TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Kool, J. J., Kanning, W., Jommi, C., & Jonkman, S. N. (2020). A Bayesian hindcasting method of levee failures applied to the Breitenhagen slope failure. *Georisk*, *15*(4), 299-316. https://doi.org/10.1080/17499518.2020.1815213

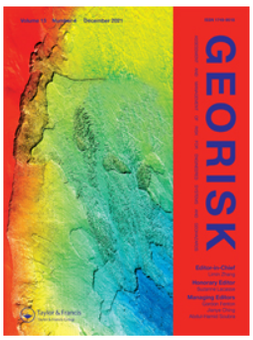
Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

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.



Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ngrk20

A Bayesian hindcasting method of levee failures applied to the Breitenhagen slope failure

J. J. Kool, W. Kanning, C. Jommi & S. N. Jonkman

To cite this article: J. J. Kool, W. Kanning, C. Jommi & S. N. Jonkman (2021) A Bayesian hindcasting method of levee failures applied to the Breitenhagen slope failure, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 15:4, 299-316, DOI: 10.1080/17499518.2020.1815213

To link to this article: https://doi.org/10.1080/17499518.2020.1815213

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 03 Sep 2020.

Submit your article to this journal

Article views: 726

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ngrk20

GEORISK
2021, VOL. 15, NO. 4, 299–316
https://doi.org/10.1080/17499518.2020.1815213

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

A Bayesian hindcasting method of levee failures applied to the Breitenhagen slope failure

J. J. Kool[a], W. Kanning[a,b], C. Jommi[a,c] and S. N. Jonkman[a]

[a]Faculty of Civil Engineering and Geoscience, Delft University of Technology, Delft, The Netherlands; [b]Department of Dike Safety, Deltares, Delft, The Netherlands; [c]Soil Mechanics and Geotechnical Engineering, Politecnico di Milano, Milan, Italy

ABSTRACT

Hindcasting of past levee failures enhances insights in the performance and vulnerability of levees. The scarcity of field evidence makes identifying the cause(s) of failure difficult. Under these circumstances, multiple scenarios and model choices are possible to characterise and to model the failure.

This paper shows how probabilistic Bayesian techniques advance the procedure of hindcasting of levee failures. In the developed approach, a-priori levee information, and failure observations are systematically taken into account to determine the most likely scenario and the most representative model choices to characterise the failure most accurately. Observations, such as the slip surface, are taken into account in the probability estimates.

The levee failure near Breitenhagen, Germany (2013) is used as a case study. The levee failed during river floods due the instability of the landside slope. The levee failure was most likely triggered by locally weak soil conditions and unexpected high water pressures due a connection between a pond on the riverside of the levee and the aquifer. These conditions were likely caused by the occurrence of a previous breach at this location.

The approach developed in this paper is expected to support a more systematic and objective method of analysis of other levee failures.

ARTICLE HISTORY
Received 6 December 2019
Accepted 11 August 2020

KEYWORDS
Hindcasting; forensic analysis; levee failure; flood defences; slope stability; Bayesian updating

1. Introduction

Consequences of flood defence failures are very large and it is therefore of great importance to gain insights in the performance and vulnerability of these systems flood defence (Vorogushyn, Merz, and Apel 2009; Jonkman and Schweckendiek 2015). Hindcasting of failed levees provides valuable insight into general levee performance, quality of strength models and dominant factors that contribute to levee failure. Examples of the hindcasting of past failures include analyses of levee failures in New Orleans, USA (Seed et al. 2006), Breitenhagen, Germany (Grubert 2013) and Yabe-river, Japan (Honjo et al. 2015). These insights can be used to advance the design and engineering of flood defences. As part of these hindcasts, various model choices have to be made to calculatively characterise the failure as accurately as possible. However, these model choices are often uncertain. Furthermore, field evidence is scarce: subsoil information is often "washed away" and there are often no direct visual observations during the failure. This makes identification of the cause(s) of failure difficult.

Forensic analysis provides a systematic procedure of analysis for the investigation of failures. The procedure of forensic analysis roughly consists of three stages: (1) collecting and reviewing of evidence, (2) utilising calculative models for a back analysis (hindcasting) to identify the cause, and (3) reporting the findings (Carper 2000). In previous work, it appeared that hindcasting using a deterministic sensitivity analysis identifies several possible causes of failure (Zhang, Tang, and Zhang 2010; Kool et al. 2019). However, the deterministic sensitivity approach of hindcasting did not provide an explicit insight into the likelihood of various failure scenarios.

In order to enhance levee hindcasting, probabilistic Bayesian techniques are used in this paper to identify the most likely scenario of failure, the most representative model choices to characterise the failure, and the most dominant parameters triggering the failure. The prior and posterior probabilities reflect the probability that a model best describes the reality and that a scenario of failure was present given the evidence. The combination of the most likely scenario of failure and most representative model choices is referred to as "MLC" in

CONTACT J. J. Kool j.j.kool@tudelft.nl

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the remainder of this paper. In the past, Bayesian techniques have been used to characterise the lifetime reliability of geotechnical structures (Zhang, Tang, and Zhang 2010; Schweckendiek 2014b; Baecher 2017), and have been applied to the hindcasting of different types of geotechnical related structures (Luckman, Der Kiureghian, and Sitar 1987; Gilbert, Wright, and Liedtke 1998; Zhang, Tang, and Zhang 2010; Gilbert 2016). However, to our knowledge, Bayesian techniques have not been applied to hindcasting of specific levee failures. Another novel element in the proposed approach in this paper is that Bayesian updating is carried out using specific information on the geometry of the observed failure, in this case, the slip surface. This approach allows us to account for the scarcity of evidence, as the available information after failures is generally limited to a small number of observations, photos and videos.

The article is structured as follows: The background of the method for the hindcasting of levee failures using Bayesian techniques is presented in section 2. In section 3 the case study of the levee failure near Breitenhagen in the year 2013 is analysed using the Bayesian hindcasting method. The results are discussed in section 4, and the conclusion and recommendations in section 5.

2. Probabilistic hindcasting of slope instability using Bayesian updating

2.1. Method background for Bayesian levee hindcasting

The proposed approach for hindcasting levee failures consists of defining scenarios that could explain the failure and assign prior probabilities of occurrence to these scenarios. Subsequently, the probability of failure given a scenario is calculated (which is used as likelihood) and Bayesian updating is used to update the prior probabilities with the likelihood and failure observation (see e.g. [Gilbert, Wright, and Liedtke 1998; Schweckendiek 2014b]). The objective is to find the most likely scenario (S_i) that has resulted in the failure event (F).

2.1.1. Prior probabilities

The first step of the proposed hindcast approach is to use the collected information for the identification of possible scenarios (S_i) that could have resulted in failure of the levee. A scenario typically represents the loading conditions possibly triggering the failure as well as the subsoil conditions (e.g. pore water pressures, stratigraphy). Prior scenario probabilities (i.e. $P(S_i)$) are assigned in such a way that they total a sum of 1 (e.g. Schweckendiek et al. 2017).

2.1.2. Likelihood function based on failure and observed evidence

The Bayes' theorem makes it possible to incorporate evidence in the analyses, such as observed failure information (Schweckendiek 2014a). To apply Bayesian updating of the previously assigned prior scenario probabilities, two pieces of information are incorporated using a likelihood function in the next steps, i.e. (A) the actual occurrence of failure and (B) the geometry of the observed slip surface (when available).

Hence, likelihood function in step (A, referred to as likelihood A) is calculated by determining the probability of failure (F) per scenario $P(F|S_i)$ according to:

$$P(F|S_i) = P(Z(X)\langle 0|S_i) \tag{1}$$

In this equation, the limit state function Z is used to describe when failure occurs. This is evaluated in this paper by using a slope stability model. The joint probability density function f(x) of the random variables (X) is used to describe Z. Negative values of the limit state function (Z) describe which combinations of random variables (X) combination results in failure of a scenario (S_i). Hence, the likelihood function A shows how likely it was that, according to a model, failure would have occurred given a certain scenario. And thus how likely it was that failure was observed given a certain scenario. This likelihood is expressed as a probability of failure given the scenario. This probability of failure is the results of all the uncertain input parameters (X) in the model.

The First Order Reliability Method (FORM) is used in this paper to estimate the probability of failure, Equation (1). Subsequently, Equation (2) is used to find the design point at Z(X) = 0 which is the combination of parameter values that provide the highest probability density (see Rackwitz 2001). Hence, this is the most likely combination of parameters values triggering a failure. When this design point is found, the probability of failure is calculated, using $P_f = 1-\phi(\beta)$.

The design point is found using a process of iterations in evaluating:

$$Z = \beta - \sum_{m=1}^{n} \alpha_m u_m \tag{2}$$

Where β is the reliability index and α_m is the influence coefficient, or the FORM sensitivity coefficient (for which $\Sigma {\alpha_m}^2 = 1$) of variable X_m; u_m is the standard normally distributed variable representing a normalised stochastic variable. The value of the influence coefficient of a variable provides a measure of its contribution to the reliability. The highest value of the influence coefficients (max α_m) identifies the dominant basic variables. The

process of iterations in the used FORM analysis using the Probabilistic Toolkit is described in (Brinkman 2015).

Additionally, for slope instability analyses, the observed shape of the slip surface of slope failure (*h'*) can be combined with the likelihood function (A) and, thereby as a next step, can be used for updating the scenario probabilities. This combination of likelihoods is further referred to as the likelihood (B). Evidence is expressed by an observation function *h'* and limit state function (Z), as in Equation (3). Negative values of the *h'* conditioned limit state function (Z) describe which random variable (X) combination result in both potential failure of a scenario (S_i) and observed slip surface shape (*h'*). The following likelihood function combines the influence of failure and slip surface (*h'*) (into likelihood function (B)):

$$P(F \cap h'|S_i) = P(Z(X)\langle 0|S_i) \tag{3}$$

In which Z(X) is the limit state function that includes the observed slip surface in the slope stability computation.

2.1.3. Posterior distribution of scenarios probabilities with observed failure

Subsequently, the earlier assigned prior probabilities of scenario S_i are updated using both the information that the slope has failed (A) and the geometry of the slip surface that is observed (B). This posterior distribution is shown in Equation (4):

$$P(S_i|F \cap h') = \frac{P(F \cap h'|S_i)P(S_i)}{P(F \cap h')} \tag{4}$$

2.2. Bayesian hindcasting of slope instability

2.2.1. Steps of Bayesian hindcasting of slope instability

A Bayesian method for the hindcasting of slope instability is proposed in six steps, based on the method by Kool et al. (2019) and described in Figure 1.

