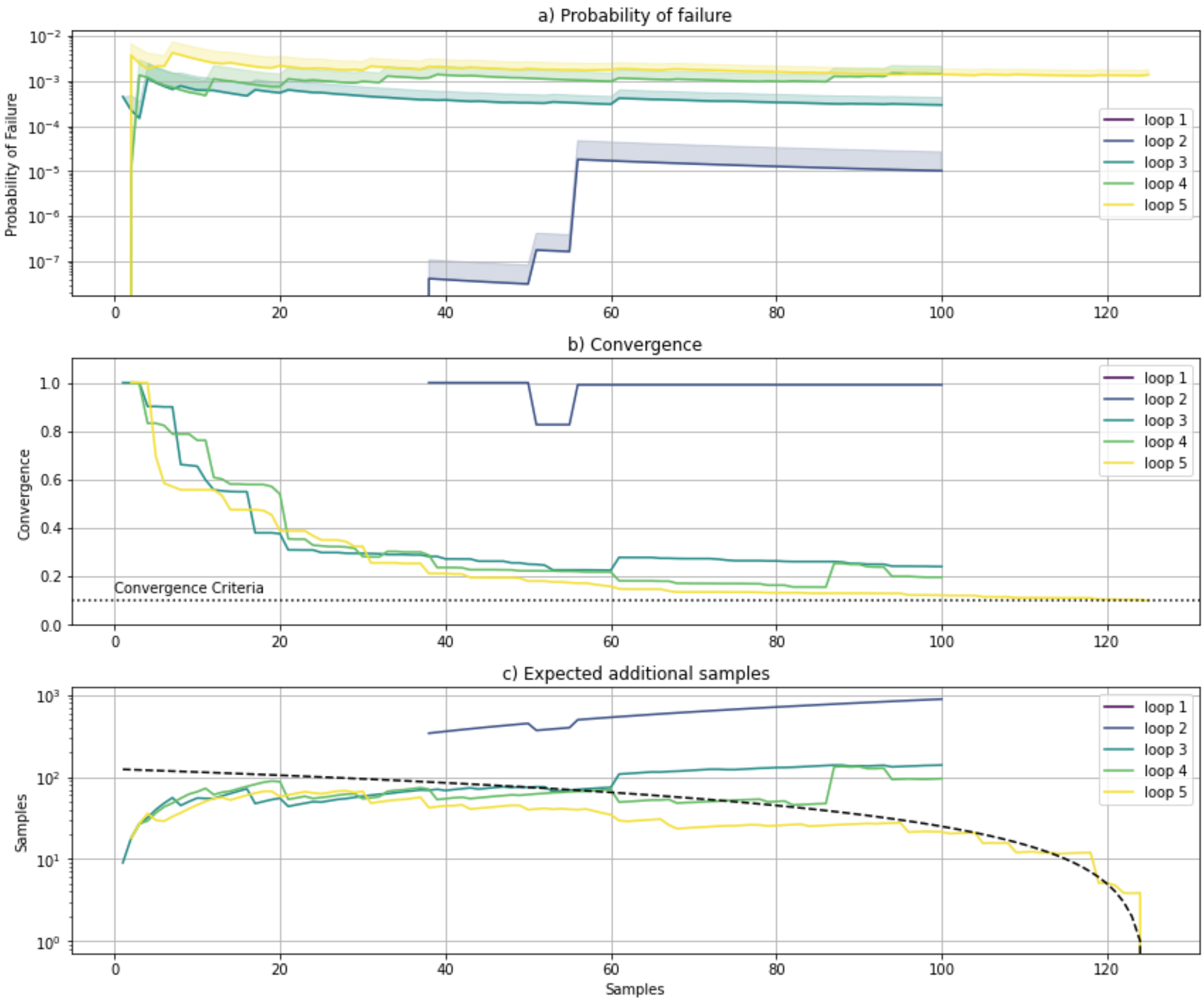

Figure 3. a) The probability of failure, Eq. (2), of the consecutive loops. The shadow gives the 95th percentile of the probability of failure (precision). **b)** The convergence criterium of the consecutive loops, Eq (3). If there is only one failing sample or one failing sample with a significant larger weight than the others, the convergence is (almost) one. The convergence criterium is given with the horizontal line. **c)** The expected number of additional samples required for convergence of the consecutive loops, Eq. (4). The estimated number of samples required in case the importance distribution is shifted to the design point, Eq. (5), is for the given reliability index around 100 samples. With the striped line the actual number of additional samples that were required in the last loop is given. Note that the line for loop 1 is not visible in the graph since no failing samples ($Z < 0$) are found in this loop.

