TU Delft

Delft University of Technology

Document Version
Final published version

Citation (APA)
Huang, W., A. Hicks, M., & Guan, Z. (2026). Calibration of 2D load for railway embankment stability assessment based on 3D reliability analysis using RFEM. In J. Pistrol, D. Adam, & H. F. Schweiger (Eds.), *Proceedings of the 21st ICSMGE* (pp. 5483-5486). Article 1395 ÖGG. https://doi.org/10.53243/ICSMGE2026-1395

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-1395

Calibration of 2D load for railway embankment stability assessment based on 3D reliability analysis using RFEM

Wei Huang, Michael A. Hicks, Zheng Guan
Department of Geoscience & Engineering, TU Delft, Delft, The Netherlands, W.Huang-1@tudelft.nl

ABSTRACT: The Random Finite Element Method (RFEM) is used to calibrate the 2D strip load for the deterministic stability assessment of a railway embankment, considering the effects of 3D train loads and spatial variability of the soil undrained shear strength. 3D random fields are generated to simulate the soil strength variability using different scales of fluctuation in the horizontal plane, and finite element analyses are conducted as part of a Monte Carlo simulation. Each RFEM analysis comprises 500 realizations, from which the probability distribution of undrained bearing capacity is obtained. This allows the 2D design load and resistance to be calibrated against a target reliability level in light of 3D soil spatial variability and load distribution. The results highlight the importance of 3D reliability analysis for the assessment of railway embankments, and the proposed calibration approach provides a reliability-informed 2D solution with improved accuracy.

KEYWORDS: Embankment reliability analysis, Random Finite Element Method, 3D soil spatial variability, 2D load calibration.

1 INTRODUCTION

In engineering practice, the bearing capacity of railway embankments is generally computed using two-dimensional (2D) plane strain analysis, including simplifications such as homogeneous soil layers and uniform strip loads. However, in reality, soil properties and applied loadings vary in three-dimensional (3D) space, leading to complicated 3D failure mechanisms. Therefore, 3D full probabilistic analysis is desirable to evaluate the reliability of railway embankments, in order to contribute to risk-informed and potentially more cost-effective design and maintenance decisions.

In this study, the design parameters to be used in 2D deterministic analysis are calibrated based on 3D full probabilistic analysis. The Random Finite Element Method (RFEM) (Fenton & Griffiths, 2008) is employed to determine the probabilistic undrained bearing capacity of an idealized railway embankment, considering two key parameters that characterize 3D effects in the system: 1) the length of the surcharge load applied on the embankment in the longitudinal direction, and 2) the horizontal scale of fluctuation that quantifies spatial variability of the soil shear strength. An in-house RFEM code has been developed for this purpose, based on that previously used in 2D (Hicks & Samy, 2002) and 3D (Hicks & Spencer, 2010; Hicks et al., 2014; Hicks & Li, 2018; Varkey et al., 2023) slope reliability analyses. By post-processing the RFEM results for the embankment bearing capacity, the reliability-based calibration of the 2D deterministic design load is demonstrated.

2 PROBLEM AND METHODOLOGY

2.1 *Problem definition*

The railway embankment is simplified as a slope loaded by its self-weight and a rectangular uniform surcharge on the top, as illustrated in Figure 1. The slope height of 2 m and the slope angle of 45° are within the typical ranges for railway embankments in the Netherlands, and the distance (in the x-direction) from the nearest edge of the surcharge q to the slope crest is assumed to be 2 m, which is a reasonable value for the location of the track structure that distributes operational loads over a width of approximately 3 m on the embankment top (CEN, 2021). Note that the dead weight of the track structure has been omitted for simplicity. The length (in the y-direction) of the embankment is set as 100 m, while the surcharge length L_q is varied (16, 30, 60, 100 m) to simulate different lengths of trains in operation.

2.2 *Finite element model*

In the finite element (FE) model, 20-node hexahedral elements of size 1.0×1.0×0.5 m (in the x-, y-, and z-directions) are adopted, except for along the sloping face where elements have been distorted, and a reduced integration scheme involving 2×2×2 Gauss points in each element is used. The mesh is fully fixed at the base, while the two vertical side faces (in the y-z plane) are constrained only in the x-direction and the two end faces (in the x-z plane) are allowed to move only in the z-direction. Such boundary conditions have been adopted and validated in previous research on 3D slope reliability (Hicks & Spencer, 2010; Hicks et al., 2014).