The presented Bayesian theory and the model for slope instability analysis (introduced in the following sections) are used to identify the combination of most likely scenario triggering the failure and the most representative model choice (MLC), step by step:

(1) System description and identification of possible failure scenarios (e.g. stratigraphy and water pressures) with the help of the collected evidence. Prior probabilities are assigned to all scenarios (S_i), a uniform prior can be chosen in case no prior knowledge is available to discern the scenarios:

$$P(S_i) \tag{5}$$

(2) Introduction of alternatives for model choices (M_j) and parameter choices (X) into the scenarios (S_i) and building slope stability models (noted as ($S_i \cap M_j$)). Prior probabilities are assigned to all model choices:

$$P(M_j|S_i) \tag{6}$$

(3) Evaluation of the limit state function using the generated slope stability models ($S_i \cap M_j$), to calculate the likelihood, one by one:

A. The conditional probability of event F (likelihood (A)):

$$P(F|S_i \cap M_j) \tag{7}$$

B. The conditional probability of F (likelihood (B)) including the field observation information when available (observed slip surface information (*h'*)):

$$P(F \cap h'|S_i \cap M_j) \tag{8}$$

(4) Update of prior scenario and model choices probabilities into posterior probabilities using likelihood (A) given failure (F) and (B) slip surface information (*h'*):

$$P(S_i|F \cap h') \tag{9}$$

$$P(M_j|F \cap h') \tag{10}$$

(5) Identification of the combination of most likely scenario and model choices: the scenario and model choices resulting in the highest posterior probabilities, given failure (F) and slip surface information (*h'*), are considered the most likely to characterise the failure most accurately (MLC):

$$\max P(S_i|F \cap h') \tag{11}$$

$$\max P(M_j|F \cap h') \tag{12}$$

(6) Identification of the most dominant parameter contributing to failure: the most dominant variables are found by assessing the influence factors of each basic variable for the combination of most likely scenario and most representative model choices (MLC):

$$\max \alpha_m^2 \tag{13}$$

Please, note that the prior probabilities refer to the probability that a model (M) best describes the reality, and that a scenario (S) was present given the information. In the next paragraphs step 1 to step 6 will be explained in more detail.

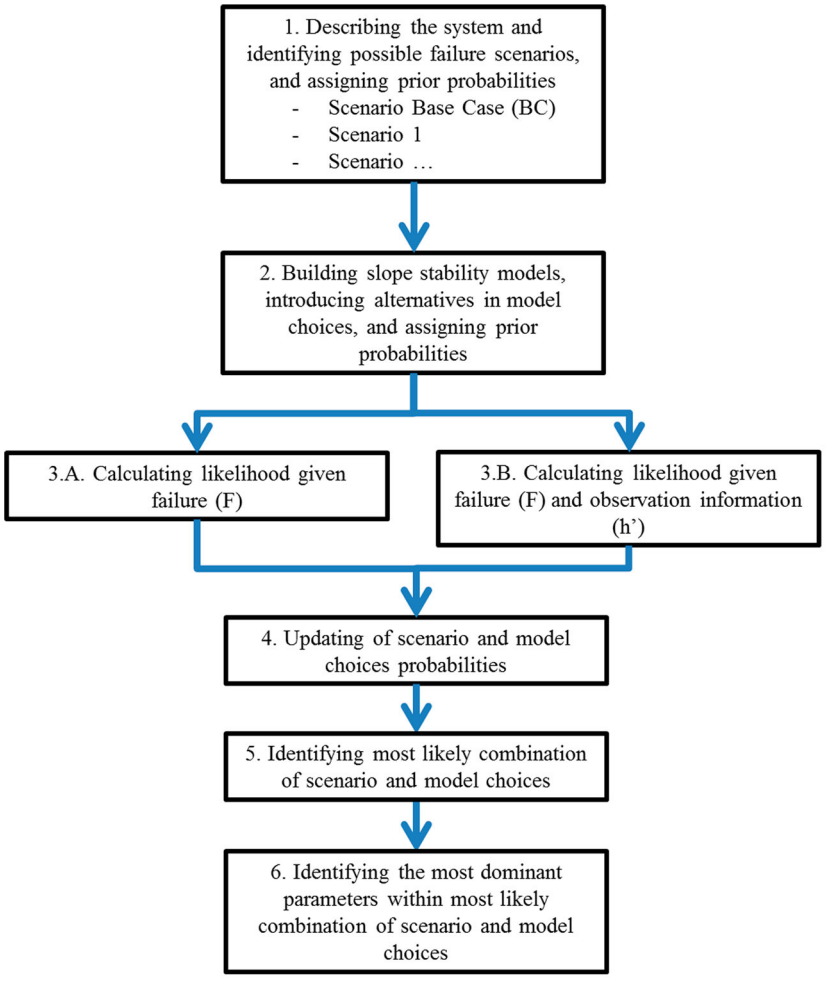


Figure 1. Method of hindcasting of levee failures using Bayesian techniques.

2.3. Elaboration of the hindcasting steps

2.3.1. System description and inputs (Step 1. and Step 2.)

A general model of slope stability analysis of a levee is used to analyse the failure of the levee. The analysis calculates the Factor of Safety (FoS) which is a function of the resisting moment (M_R) and the driving moment (M_S). Higher pore water pressures increase M_S and decrease the available shear strength of the soil and lead to a reduction of M_R. The limit state of the slope stability model is used with the purpose to determine the conditional failure of the levee and to incorporate the influence of the observation information (Step 3.A. and 3.B).

The performance of a levee depends on various typical variables (van Deen and van Duinen 2016) (Table 1). Collected evidence is thoroughly studied on the typical variables influencing the performance of the levee with the purpose of building a slope stability model. Further, the collected evidence is used to identify scenarios which might support conditions for triggering the levee to fail. Generally accepted literature is used to collect possible alternatives to typical variables supplementing the required set of variables in building the slope stability model.

Each discrete variable (Table 1) contributes a finite set of possible choices of model and parameter choices. All possible combinations of choices are used to generate slope stability models covering the possible characteristics of the failure.

A uniform prior can be chosen in case no prior knowledge is available to discern the scenarios of failure and alternatives of typical discrete variables. For example; the evidence is used to construct four scenarios possibly triggering the failure. Each of them is assigned a probability of 0.25 (Equation (5)).

2.3.2. Likelihood given observed failure (Step 3.)

The generated slope stability models (step 1 and step 2) are used in step 3A to determine the conditional probability of failure, which is taken as likelihood given failure (step 3A). The probability of slope instability is

Table 1. Overview of uncertainties and related typical variables of influence on slope instability performance. Variables are introduced as either a discrete or continuous variable. *chosen to exemplify the method on the Breitenhagen case, discussed in the next section (see Figure 6).

Uncertainties slope instability (Baecher and Christian 2003)	Variables (van Deen and van Duinen 2016)	Variable type	Specific variables/ choices related to Breitenhagen levee
Site characterisation	Geometry of the levee	Discrete	[–]*
	Stratigraphy of the soil layers and the levee body		[–]*
	Water level (landside and waterside)		[–]*
Model	Limit Equilibrium Model (LEM)		Bishop Spencer Uplift-Van
	Soil Shear Strength Model		Drained Undrained
Parameters	Soil parameters	Continuous	Cohesion Friction angle Shear strength ratio Strength increase exponent Pre-Overburden Pressure
	Hydraulic pressure of relevant layers		Simplified model (steady-state flow)

calculated using the limit state function Z(X), see Equation (14). This is repeated for each generated slope stability model and its corresponding random FoS(X); with each model based on a combination failure scenario, model and parameter choices:

$$Z(X) = FoS(X) - 1 \tag{14}$$

Equation (15) is used to calculate the conditional probability of failure (step 3A.).

$$P(F|S_i \cap M_j) = P(Z(X) < 0|S_i \cap M_j) \tag{15}$$

The collected evidence is subsequently used in step 3B. to estimate the geometry of the observed slip surface (Figure 2) or a zone in which the slip surface might have occurred. The slope stability models that produce a slip surface that correspond the observed information (shaded area in Figure 3 and denoted with *h'*) are taken into account of the analysis and expressed by Equation (16) (step 3B.).

$$P(F \cap h'|S_i \cap M_j) = P\big(Z(X)\big(0 \cap h'\big|\ S_i\ \cap\ M_j\big) \tag{16}$$

In practice, the observations of a slip surface are not exact. Therefore *h'* represents an estimated area with an upper boundary (S_{R1}) and lower boundary (S_{R2}), see Figure 2. The yellow area shows the envelope of all possible slip surfaces that is the result of step 3A. The shaded area shows the boundaries that are introduced to incorporate the observations, as done in step 3B.

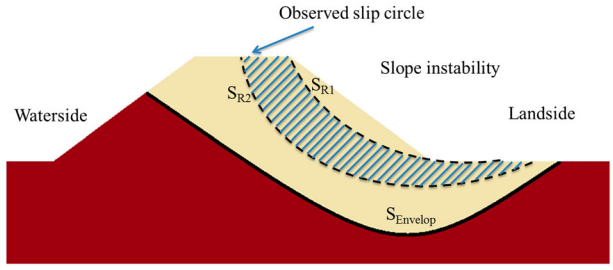


Figure 2. Yellow ($S_{Envelop}$): Yellow area marks the envelope where the critical slip surface can be found regardless whether this corresponds to the observed information. Shaded area marks the possible zone of the slip surface (*h'*).

2.3.3. Posterior probabilities identifying the most likely scenario and model choices (Step 4. and Step 5.)

Posterior probabilities are used to identify the MLC for hindcasting purposes. Equation (4) is calculated by applying a tree diagram in which all combinations of scenario and model choices are shown. For each combination of scenario and model choice, the likelihood is calculated in a slope stability computation by computing the probability that failure would have occurred. Combining these likelihoods with the prior probabilities of the scenarios finally gives the posterior distribution. For explanatory purposes, a simplified example is shown, in this section.

The simplified example takes only two discrete choices ($n = 2$) into account per discrete variable, that is two scenarios of failures ("S_1","S_2",) and two model choices ("M_1","M_2"). An overview of the resulting four possible slope stability models (noted as $(S_i \cap M_j)$) is shown both by an event tree (Figure 3) and an Edwards-Venn diagram (Figure 4). The slope stability models are used to determine $P(F|S_i \cap M_j)$. Prior probabilities are assumed $P(S_1) = P(S_2) = 0.5$ and $P(M_1) = P(M_2) = 0.5$; since $P(S_i)$ is independent of $P(M_j)$ result in $P(S_1 \cap M_1) = P(S_1)P(M_1) = 0.25$.