5 Application in Levee Reinforcement Projects

The procedure of Adaptive Monte Carlo Importance Sampling is integrated in the Probabilistic Toolkit (software), Deltares (2021), and in the past years in a supervised setting applied in different levee assessment and levee reinforcement projects. In the first project the adaptive part was executed manually, and the convergence criterium of Monte Carlo Importance Sampling was based on the standard deviation of the calculated probability of failure. Roughly 1 000 calculations were required for one probabilistic calculation. The procedure enabled the use of a full probabilistic design approach instead of the semi-probabilistic design approach generally used in The Netherlands, see Ministerie van Infrastructuur en Milieu (2017). The use of a full probabilistic assessment led in this project, Tao et al. (2021), to a significant reduction of the required design solution (33% of the berm length). Presenting the slip plane of all the failing samples with an opacity based on the weight of the sample, as in Figure 2, led to understanding by the decisionmakers and gave a useful indication of the type of design solutions required.

In the consecutive two projects, the algorithm was further modified to add on the adaptive scheme, and the shift of the importance distribution of the initial loop was based on a sensitivity study, Deltares (2021). The choice of when starting a new loop was based on the ratio of failing samples, which in many cases led to continuation of a loop although it would have been more effective to start a new loop.

With the knowledge and experience gathered in those projects the convergence and decision criteria are optimized as presented in this paper. Those criteria are now being implemented in the Probabilistic Toolkit, to be used in future levee reinforcement projects. The objective is to use the procedure integrated in D-Stability in an unsupervised setting which allows for fast computations involving complex geometries and multiple critical slip circles.

6 Discussion and Future Developments

The aim is to use the probabilistic method in practice in an unsupervised setting with a limited calculation time. In several projects it is already shown that the procedure can be used in practice in a supervised setting. The new convergence and decision criteria enable an unsupervised implementation.

In between the decision criteria, the samples within a loop can be executed parallel, further reducing the calculation time. The suitability for parallelization is one of the main advantages with respect to surrogate modelling. The Probabilistic Toolkit enables already a Cloud connection to facilitate parallelization.

Since the limit state function in the considered case is monotonous the risk of obtaining the probability of a local minimum of the limit state function (which would hinder both MC-IS and especially FORM application) is limited. Optionally Gaussian Process Regression (GPR) can be used to explore the parameter space using the samples of all the consecutive loops (and eventually adding samples in unknown regions). GPR is already used in a study to exclude the existence of minima with a higher probability of failure than obtained with Adaptive Monte Carlo Importance Sampling. The use of GPR to initiate Multiple Importance Sampling is subject for further study.

7 Conclusion

The use of the Adaptive Monte Carlo Importance Sampling approach presented enables levee reinforcement projects to use a full probabilistic assessment for slope stability. Within 1 000 samples an acceptable convergence is reached, and the procedure is suitable for parallelization. With a calculation time of maximum tens of seconds for one sample and a common amount of (virtual) CPUs, a probabilistic result can be obtained within several minutes, the time for a cup of coffee.

References

Deltares 2020. D-Stability. Installation manual, tutorial and scientific manual. *D-GEO Suite Stability*, Version 2020.03.3, September 2020. https://www.deltares.nl/en/software/d-stability/.

Deltares 2021. Probabilistic Toolkit. *User Manual*. Version 2, 20 April 2021. https://www.deltares.nl/en/software/probabilistic-toolkit-ptk/.

Kahn, H. and Harris, T. E. (1951). Estimation of particle transmission by random sampling. *National Bureau of Standards applied mathematics series*, 12:27–30.

Kahn, H. and Marshall, A. (1953). Methods of reducing sample size in Monte Carlo computations. *Journal of the Operations Research Society of America*, 1(5):263–278.

Kanning, W, Schweckendiek, T 2016. Handreiking Faalkansanalyse en Faalkans Updating. Groene Versie - Macrostabiliteit Binnenwaarts. *Deltares Report* 1230090-032-GEO-0016 (in Dutch).

Kind, M. 2014. Economically efficient flood protection standards for the Netherlands. *Journal Flood Risk Management,* Volume 7, Issue 2 (June 2014), 103–117. https://doi.org/10.1111/jfr3.12026.

Ministerie van Infrastructuur en Milieu 2017. Handreiking ontwerpen met overstromingskansen. *Veiligheidsfactoren en belastingen bij nieuwe overstromingskansnormen*. February 2017 (in Dutch).

Owen, A.B. 2013. Monte Carlo Theory, *Methods and Examples*.

Tao, P.Y., Kwakman, L., Van der Meer, A.W., Van Zuylen, J.A., Van Veen, C.J. 2021. Probabilistisch ontwerpen voor stabiliteit bij dijkversterking SAFE. *Land en Water*. Volume 5. May 2021 (in Dutch).

Van, M. 2001. New approach for uplift induced slope failure. *In XVth International Conference on Soil Mechanics and Geotechnical Engineering, Istanbul*, pages 2285-2288.

Van der Most, H., Tanczos, I., De Bruijn, K. M., & Wagenaar, D. 2014. New risk-based standards for flood protection in the Netherlands. *In Proceedings of 6th International Conference on Flood Management, Sao Paulo, Brazil*, Sept 2014.