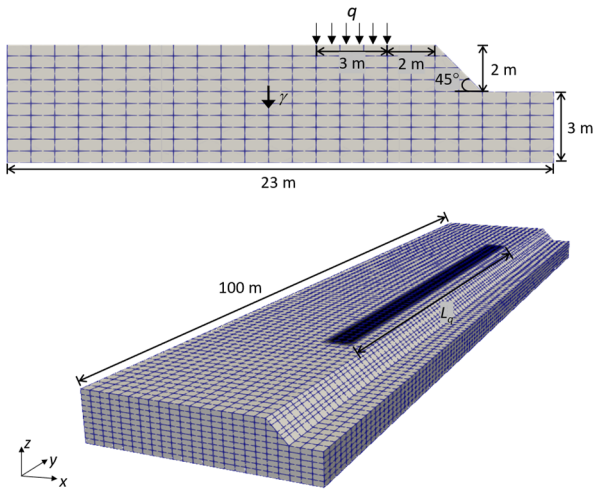


Figure 1. Problem geometry and finite element mesh in 2D and 3D.

Table 1. Soil parameter (mean) values.

Parameter	Symbol	Value	Unit
Unit weight	γ	20	kN/m^3
Undrained shear strength	c_u	20	kPa
Young's modulus	E	3×10^4	kPa
Poisson's ratio	ν	0.3	-

The embankment material is modelled using an elastic-perfectly plastic Tresca soil model under undrained (total stress) conditions, and the (mean) values of the relevant parameters are listed in Table 1. By incrementing the magnitude of q until failure has occurred, signaled by non-convergence after 500

equilibrium iterations, the limit load is taken as the undrained bearing capacity q_u of the embankment.

2.3 *RFEM reliability analysis*

In RFEM, the stochasticity of a spatially variable soil property is represented by a sufficiently large number of random fields (RFs), for which a variety of generation approaches exist (Fenton & Griffiths, 2008). Here, Local Average Subdivision (LAS) (Fenton & Vanmarcke, 1990) is adopted to discretize the RFs based on the probability density function (PDF) and the spatial correlation function of the material property. To characterize the RFs, the point statistics (i.e., mean and standard deviation), as well as the scales of fluctuation in the different spatial coordinate directions, are needed as input.

The undrained shear strength c_u is the only stochastic parameter in the current RFEM model. Based on various evidence from laboratory and field test data, as discussed by Hicks and Samy (2002), an underlying normal distribution is assumed for c_u, along with a spatial correlation structure modeled by a Gauss-Markov correlation function. The normal PDF of c_u is truncated at a lower threshold of 5 kPa, a typical lower-bound value of c_u for soft soils in the Netherlands, to avoid negative values, which eliminates less than 0.135% of the possible lower-tail values for relatively low coefficients of variation (i.e., COV < 0.25). The input statistics of c_u are selected from typical ranges and are listed in Table 2. Multiple values of the horizontal scale of fluctuation are included to investigate its influence on the results, as an accurate estimate is often unavailable due to sparse field data (Phoon, 2023).

Table 2. Point and spatial statistics of c_u.

Parameter	Symbol	Value	Unit
Mean	μ_c	20	kPa
Standard deviation	σ_c	4	kPa
Coefficient of variation	$COV_c=\sigma_c/\mu_c$	0.2	-
Horizontal scale of fluctuation (in x & y)	θ_h	4, 8, 16, 30, 60, 120	m
Vertical scale of fluctuation (in z)	θ_v	1	m

Following previous studies (e.g., Hicks et al., 2014), stationary anisotropic RFs are obtained for c_u by distorting (i.e., squashing in the vertical direction and/or stretching in the horizontal directions to give $\theta_v < \theta_h$) LAS-generated, isotropic, standard normal RFs, and then transforming them to the physical space of c_u based on μ_c and σ_c. These RFs are discretized in such a way that each Gauss point in the FE mesh corresponds to a unique RF cell. A total of 500 RF realizations have been generated for each RFEM analysis, which is sufficient for convergence of the output statistics (i.e., the mean and standard deviation of q_u). Subsequently, for a target probability of failure $P_{f,T}$, the design load q_d can be derived from the distribution of q_u based on the RFEM results, such that

$$P_{f,T} = P(q_u < q_d) \approx \sum_i^n I(q_u^i < q_d) \Big/ n \tag{1}$$

where n is the total number of realizations, q^i_u is the value of q_u from the i-th realization, and $I(\bullet)$ is an indicator function equal to 1 if the inequality inside is true or 0 otherwise.