The four slope stability models $(S_i \cap M_j)$ are assumed mutually exclusive, and the total probability of failure is the union in event F (Figure 4). This total probability indicates whether the failure was likely (could be expected) or not (a surprise); the total probability of failure is calculated as:

$$\begin{aligned} P(F) = {} & P(F|M_1 \cap S_1)P(M_1|S_1)P(S_1) \\ & + P(F|M_2 \cap S_1)P(M_2|S_1)P(S_1) \\ & + P(F|M_1 \cap S_2)P(M_1|S_2)P(S_2) \\ & + P(F|M_2 \cap S_2)P(M_2|S_2)P(S_2) \end{aligned} \tag{17}$$

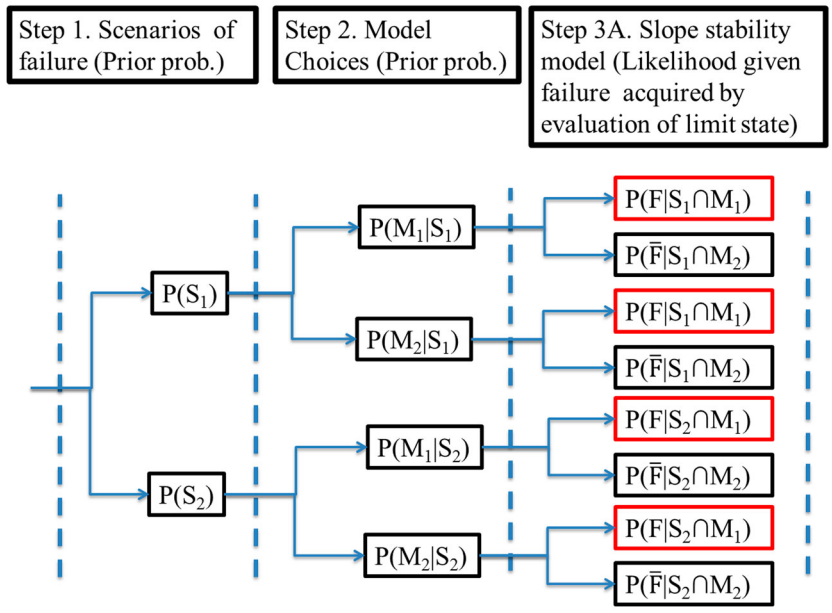


Figure 3. Event tree with the incorporation of the scenarios of failure ("S_1", "S_2"), alternatives in model choices ("M_1", "M_2"). The slope stability models are put to purpose to determine the related likelihood given failure ("$F|S_i \cap M_j$") and given non-failure ("$F|S_i \cap M_j$").

Bayes rule is used to update the prior model and scenario probabilities by the information that failure occurred (step 3A). When slip surface information is available, the set of likelihoods given failure and observation information (step 3B) replaces the set of likelihoods given failure (step 3A, Figure 3). The posterior probabilities of the scenarios of failure and model choices are obtained by alternately evaluating Equation (18) and Equation (19):

$$P(S_2|\ F) = \frac{\sum_{j=1}^{n} P(F|M_j \cap S_2)P(M_j|S_2)P(S_2)}{P(F)} \quad (18)$$

$$P(M_2|\ F) = \frac{\sum_{i=1}^{n} P(F|S_i \cap M_2)P(S_i|M_2)P(M_2)}{P(F)} \quad (19)$$

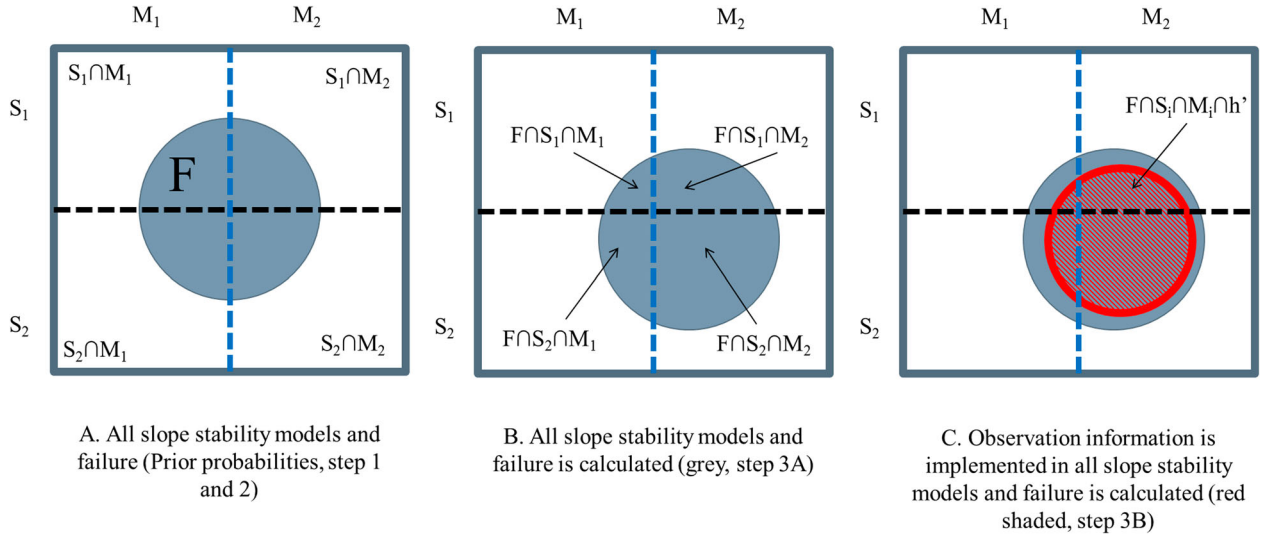


Figure 4. (A.) Edwards-Venn based diagram to illustrate the relation between the scenarios of failure (S_1, S_2), the model choices (M_1, M_2) and all the four resulting slope stability models that are built ($S_i \cap M_j$) in order to calculate the probability of failure in the next step (noted as F and marked as the grey area). Prior probabilities of scenarios and model choices are included. (B.) Probability of failure per slope stability model is calculated and implemented in the diagram ($P(F \cap S_i \cap M_j)$ marked as grey area). This illustrates intersection of the calculated F and the slope stability models. (C.) The red shaded area is a subset of F and incorporates the observed slip surface information (*h'*). The most likely scenario (S_2) and model choices (M_2) are identified.

This way, the most likely scenario leading to failure and model choice to represent the levee failure most accurate is obtained by evaluating Equation (11) and Equation (12).

2.3.4. Dominant influence factor (Step 6.)

The MLC is used to determine the most dominant variable leading to the failure of the levee. The influence factor is a by-product from the evaluation of the limit state function using FORM, as is done in step 3A and step 3B, and provides a measure of the contribution of each variable ($\Sigma_m \alpha^2 = 1$) to failure. The most dominant variable is identified utilising Equation (13).

3. Case study of levee failure near Breitenhagen, Germany in 2013

3.1. The Breitenhagen case: levee failure and input data

3.1.1. General failure information

In 2013, the levee near Breitenhagen (Germany) failed and caused considerable economic damages to the area, see e.g. Figure 5 (Weichel 2013). Previous analyses of the levee failure show that the levee would be calculatively stable (FoS > 1) under conditions of best estimates of levee characteristics (Grubert 2013; Kool et al. 2019). This indicates that possibly an anomaly (e.g. old breach or conductive layer) dominated the outcome (Grubert 2013; Kool et al. 2019). Details on the interpretation of the collected data and the identification of anomalies, which are introduced in the analyses as possible scenarios leading to failure, are discussed by Kool et al. (2019). The collected data are summarised in Table 2. The data that support the findings of this study are available in the International Levee Performance Database (ILPD) – a public database with information on levee failure and performance cases (at https://leveefailures.tudelft.nl/, failure-id 121001: Breitenhagen, 2013).

3.1.2. Step 1: System description, identification of possible failure scenarios

3.1.2.1. General. The collected evidence is examined to identify possible failure scenarios, i.e. loading and subsoil conditions that possibly explaining the failure. The Breitenhagen levee has a typical cross-section of a levee on top of an aquitard on both land- and riverside, consisting of layers of cohesive and alluvial soil, and a high permeable aquifer consisting of sand and gravel (Figure 6; based on [Grubert 2013]). At the location of the breach, the levee was about 3.50 m high and had a crest width of 3.00 m. A possible failure scenario is caused by the trees growing at the section between the levee and the pond near the toe of the levee at the waterside. Furthermore, this particular section of the levee has likely breached in the past, resulting in a pond in front of the levee (Sixdorf 2016).

3.1.2.2. Failure scenarios. Seepage of water from the riverside causes local high water pressures inside the levee to reduce the shear strength and possibly trigger slope instability. The best-estimated situation is referred to as the "Base case" (BC). Subsequently, three possible scenarios of water pressures are identified as shown in Figure 7. To emphasise the scenario-specific dominant water pressures, the influence of each scenario-specific water pressure is introduced in relative to the BC. Kool et al. (2019) further elaborates and motivates the different scenarios:

- Scenario "Base case" (BC): the best estimate of the base case is based on normal levee investigation data (borings, etc.) and best estimates of water pressures,
- Scenario "Saturated levee" (S1): higher phreatic as a result of more permeable levee than expected, based on evidence from photos,
- Scenario "Conductive layer" (S2): higher water pressures inside levee due to a conductive layer inside the levee as a result of tree roots, based on evidence on photos,
- Scenario "Pond connection with aquifer" (S3): higher water pressures in the aquifer under the levee due to a close connection of the outside water level with the aquifer, based on evidence on photos.

3.1.2.3. Water pressures in failure scenarios. This analysis adopts a pragmatic implementation of water pressures using simplified models of quasi-steady-state flow to incorporate the transient flow (Figure 7; [TAW 2004]). Three aspects are considered: A. the phreatic line (PL), B. the aquifer head (HH1) and C. the interpolation between A. and B. (this is done linearly and not further discussed). Whether steady-state conditions are reached depends on the duration of the water level and the hydraulic conductivity of the soil. In this case, the higher water event lasted over a month, of which 12 days the water was relatively close to the maximum (see Kool et al. 2019). Analysis by Drews (2015) shows that the aquifer conditions mostly reach steady conditions for the various scenarios, which is incorporated in the modelling in this study. For the phreatic line, this is highly dependent on the hydraulic conductivity of the levee material, see Drews (2015). For the reported



Figure 5. Left: overview photo of the inland while the flooding evolves, Right: overview photo of the breach at 10-jun-2013 (Grubert 2013; Weichel 2013).

low conductivity, steady-state conditions are not met. However, since photos show wet soils during, the slope failure; indicating a high phreatic line. Hence, a phreatic line that represents (close to) steady-state conditions is used in the base case.

Table 2. Collected data on the levee near Breitenhagen that failed. The data that support the findings of this study are available in the International Levee Performance Database (ILPD) (at https://leveefailures.tudelft.nl/, failure-id 121001: Breitenhagen, 2013).

Document	Information
Design of upgrade 1846 (Sixdorf 2016)	Design drawings of the cross-sections. The design was not realised.
Photo report 2003 (Sixdorf 2016)	Photos of the installation of sheet piles to prevent seepage near the pumping station
Photo report 2004 (Sixdorf 2016)	Photos of the construction of the road on top of the crest
Video footage, 2013 (Weichel 2013)	Video footage during the breach by a drone flight
Saaledeich bei Breitenhagen, geotechniscche untersuchungen der Bruchstelle Empfehlungen zut Sanierung, (Grubert 2013)	Photo report (during and after the breach) Hindcasting based on calculations Location overview Levee profile (measurements) Levee profile km 0 + 590 (incl. borings) Data based on laboratory test of local samples: water content w[%], yield point wL [%], Roll-out limit wP [%], Plasticity number IP [%], Consistency number IC [−] Moist sample [gr], dry sample [gr], pore water [gr], water content [%] Sieve curves Geotechnical stability calculations a variety of scenarios Underground hydraulic analysis (stationary and transient pressures) Uplift assessment
Photo report and paper clippings of the repair 2013 (Sixdorf 2016)	Photo footage of the repairing of the levee Clippings of the plans in the area Photo footage right after the repair

It should be noted that using steady-state conditions is usually pragmatic and conservative for design considerations; for hindcasting, this may lead to the identification of the wrong dominant scenario in case there are significant transient effects. A sensitivity analysis shows that using transient water pressures based on a low conductivity dike body does not influence the outcomes much in this case since the aquifer pressures are much more important, see "Discussion".

3.1.2.4. Modelling water pressures. The water pressures are modelled in each scenario using probability distributions to reflect uncertainty in actually occurred pressures. This is shown in Figure 7 and summarised in Table 3. The red points in Figure 7 are assigned the probability distribution, linear interpolation is used to connect this location with waterside and landside boundaries. The following choices are made:

- Base Case (BC): the mean of the PL (dark blue) and aquifer (light blue) are shown in Figure 7 and are based on quasi-steady-state conditions. For the PL, there is a small uncertainty modelled (0.3 m; see (Rozing 2015)). Also, a truncated distribution is used, similar to the dashed line of S1, to reflect the physical boundaries as the water level cannot exit the levee body. The PL can be higher than steady-state conditions though due to e.g. rainfall. For the aquifer, blanket theory (USACE 2000) is used to compute the water pressures (HH1). In this case, a semi-pervious top stratum to calculate the head is shown in Figure 6. The uncertainty is 0.3 m in both x-, and z-direction. Again a truncated distribution is used with upper bound equal to the maximum river level and lower bound equal to the pre-flood ground water table.
- Scenario "Saturated levee" (S1): The only difference with BC is the increased standard deviation of the

Table 3. Uncertainties in water pressures. The hydraulic head distribution is modelled in the red nodes (Figure 7) and modelled as a truncated normal distribution. The coordinates correspond with Z and X in Figure 7. *S1, S2, and S3 have the same distribution for PL and HH as the base case, except for what is shown in the "water pressure" column.

Scenario of failure	Line	Axis	μ [m]	σ [m]	Truncating Bounds [m] – [m]
BC: "Base Case"	Phreatic line in levee (PL)	Z	54	0.3	51–54.7
BC: "Base Case"	Phreatic line in levee (PL)	X	36.5	0.3	35–38
BC: "Base Case"	Hydraulic head in aquifer (HH1)	Z	52.8	0.3	49–54.7
S1 "Saturated levee"*	Phreatic line in levee (PL)	Z	54	1.1	51–54.7
S1 "Saturated levee"*	Phreatic line in levee (PL)	X	36.5	1.1	35–54.7
S2: "Conductive layer"	Extra hydraulic head rooted conductive layer (HH2)	Z	52.8	1.11	51.5–54.7
S3: "Pond connection with aquifer"*	Hydraulic head in aquifer (HH3)	Z	54	1.11	49–54.7

PL (of 1.1 m) to reflect the higher potential for full saturation.

- Scenario "Conductive layer" (S2): The only difference with BC is the introduction of an extra piezometric head (HH2) at NHN + 51.2 m with increased head. There is a linear interpolation between HH2 and both the PL and HH1 which have the same properties as the BC (but are not shown in the figure for clarity)
- Scenario "Pond connection with aquifer" (S3): The only difference with BC is the increased mean of the aquifer head (HH3) due to the pond connection, resulting in a lower L1 in the blanket equations, see Figure 6.

3.1.3. Step 2: Alternatives for model choices and parameter choices

The possible model choice and parameter choices are derived that, together with the failure scenarios of step 1, will characterise the levee failure most accurately. The different scenarios are implemented with alternatives of Limit Equilibrium Methods (LEM), soil reaction behaviour and properties of soil parameters, to generate a set of slope stability models that include all possible characterisations of the failure (Table 1).

The LEM of Bishop (M_{Bishop}), Uplift Van (M_{Uplift}) and Spencer ($M_{Spencer}$) are applied as possible analytical slope stability methods (Bishop 1955; Spencer 1967; CIRIA 2013). LEMs are chosen over more complicated methods such as Finite Elements because of their short running time and sufficient accuracy for the not-complex slip surface as observed in this failure. The different LEM mainly differ in the various shapes of slip surfaces that they take into account. Moreover, the LEM of Bishop is considered less accurate than the LEM of Uplift Van and Spencer but is incorporated as a control method. The uncertain typical variables

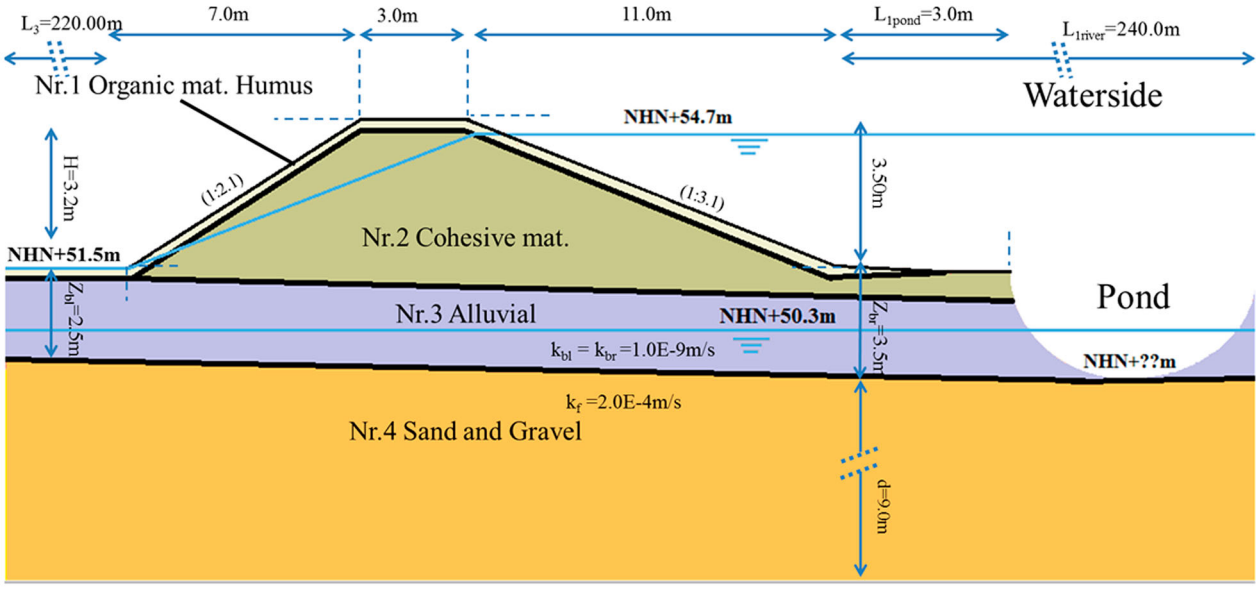


Figure 6. Simplified soil layering of the failed levee near Breitenhagen, including dimensions (Grubert 2013; Drews 2017). Coordinates are expressed in a vertical datum: Normalhöhennull (NHN), which is roughly sea level. The observed high water level NHN + 54.7 m, measured groundwater of NHN + 50.3 m. Aquitard extends up to L_3 = 220 m on the land side and up to L_{1river} = 240 m on the riverside. Furthermore, k_{bl}, k_{br} and k_f indicate the hydraulic conductivity of respectively the landside blanket, waterside blanket and aquifer; while Z_{bl} and Z_{br} indicate the landside and waterside blanket thickness.

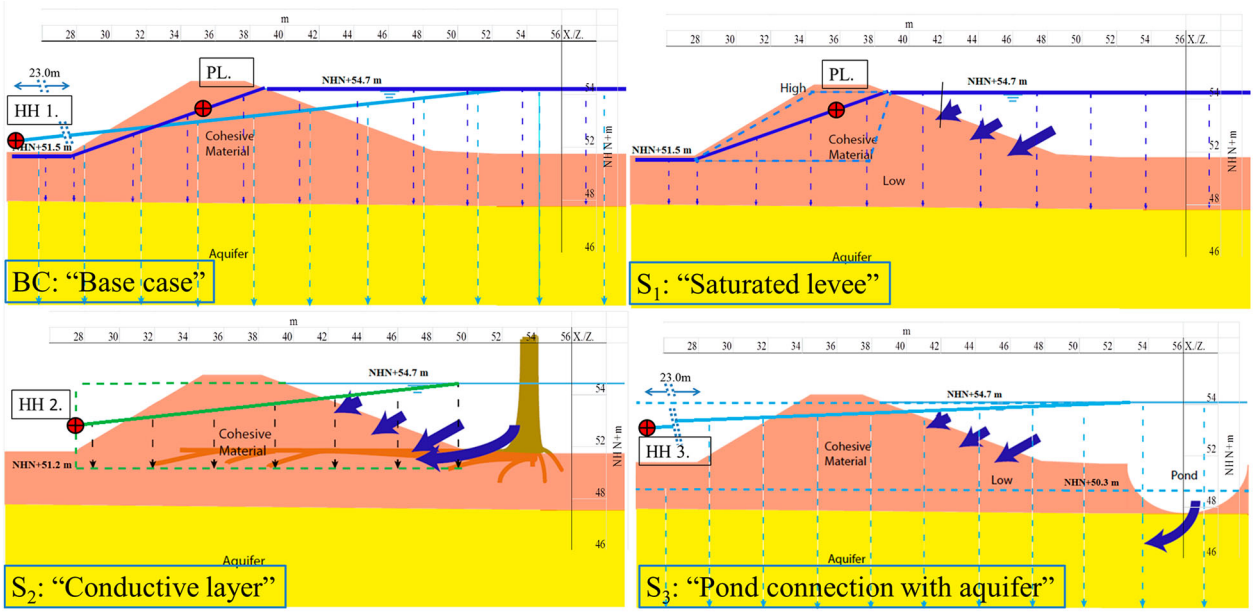


Figure 7. Modelling of the four scenarios of water pressure. Horizontal coordinates are project-specific, vertical coordinate relative to NHN. The dashed lines are (high and low) boundaries of the piezometric lines. The red node is the point of manipulation in the X-axis and Z-axis to incorporate the uncertainty in the phreatic line (Table 3).

(step 1) are the subject of this study as presented in Table 1. Therefore, the overall imperfection of the simulation is accepted and compensatory measures, such as an overall model uncertainty factor, are left out of the assessment.

How soils react to loading depends on the rate of the loading and the conductivity of the soil. Often, the shear strength of the soil is modelled as drained or undrained, while in reality the soil likely reacts as partially drained. Initially, drained (SB_{dr}) and undrained (SB_{und}) (Schofield and Wroth 1968; Ladd 1991) are used in this paper as these are most commonly used. Partially drained behaviours should be considered in case this is expected to have a large influence on the results; which is not expected for this case. Undrained soil behaviour takes the possibly generated water pressures by deformations into account, in contrast to drained soil behaviour (van Deen and van Duinen 2016).