2.4 *Reliability-based calibration*

Based on Equation (1), design load values corresponding to the same $P_{f,T}$ can be determined and compared across different scenarios, especially with respect to the q_d obtained from the 2D plane strain analysis that implies $L_q = \infty$, for the purpose of reliability-based design load calibration. To be more specific, let $q_{d,2D}$ and $q_{d,3D}$ denote the respective q_d values derived from 2D and 3D RFEM analyses of the embankment for the same value of $P_{f,T}$. Hence, a "load calibration factor" η_L can be defined as the ratio between $q_{d,2D}$ and $q_{d,3D}$ in order to convert a limited-length 3D load to a "reliability-equivalent" 2D strip load. This factor is expected to be a function of L_q and θ_h, as well as $P_{f,T}$, and is analogous to the load factor in the load and resistance factor design method (Salgado et al., 2011) if considering uncertainty in the 2D load due only to 3D effects of the load length and soil strength variability. Hence,

$$\eta_L = \frac{q_{d,2D}(\theta_h, P_{f,T})}{q_{d,3D}(L_q, \theta_h, P_{f,T})} \text{ or } q_{d,2D} = q_{d,3D} \cdot \eta_L(L_q, \theta_h, P_{f,T}) \tag{2}$$

On the other hand, the resistance (i.e., bearing capacity) in the 2D deterministic design can also be calibrated using a resistance factor η_R that takes into account the uncertainty due to soil spatial variability. In the same context as above, if $q_{u,det}$ is the deterministic solution obtained for the 2D homogeneous model characterized by the mean strength (i.e., $c_u = \mu_c$), the design resistance $q_{u,d}$ benchmarked against the 2D stochastic solution should ensure a probability of failure not exceeding the target one, which, in consistency with Equation (1), gives

$$q_{u,d} = q_{u,det} \cdot \eta_R\left(\theta_h, P_{f,T}\right) = q_{d,2D} \text{ or } \eta_R = \frac{q_{d,2D}\left(\theta_h, P_{f,T}\right)}{q_{u,det}} \tag{3}$$

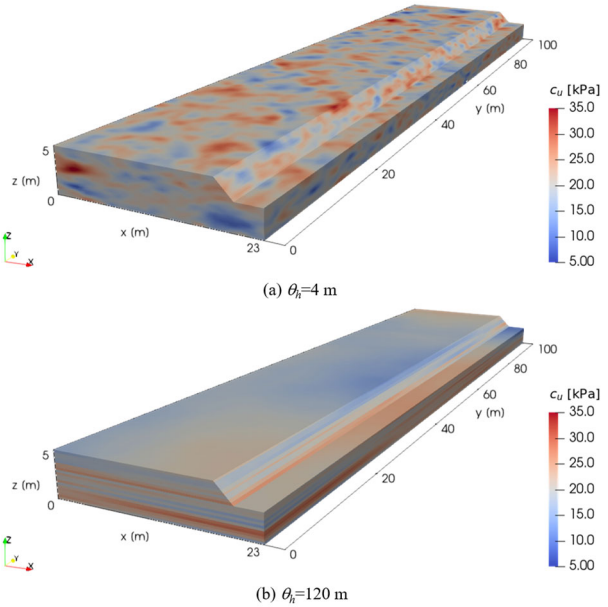
As such, the calibration factors η_L and η_R derived according to Equations (1)-(3) based on the RFEM results will lead to a reliability-informed 2D deterministic design, where the load $q_{d,2D}$ and resistance $q_{u,d}$ have incorporated the effects of 3D load length and soil variability.

3 RESULTS AND DISCUSSION

Figure 2 shows typical realizations of spatially variable c_u in the 3D domain for small and large values of θ_h relative to the embankment size. According to the observation of three distinct categories of failure mode for 3D heterogeneous slopes that are loaded by gravity only (Hicks & Spencer, 2010), a smaller θ_h results in a higher degree of soil property averaging over potential failure surfaces, whereas a larger θ_h indicates the soil has more of a layered status that forces rupture to propagate through horizontally semi-continuous weak zones. Hence, a relatively narrow spread of possible solutions with the mean response closer to the deterministic counterpart is expected for cases with a small θ_h. When a surcharge is added on top of the slope, the load length L_q relative to θ_h also plays a role in the probabilistic response, as demonstrated by the histograms in Figure 3 for θ_h = 16 m. The case of L_q = 16 m yields a similar distribution of q_u to that from the 2D stochastic model, which can be intuitively explained by potential failures being simultaneously constrained by L_q and a similar length of semi-continuous weak layers ($\approx \theta_h$) under the loaded area, so tending towards the plane strain condition. In contrast, a much larger L_q (e.g., 100 m) with respect to θ_h enables failure paths to leverage more spatial averaging.