The undrained soil response for low permeability materials is implemented using the SHANSEP implementation (Ladd 1991) of the Critical State Soil Mechanics (Schofield and Wroth 1968):

$$s_u = \sigma'_{v,i} S\, OCR^m \; with \; OCR = \frac{\sigma'_{v,y}}{\sigma'_{v,i}} \; and \; \sigma'_{v,y} = \sigma'_{v,i} + POP \tag{20}$$

Where s_u [kPa] is the undrained shear strength ratio, $\sigma'_{v,i}$ [kPa] is the in-situ vertical effective stress, S[−] the undrained shear strength ratio, OCR [−] the over-consolidation ratio, m[−] the strength increase exponent, σ'_{vy} [kPa] the vertical yield stress and POP [kPa] the pre-overburden pressure.

The report of Grubert (2013) describes the local soil types and characteristics but does not report the probability distributions of the strength parameters that represent the diversity of shear strength in space (TAW 2001; Calle 2008). Based on the descriptions of the soil characteristics by Grubert (2013) and Normcommissie (2011), the corresponding values of soil strength are collected for both drained and undrained soil behaviour (Tables 4 and 5). The probability distributions of the parameter input are adopted from Baker and Calle (2006).

Conform steps 1 and 2, failure scenarios are identified, a finite set of interchangeable alternatives of model choices and parameter choices are systematically collected, which results in 24 combinations of failure scenarios and model choices ($S_i \cap M_j \cap SB_n$). Each combination is used to build a slope stability model (see Figure 8), which is evaluated for the probability of failure given the a-priori conditions. When observed slip surface information available (see step 3 below). The observed slip surface information is included in all 24 slope stability models by conditioning the search area of the critical slip surface to the observed slip surface. Both the soil parameters and hydraulic pressures of relevant soil layers are implemented as continuous probabilistic distributions (Tables 3–5). An overview of all

Table 4. The probability distribution of soil properties (parameters of drained soil behaviour), soil weight (γ), wet soil weight (γ_{wet}), soil friction angle (φ), cohesion (c'). *Grubert (2013) used local soil investigations to identify the soil types, **Normcommissie 2011 provides typical mean values of strength parameters which are common in the Netherlands with similar soil description, ***Baker and Calle (2006) provide default typical values of variation coefficients with similar soil description. The probability distributions are log-normal with a mean (μ) and a standard deviation (σ) (Schweckendiek et al. 2017).

Soil type *	Description of a soil layer **	μ_γ [kN/m^3] **	σ_γ [kN/m^3] $v = 0.05$ ***	$\mu_{\gamma,wet}$ [kN/m^3] **	$\sigma_{\gamma,wet}$ [kN/m^3] $v = 0.05$ ***	μ_φ [deg]**	σ_φ [deg] $v =$ 0.2 ***	$\mu_{c'}$ [kPa] **	$\sigma_{c'}$ [kPa] $v = 0.5$ ***
Nr. 1. Organic mat Humus	Clay, clean and moderately stiff	17	0.85	17	0.85	21.25	4.25	7.5	3.75
Nr.2. Cohesive mat.	Clay, little bit of sand, stiff	18	0.90	18	0.90	25	5	7.5	3.75
Nr. 3. Alluvial	Clay, little bit of sand, stiff	18	0.90	18	0.90	25	5	7.5	3.75
Nr.4. Sand and gravel	Gravel, bit silty, clean	18,5	0.93	20.5	1.03	36.25	7.25	0	0

Table 5. Probability distributions of soil properties (undrained soil behaviour), shear strength ratio (S), strength increase exponent (m), Pre-overburden Pressure (POP). *Grubert (2013) used local soil investigations to identify the soil types, **Helpdeskwater 2016 provides typical values of strength parameters which are common in the Netherlands with similar soil description. The probability distributions are log-normal with a mean (μ) and a standard deviation (σ) (Schweckendiek et al. 2017).

Soil type *	Description of soil layer **	μ_S [−] **	σ_S [−] **	μ_m [−] **	σ_m [−] **	μ_{POP} [kPa] **	σ_{POP} [kPa] **
Nr. 1. Organic mat Humus	Sandy and Silty Clay	0.30	0.03	0.9	0.03	22	6.6
Nr.2. Cohesive mat.	Levee material	0.31	0.06	0.9	0.03	30	9
Nr. 3. Alluvial	Levee material	0.31	0.06	0.9	0.03	30	9

combinations is presented using an event tree in Figure 8 Appendix A, which is enclosed in the supplemental material of this paper, presents an elaborate overview of all combinations and calculations results, and a calculation example that demonstrates how the posterior probability of S_3 is calculated.

3.1.4. Step 3: Likelihood given failure and field observational information

The assembled 24 slope stability models are used to evaluate the accompanied limit state functions with FORM. This results in the conditional probability of failure of each slope stability model, which is adopted as likelihood in this analysis. The likelihood is determined for two situations:

A. Field observation information is not available and the search area of the critical slip surface is not defined in the slope stability models, resulting in $P(F|S_i \cap M_j \cap SB_m)$ (step 3A.),
B. Field observation information is available and the search area of critical slip surface is conditioned to the observed slip surface information in the slope stability models, resulting in $P(F \cap h' | S_i \cap M_j \cap SB_m)$ (step 3B.).

Figure 9 shows the field observation information that is available. The slope stability models are conditioned (h') according to the observed shape of the initial slip surface as part of step 3B. (Equation (8)). The exact geometry of the slip surface is difficult to estimate, and therefore estimated by an upper and lower boundary that are 2 m apart, as suggested in Figure 2.

3.1.5. Step 4, 5, and 6: Systematic analysis of uncertainties

Both sets of the 24 likelihoods (step 3A. and 3B.) and the prior probabilities (step 1 and 2) are implemented alternately in Bayes rule and used to determine the posterior probabilities for each scenario of failure and model choice (step 4). The resulting posterior probabilities substantiate the identification of the MLC (step 5) which is used to identify the parameters that contributed to the failure the most.

Evaluating the limit state with FORM enables to estimate the conditional probability of failure, but also determines the contribution of the individual parameters to failure ($\Sigma \alpha_i^2 = 1$). Because FORM might introduce approximation errors on the probability of failure, the results have been verified with Monte Carlo which gave very similar results for the considered cases.

3.2. Results

3.2.1. Updated scenario probabilities

The two sets of the 24 likelihoods (step 3A. and 3B.) are used to update the prior probabilities of earlier identified

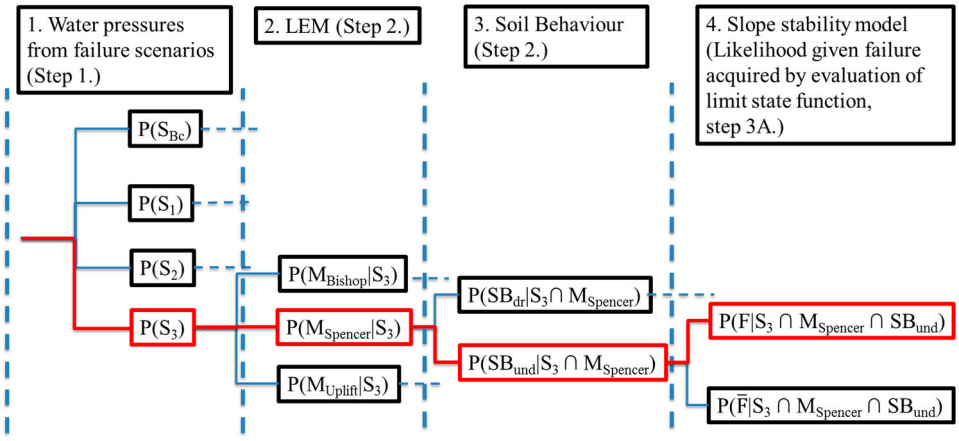


Figure 8. Presents an overview of the different combinations of scenarios and model choices. Level 1: Water pressure from failure scenarios (i.e. S_{Bc}:"Base case", S_1:"Saturated levee", S_2:"Conductive layer", and S_3:"Pond connection with aquifer"). Level 2: LEM (i.e. M_{Bishop}, $M_{Spencer}$ and M_{Uplift}). Level 3: Soil reaction behaviour (i.e. SB_{dr}: drained, and SB_{und}: undrained). Level 4: Likelihood given failure (F) and non-failure (F). Red branch contains MLC, as an example.

scenarios that possibly trigger a failure (step 1). The prior and the resulting posterior probabilities are shown in Figure 10, step by step.

Updating the prior probabilities of the scenario BC, S_1, S_2, and S_3, based on the information of the actual failure (step 3A) emphasises S_3 as the most likely scenario ($P(S_3) = 0.25$ to $P(S_3|F) = 0.72$). When the observed slip surface information is included in the update (step 3B), the posterior probability of S_3 reduces slightly (from $P(S_3|F) = 0.72$ to $P(S_3|h' \cap F) = 0.64$). The results identify S_3 as the most likely scenario of failure.

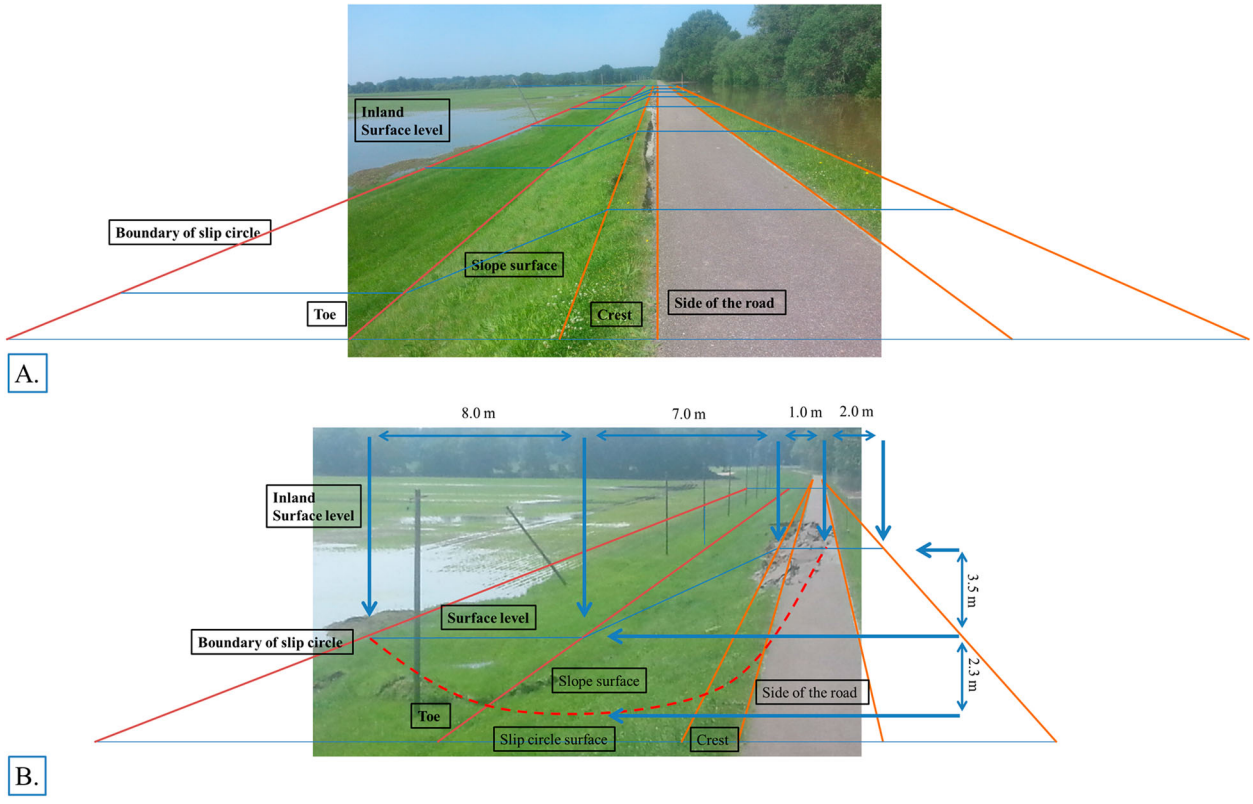


Figure 9. Geometric analysis of the failed levee. Longitudinal guides to the levee (in Orange, Red), guides across to the levee (Blue). (A.) First distortions are apparent, (B.) Larger distortions are apparent and used to estimate the geometry of the slip surface (Red dotted).

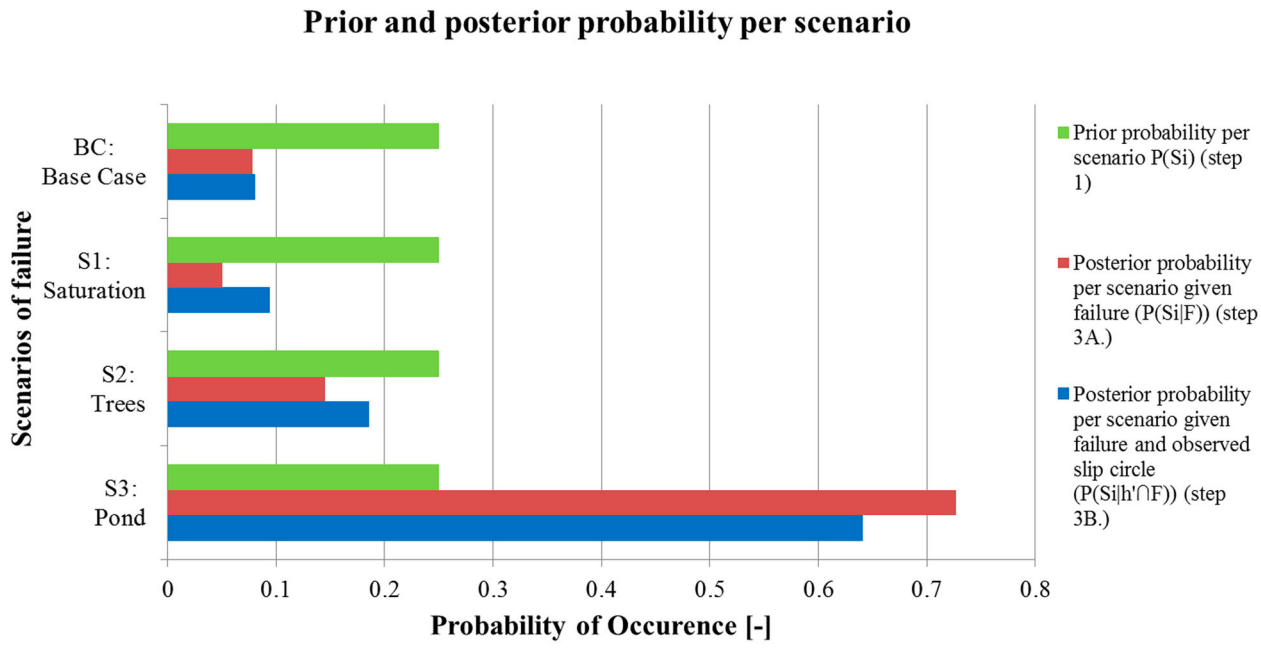


Figure 10. Prior probability per scenario BC, S_1, S_2, and S_3 (equally distributed probabilities (green)), the posterior probabilities per scenario given failure (red) and observed geometry of slip surface (blue).

3.2.2. Systematic analysis of uncertainties and most dominant parameters

Additionally, the set of the 24 likelihoods (step 3A. and 3B.) is used to update the prior probabilities of alternatives in the model choices (step 2), in the process of identifying the MLC. Both the prior and the resulting posterior probabilities are shown in Figures 11 and 12. Then, the MLC is deployed to determine the most dominant contributor to the failure (step 6).

Figure 11 shows that updating of the prior probabilities on the information that the actual failure happened (step 3A), emphasises the LEM of Uplift Van as the most likely choice of LEM ($P(M_{Uplift}) = 0.33$ to $P(M_{Uplift}|F) = 0.43$). However, when the observed slip surface information is included in the update (step 3B), the resulting

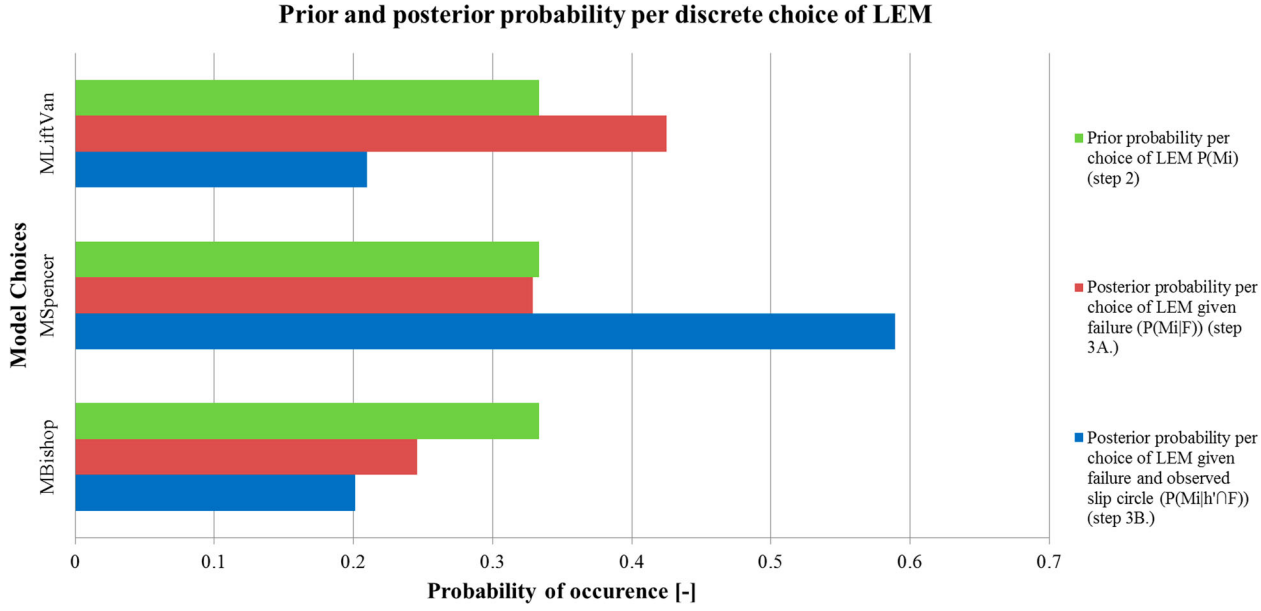


Figure 11. Prior probability per choice of LEM Uplift Van, Spencer, and Bishop and equally distributed probabilities (green), the posterior probabilities per choice of LEM given failure (red) and observed geometry of slip surface (blue).

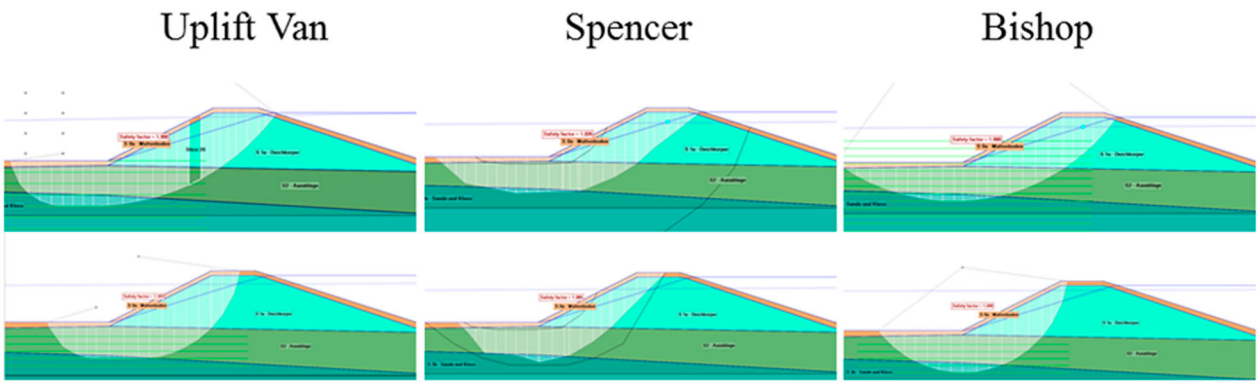


Figure 12. Slip circle in the design point of the most likely scenario (S3: Pond) and soil behaviour (undrained soil behaviour) LEM Uplift Van (left), Spencer (middle), and Bishop (right). Upper figures: search area of the critical slip surface is not defined in the slope stability models (step 3A). Lower figures: search area of critical slip surface is conditioned to the observed slip surface information in the slope stability models (Step 3B).

posterior probability emphasises the LEM of Spencer as a more likely choice than the LEM of Uplift Van, i.e. $P(M_{Spencer}|F) = 0.33$ to $P(M_{Spencer}|h' \cap F) = 0.59$. The posterior probabilities identify the LEM of Spencer to characterise the failure most likely most accurately. Figure 12 shows the most likely resulting slip circle (in the design point) when S3: Pond and undrained soil behaviour is assessed using LEM of Uplift Van, Spencer and Bishop without observational information (step 3A, upper row) and with observational information (step 3B, lower row) included.

Figure 13 shows that updating prior probabilities of soil response behaviour on the information that the levee failed (step 3A.) emphasises the posterior probabilities of the undrained soil response behaviour as the most likely soil behaviour, i.e. $P(SB_{undr}) = 0.50$ to $P(SB_{undr}|F) = 0.65$. Including the observed slip surface information in the update (step 3B.) has little effect on the posterior probabilities. The posterior probabilities identify the undrained soil response behaviour to characterise the failure most likely most accurately; however, the differences are small and no firm conclusion can be drawn.