Quantitatively, the probability of failure in these cases for a range of design loads is calculated according to Equation (1) and plotted in Figure 4, which exhibits essentially the cumulative distribution functions derived from the histograms. The curves for the different L_q values show similar shapes but are offset from each other for q_d approximately in the range from 40 to 90 kPa, where a larger L_q is associated with a higher P_f for a given q_d, with the 2D (dashed) curve lying amidst the 3D curves and not necessarily giving (un)safer predictions.

Outside the range, q_d < 40 kPa ensures negligibly small P_f, while q_d > 90 kPa leads to almost certain failure across all cases.



(a) θ_h=4 m

(b) θ_h=120 m

Figure 2. Examples of 3D random fields of c_u for different θ_h.

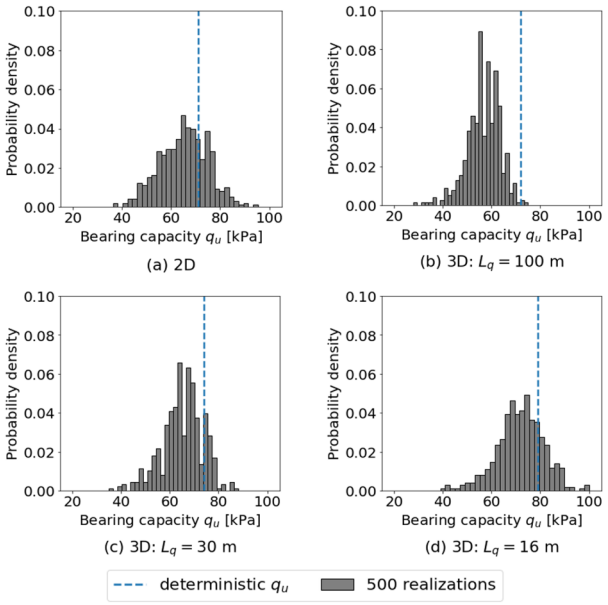


Figure 3. Histograms of q_u for 2D and 3D cases with θ_h=16 m.

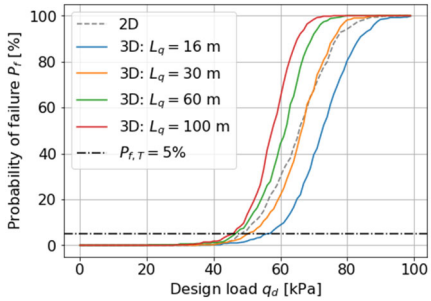


Figure 4. Plots of P_f vs. q_d for 2D and 3D cases with θ_h=16 m.

Drawing a horizontal line representing the value of $P_{f,T}$, the intersections of it with the curves in Figure 4 are the reliability-equivalent "design points" that correspond to different q_d values across the 2D and 3D cases, which are used to calculate the load factor η_L according to Equation (2) as a function of L_q (as θ_h is fixed). Figure 5 shows the increasing trend of η_L with larger L_q, for three $P_{f,T}$ values. It is observed that $\eta_L \leq 1$ for $L_q \leq 30$ m, implying an overall higher load-bearing capacity owing to the beneficial 3D effects when L_q is small relative to the embankment length as well as θ_h. In contrast, when L_q is larger, the relatively small θ_h makes it easier for the region under the influence of the load to encompass weak zones that promote semi-continuous or discrete failures, leading to a lower q_u than under plane strain conditions. Besides, it seems that the critical L_q at which the 3D load-bearing capacity ceases to be greater than for 2D (i.e., where η_L = 1) is smaller under a higher reliability requirement (i.e., lower $P_{f,T}$). Thus, it is important to identify the longest maximum load that the embankment is designed to sustain in its service life, and the corresponding load (amplification) factor needs to be applied if a 2D design approach is used.

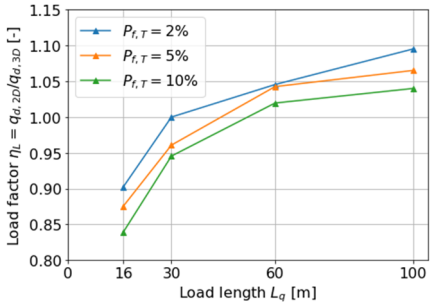


Figure 5. Calibrated η_L vs. L_q against $P_{f,T}$ for 3D cases with θ_h=16 m.

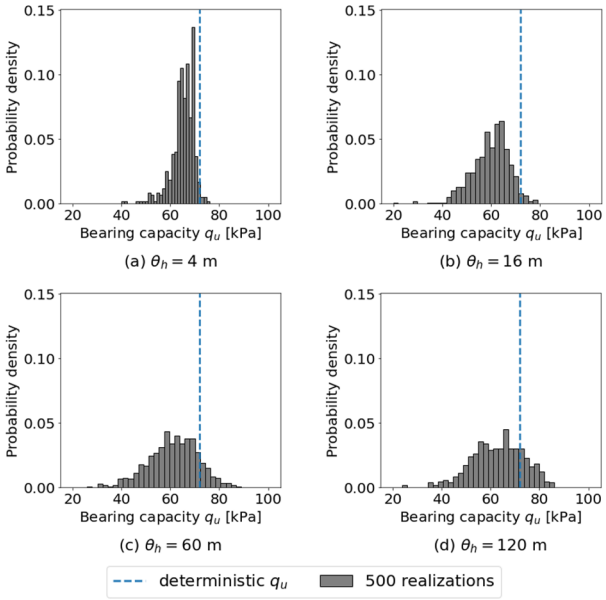


Figure 6. Histograms of q_u for 3D cases with L_q=60 m.

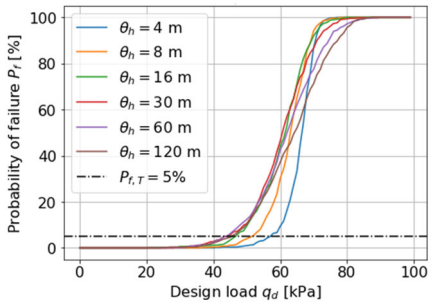


Figure 7. Plots of P_f vs. q_d for 3D cases with L_q=60 m.

The same procedure is also applied to cases with the same L_q value (e.g., 60 m) to study the influence of varying θ_h. The

histograms of q_u displayed in Figure 6 again confirm that a wider range of solutions is possible with a larger θ_h. The corresponding cumulative probability curves are shown in Figure 7, and the calibrated load factor against $P_{f,T}$ as a function of θ_h in Figure 8. It is observed that, for a fixed L_q, η_L increases with θ_h approximately to a point where $\theta_h \approx L_q$. The reason for the following slight drop with even larger θ_h is not clear and requires further investigation. However, $\eta_L < 1$ for small θ_h is attributed to the spatial averaging effect in the third dimension in the vicinity of the load, while larger θ_h favors 3D failure and lowers the overall bearing capacity as it becomes more likely to encounter semi-continuous weak zones in the soil.

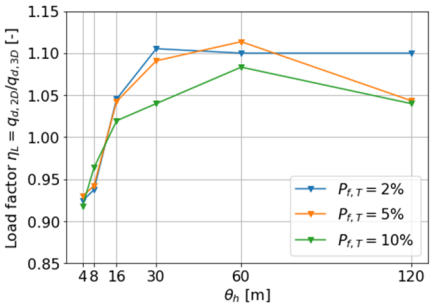


Figure 8. Calibrated η_L vs. θ_h against $P_{f,T}$ for 3D cases with L_q=60 m.

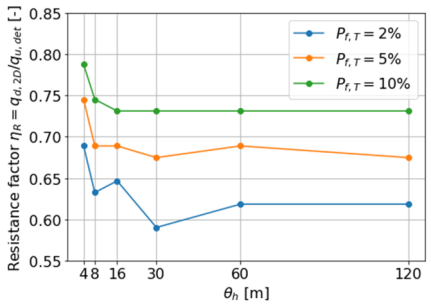


Figure 9. Calibrated η_R vs. θ_h against $P_{f,T}$ for 2D cases.