The slope stability model implemented with the combination of scenario pond connection with the aquifer (S_3), method of Spencer ($M_{Spencer}$) and undrained response soil behaviour (S_{udr}) characterises the failure most likely most accurately. This slope stability model is used to evaluate the limit state and to estimate the contribution of the individual variables to the failure. Figure 14 shows that overall the shear strength ratio and the hydraulic head in the aquifer are considered the most dominant parameters contributing to the failure. The influence of the Pre-overburden pressure seems to decrease as a result of the incorporation of the observed slip surface information.

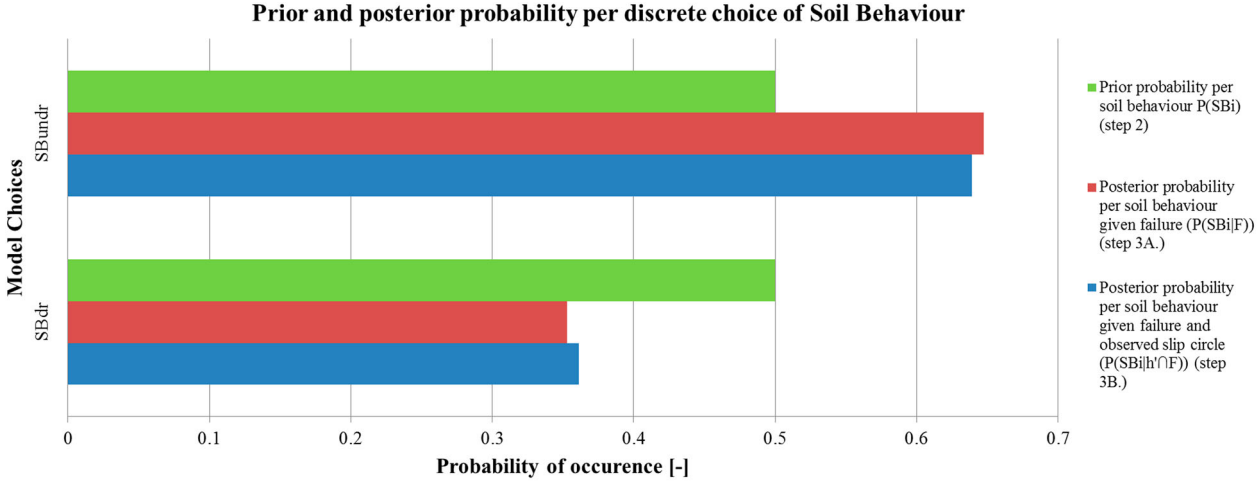


Figure 13. Prior probability per choice of soil behaviour both drained (SB_{dr}) and undrained (SB_{undr}) and equally distributed probabilities (green), the posterior probabilities per choice of soil behaviour given failure (red) and observed geometry of slip surface (blue).

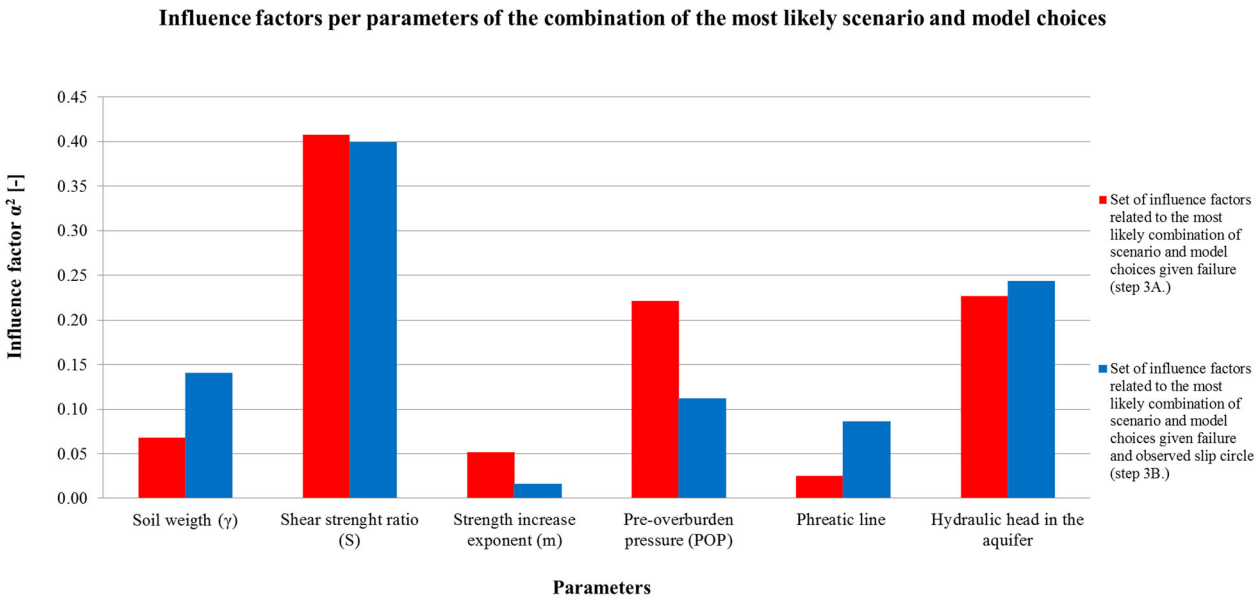


Figure 14. Set of influence factors as a result of utilising the MLC. Comparing the influence factors of the individual parameters without (red) or with (blue) incorporation of observational information ($\Sigma\alpha^2 = 1$).

4. Discussion

4.1. Framework for probabilistic hindcasting of levee failures

The six steps of hindcasting are based on a framework which explicitly and transparently accounts for significant uncertainties and model choices by quantitative means that reflect on the likelihood of occurrence. Although, consideration of all possible combinations of scenarios of failure and model choices is time-consuming; it does provide a thorough analysis of the failure possibilities.

Even more, the suggested Bayes technique based approach enables the implementation of observational information in the hindcasting. The approach provides insights on how each piece of evidence influences the scenario and model choice related posterior probabilities, building up to the identification of the slope stability model best representing the failure, piece by piece. The proposed slope stability model facilitates the identification of the most dominant parameter contributing to failure. The information that the event of failure actually happened is decisive for the identification of the most likely failure scenario. The information on the shape of the slip surface is decisive for the identification of the most representative model choices and the most dominant parameter contributing to failure.

The First Order Reliability Method (FORM) is used in this paper to estimate the probability of failure for every generated slope stability model. However, FORM gives an approximation of the probability of failure meaning that the exact value of the probability of failure is unknown. FORM has the benefit of providing a quick answer and insight into the influence factors. It also has (potential) drawbacks. For instance, the found local design point might correspond to a non-representative failure probability. In such cases, the calculated influence factors are non-representative as well. Also, the iterative process might not converge to one answer due to numerical problems. However, the software that was used (Probabilistic Toolkit of [Brinkman 2015]) provides the option of detecting numerical problems and this was not an issue for this assessment. Also, several results of FORM calculations have been verified with parallel Monte Carlo simulation, providing very similar results.

Scenarios of failure and alternatives of model choices are assumed equally probable as a first estimate, despite that some of the scenarios and the model choices are considered less likely, such as the introduction of Bishop's method. Experts consider Bishops' method to be less accurate than both Spencer's and Uplift Van method. Moreover, slope stability computations are the bases of the likelihoods and result in posterior probabilities to determine the most likely causes. Field observations can be implemented into the analyses by including these in the slope stability computations, such as the geometry of the slip surface or the actual slope failure. Absence of clear evidence impedes the estimation of prior probabilities, but appears less important for results.

Furthermore, by including the likelihood of occurrence in the hindcasting, the outcome is considered to

reflect the actual situation more objectively than a hindcasting based on a deterministic sensitivity analysis. With a deterministic sensitivity analysis, the results are highly determined by the chosen input parameter bounds. Also, despite the equally assigned prior probabilities over the discrete choices as a first estimate, which is considered as a very rough estimate, experts in the field might be able to suggest more appropriate prior probabilities. Altogether, it is still possible that the actual cause was not identified as part of the considered scenarios.

4.2. Application of probabilistic hindcasting to the Breitenhagen case

The six steps of hindcasting are applied to the levee failure near Breitenhagen. Several elements in the analysis are based on local data such as the scenarios of water pressure and part of the soil parameters. Other elements rely on generic inputs such as some shear strength parameters; see Tables 4 and 5. The more local data are available, the more accurate the hindcast. In this case, it was not feasible to collect more local data and we believe there was sufficient data available to draw conclusions. But more local data might yield in even more conclusive conclusions.

The connection between the river and the aquifer, as presented in "Pond connection with aquifer" Scenario (S_3), results in high water pressures building up underneath and in the toe area of the levee. High water pressures in these areas reduce the local shear strength. These conditions influence the toe area (passive area) of the levee the most, resulting in a non-linear and stretched slip surface when slope instability occurs. Observations, illustrated in Figure 9, confirm that the passive area of the slip surface stretches up to the electrical pole at the toe of the levee and tilts the pole. Thus, the observed stretched passive part of the slip surface indicates high pore pressures underneath the levee. Other scenarios of water pressures concentrate on the active area of the slip surface, leaving the passive area relatively intact and result in a slip surface which is less stretched. Even though the other scenarios are not very probable, they cannot completely be excluded. Also, the method of Spencer is most robust and accurate when calculating critical slip surfaces that are not typically circular shaped. This explains why the posterior probability favours the method of Spencer when incorporating the observed slip surface information (h') (Figure 11).

Whether the soil is best characterised by drained or undrained behaviour is, in reality, less binary than the models suggest. The soil behaviour is probably best characterised as partially drained which explains the relatively small difference between the resulting posterior probabilities (Figure 13).

The forensic analysis of Kool et al. (2019) using a deterministic sensitivity analysis-based hindcasting, reports a collection of identified dominant parameters contributing to failure: that deviating low shear strength associated with low values of POP or cohesion justify the failure. This is in contrast to the findings of the current analysis, which identifies the low shear strength associated with low values of Shear strength ratio. The difference is explained by: (1) Undrained soil behaviour is identified as the most likely model choice in the analysis and, therefore, automatically excludes the cohesion as the most dominant parameter, (2) The applied range of the values of POP (by Kool, 2019) covers situations that are less relevant for this specific situation and results in FoS << 1. This low value of FoS explains why the POP is identified as a dominant contribution to failure when a deterministic sensitivity analysis of hindcasting is used. However, when parameter values are related to their likelihood of occurrence, the Shear strength ratio values resulting in failure are more probable to occur, which explains the high influence factor.

In this study, the water pressures are mostly based on steady-state assumptions for the phreatic line and hydraulic head in the aquifer. While a pragmatic and conservative choice for engineering purposes, this assumption can lead to the wrong conclusions for hindcasting. In general, transient effects should be incorporated in a hindcast. In this study, steady-state conditions were likely reached for uplift due to the long duration of the high water and the high conductivity of the aquifer. For the phreatic line, steady-state may be more questionable because of the low conductivity of the levee material (though the pictures do show a mostly saturated levee). A sensitivity analysis with a lower phreatic line (reflecting transient effects) does not yield a different conclusion. This is because the stability is mainly determined by the aquifer head. For cases where the levee body is of more importance for the stability, transient effects also become more important.