Aside from transforming the design load from 3D to 2D, the uncertainty in the resistance in a 2D deterministic design also needs to be handled. Using the 2D RFEM results for different θ_h values, the reliability-based resistance correction factor η_R is calculated according to Equation (3) and plotted in Figure 9 as a function of θ_h. The curves corresponding to different $P_{f,T}$ values consistently drop and finally flatten with larger θ_h, while $\eta_R < 1$ is always the case. Therefore, to be on the safe side, a very large θ_h value, approximately the same as the maximum load length, can be assumed for the calibration of η_R as well as η_L, if no accurate estimate is available.

4 CONCLUSIONS

Surcharge-loaded slopes, such as railway embankments, are subject to uncertainties in a 3D domain, which are usually overlooked in conventional design practice involving only cross-sections and 2D strip loads. In this paper, a series of full probabilistic analyses conducted using RFEM has revealed the impact of 3D soil heterogeneity in conjunction with the out-of-plane load length on the undrained load-bearing capacity of an idealized embankment. Making use of the RFEM results, a practical approach has been proposed for the calibration of 2D load and resistance based on "reliability equivalence". Major findings include:

1. The 3D effect of a limited load length L_q and that of the horizontal scale of fluctuation θ_h cannot be isolated, as a larger L_q diminishes the benefits of the end constraints while a larger θ_h leads to a wider spread of solutions, making weak responses more probable. As such, a "worst-case" scenario in 3D for a given (2D) design load can be the maximum L_q with $\theta_h \approx L_q$.
2. The load factor η_L, defined as the ratio of 2D design load to 3D design load for the same target reliability, can be used to calibrate the strip load applied in a plane strain model while reflecting the 3D nature mentioned in point 1. As η_L can be a reduction or amplification factor depending on the combination of L_q and θ_h, the 3D "worst-case" scenario can be taken as a conservative baseline for the load calibration.
3. The resistance factor η_R is a complementary measure that corrects the 2D deterministic bearing capacity based on the same target reliability against which the 2D design load has been calibrated. It is found that η_R is in general a reduction factor, especially when θ_h is large.

To conclude, a more comprehensive parametric study involving a diverse range of embankment geometries, load characteristics, and soil properties will help to generalize the findings and contribute to future guidelines for the reliability-based design or assessment of similar geotechnical structures.

5 ACKNOWLEDGEMENTS

This work is part of the research program RESET ("Reliable Embankments for the Safe Expansion of rail Traffic") funded by ProRail, and has used the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-13223. The authors also acknowledge the use of computational resources of the DelftBlue supercomputer, provided by Delft High Performance Computing Centre (https://www.tudelft.nl/dhpc).

6 REFERENCES

CEN, 2021. *NEN-EN 15528 Railway applications – Line categories for managing the interface between load limits of vehicles and infrastructure*. Delft: NEN.

Fenton, G.A., and Griffiths, D.V. 2008. *Risk Assessment in Geotechnical Engineering*. Hoboken: John Wiley & Sons, Inc.

Fenton, G.A., and Vanmarcke, E.H. 1990. Simulation of random fields via local average subdivision. *Journal of Engineering Mechanics*, 116(8), pp.1733–1749.

Hicks, M.A., and Li, Y. 2018. Influence of length effect on embankment slope reliability in 3D. *International Journal for Numerical and Analytical Methods in Geomechanics*, 42(7), pp.891–915.

Hicks, M.A., Nuttall, J.D., and Chen, J. 2014. Influence of heterogeneity on 3D slope reliability and failure consequence. *Computers and Geotechnics*, 61, pp.198–208.

Hicks, M.A., and Samy, K. 2002. Influence of heterogeneity on undrained clay slope stability. *Quarterly Journal of Engineering Geology and Hydrogeology*, 35(1), pp.41–49.

Hicks, M.A., and Spencer, W.A. 2010. Influence of heterogeneity on the reliability and failure of a long 3D slope. *Computers and Geotechnics*, 37(7), pp.948–955.

Phoon, K.-K., 2023. What geotechnical engineers want to know about reliability. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 9(2), pp.03123001-1–03123001-29.

Salgado, R., Woo, S.I., and Kim, D. 2011. *Development of Load and Resistance Factor Design for Ultimate and Serviceability Limit States of Transportation Structure Foundations*. West Lafayette: Joint Transportation Research Program, Indiana Department of Transportation and Purdue University.

Varkey, D., Hicks, M.A., and Vardon, P.J. 2023. Effect of uncertainties in geometry, inter-layer boundary and shear strength properties on the probabilistic stability of a 3D embankment slope. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 17(2), pp.262–276.