LEM's are used in this paper to model slope stability, where a Finite Element Model (FEM) would be better in capturing complex geometries and soil behaviour. However, in this case, a sensitivity analysis shows that the scenarios of water pressures are more important for failure than the soil behaviour. Furthermore, the observed slip surface does not follow a complex shape. Therefore, in combination with its computational efficiency, we chose a LEM as this is deemed sufficiently accurate. FEM might, however, be more appropriate in case of more complex geometries and soil behaviour, and in case water pressure scenarios are less dominant. In

order to use the benefits of FEM, more local information would be needed as well (hindcasting this is difficult to collect more data as this mostly vanished), especially if spatial variability is to be incorporated.

The findings of the forensic engineering report by Grubert (2013) identified a collection of causes: unexpected saturation of the levee, steep slope of the levee and, foremost, the influence of the tree roots. However, this present study shows that a pond in front of the levee is the most likely cause of failure. The findings of Grubert (2013) are possible as well, as the computed scenario probabilities are not negligible. Both visual and historical data indicate that the pond in front of the levee is likely to be leftover from a former breach (Sixdorf 2016). Due to the reparations of this breach, the conditions of the soil might deviate from other stretches of the levee. The most likely scenario of a local high hydraulic head in the aquifer and the most dominant parameter identified by the present study can directly be connected to the former breach and the reparations, which would explain the actual location of the breach in 2013.

5. Conclusions and recommendations

5.1. Conclusions

This paper demonstrates the application of probabilistic Bayesian techniques to the hindcasting of levee failures due to slope instability in six steps. The method provides insights into the most relevant uncertainties of hindcasting by evaluating all possible scenarios of failure and model choices on the likelihood of occurrence. This results in the identification of the most likely scenario and most dominant parameters triggering a failure. The suggested steps of analysis provide a thorough, workable and transparent method of analysis. Furthermore, this method is an improvement of existing deterministic hindcasting methods (Kool et al. 2019) in the sense that it provides a better insight in the relative likelihoods of the various possible causes to failure.

The developed approach of hindcasting is applied to the levee failure near Breitenhagen, with the identified most likely scenario of failure, the most representative model choices characterising the failure, and the most dominant parameters triggering the failure as a result, despite the scarcity of evidence. The levee failure near Breitenhagen in Germany in 2013 is most likely triggered by locally weak soil conditions and unexpected high water pressures due to a connection between a pond on the riverside of the levee and the aquifer underneath the levee. The slope stability model that characterises this failure most accurately is implemented with LEM of Spencer and undrained shear strength soil behaviour. Within this combination of a failure scenario and model choices, the shear strength ratio is identified as the most dominant contributor to the failure. The contribution of the high hydraulic head due to the pond connection is identified to be the second dominant. Based on the available evidence, an old levee breach explains both the presence of a pond and the locally weak soil (Sixdorf 2016) and would thus explain the specific location of the breach.

5.2. Recommendations

Even though the probabilistic method of hindcasting is successfully applied to the levee failure near Breitenhagen as part of the forensic analysis, the method can be improved by analysing more levee failures. Historical cases such as New Orleans (2005), would make an interesting object of study, and the method can also be further developed for other failure mechanisms, such as piping. Overall, it is expected that the developed approach can support a more systematic analysis of other levee failures.

In this paper, we have used FORM for probabilistic calculations. In some cases, this method could lead to inaccurate estimates of the design point and probability of failure. In future assessments, it is recommended to investigate the (parallel) use of other methods, such as Monte Carlo, especially in case on non-linear limit state functions and complex failures.

This method shows that including observational information in the hindcasting is vital to the identification of the most dominant contributing variable. Therefore, it is recommended to explore whether it is possible to include more evidence in the hindcasting with the help of Bayesian techniques, e.g. past performance information.

In order to characterise failure, LEMs are used in this paper as the slope stability is mainly determined by water pressure scenarios and the slip surface relatively simple. Finite Elements Method (FEM) analyses can lead to a more accurate characterisation of slope instability, and should especially be considered for complex geometries and soil behaviour (see e.g. Varkey, Hicks, and Vardon 2017). However, this comes at the cost of a computational burden.

Acknowledgements

We thank both Dr.-Ing. Thilo Weichel of the Landbetrieb fur Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt and Dr.-Ing. Torsten Heyer of Dresden University of Technology for providing us with relevant data of the levee failure near Breitenhagen. This research was performed as part of the NWO TTW project SAFElevee (project number 13861). We acknowledge the valuable inputs of the anonymous reviewers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO TTW): [Grant Number 13861].

References

Baecher, G. 2017. "Bayesian Thinking in Geotechnics." Proceeding of GeoRisk 2017, Denver.

Baecher, Gregory B., and John T. Christian. 2003. Reliability and Statistics in Geotechnical Engineering. Chichester: John Wiley and Sons Ltd.

Baker, J., and E. Calle. 2006. JCSS Probabilistic Model Code, Section 3.7: SOIL PROPERTIES (revised version). Joint Commitee on Strcutural Safety.

Bishop, Alan W. 1955. "The use of the Slip Circle in the Stability Analysis of Slopes." *Géotechnique* 5: 7–17.

Brinkman, R. 2020. Probabilistic Toolkit. In User Manual. Delft: Deltares.

Calle, E. 2008. "Statistiek bij regionale proevenverzamelingen." *GEOtechniek* 1: 40–44.

Carper, Kenneth L. 2000. *Forensic Engineering*. Florida: CRC Press LLC.

CIRIA. 2013. *The International Levee Handbook*. London: CIRIA.

Drews, N. 2017. "Investigation of the dike failure near Breitenhagen (Saale River) during the flood in June 2013." Hydraulic engineering TU Dresden

Gilbert, R. B. 2016. "Important Role of Uncertainty in Forensic Geotechnical Engineering." In *Forensic Geotechnical Engineering*, edited by V.V.S. Rao and G.L. Sivakumar Babu, 493–504. New Delhi: Springer. doi:10.1007/978-81-322-2377-1.

Gilbert, R. B., S. G. Wright, and E. Liedtke. 1998. "Uncertainty in Back Analysis of Slopes: Kettleman Hills Case History." *Journal of Geotechnical and Geoenvironmental Engineering* 124: 1167–1176.

Grubert, P. 2013. Saaledeich bei Breitenhagen, Geotechnische Untersuchungen der Bruchstelle Empfehlungen zur Sanierung Magdeburg: Gesellschaft für Grundbau und Umwelttechnik mbH.

Helpdeskwater. 2016. *Schematiseringshandleiding macrostabiliteit*. 1.0. the Netherlands: Ministerie van Infrastructuur en Milieu.

Honjo, Yusuke, Hirotoshi Mori, Masanori Ishihara, and Yu Otake. 2015. "On the Inspection of River Levee Safety in Japan by MLIT." *Geotechnical Safety and Risk V* 5: 873–878.

Jonkman, S. N., and T. Schweckendiek. 2015. "Briefing: Lessons Learned from Failure of Flood Defences." *Forensic Engineering* 168: 85–88.

Kool, J., W. Kanning, T. Heyer, C. Jommi, and S. N. Jonkman. 2019. "Forensic Analysis of Levee Failures: The Breitenhagen Case." *International Journal of Geoengineering Case Histories* 5 (2): 70–92. doi:10.4417/IJGH-05-02-02.

Kool, JJ, W Kanning, T Heyer, C Jommi, and SN Jonkman. 2019. *Forensic Analysis of Levee Failures: The Breitenhagen Case*. Vol. 5, 70–92. doi:10.4417/IJGCH-05-02-02.

Kremer, R.H.J., M.T. Van der Meer, J. Niemeijer, B.A.N. Koehorst, E.O.F. Calle, and TAW. 2001. *Technisch Rapport Waterkerende Grondconstructies (TRWG)*. Delft, the Netherlands: Rijkswaterstaat, DWW.

Ladd, C. C. 1991. "Stability Evaluation During Staged Construction (22nd Terzaghi Lecture)." *Journal of Geotechnical Engineering* 117: 540–615.

Luckman, P. G., A. Der Kiureghian, and N. Sitar. 1987. "Use of Stochastic Stability Analysis for Bayesian Back Calculation of Pore Pressures Acting in a Cut at Failure." 5th Int. Conf. on Application of Statistics and Probability in Soil and Struct. Engr., Vancouver, Canada.

Normcommissie 351 006 "Geotechniek". 2011. *NEN9997-1 Geotechnisch ontwerp van constructies-deel 1: Algemene regels. Delft: Koninklijk Nederlands Normalisatie-instituut*. Delft, the Netherlands: Nederlands Normalisatie-instituut.

Probabilistic Toolkit 1.8.12.629. Deltares, Delft.

Rackwitz, R. 2001. "Reliability Analysis—A Review and Some Perspectives." *Structural Safety* 23 (4): 365–395.

Rozing, A. 2015. *Onzekerheden Waterspanningen in WTI 2017*. Delft: Deltares memo 1220083-004-GEO-0003.

Schofield, A. N., and C. P. Wroth. 1968. *Critical State Soil Mechanics*. Cambridge: McGraw-Hill.

Schweckendiek, T. 2014a. *On Reducing Piping Uncertainties*. Delft: T.Schweckendiek.

Schweckendiek, T. 2014b. *On Reducing Piping Uncertainties – A Bayesian Decision Approach*. Delft: Delft University of Technology.

Schweckendiek, T., M. van der Krogt, B. Rijneveld, and A. M. Teixeira. 2017. *Handreiking Faalkansanalsye Mactrostabiliteit*. Delft: Deltares.

Seed, R. B., R. G. Bea, A. G. Abdelmalak, A. G. Athanasopoulos, G. P. Boutwell, J. D. Bray, J. Briaud, et al. 2006. *Investigation of the Performance of the New Orleans Flood Protection System in Hurricane Katrina on August 29, 2005*. Berkeley: University of California.

Sixdorf, H. A. 2016. SAVElevee, Notes on the Dike Chronic by Heinz Armin Sixdorf. edited by N. Drews.

Spencer, E. 1967. "A Method of Analysis of the Stability of Embankments Assuming Parallel Inter-Slice Forces." *Géotechnique* 17: 11–26.

TAW. 2004. "Technisch Rapport Waterspanningen bij dijken".

USACE. 2000. Design Guidance for Levee Under-Seepage. USACE: Engineering Manual EM 1110-2-1913.

van Deen, J., and A. van Duinen. 2016. *Schematiseringshandleiding Macrostabiliteit*. Delft: Deltares.

Varkey, Divya, Michael Hicks, and Phil Vardon. 2017. *Influence of Spatial Variability of Shear Strength Parameters on 3D Slope Reliability and Comparison of Analysis Methods*. Denver, United States.

Vorogushyn, S., B. Merz, and H. Apel. 2009. "Development of Dike Fragility Curves for Piping and Micro-Instability Breach Mechanisms." *Natural Hazards and Earth System Sciences* 9: 1883–1401.

Weichel, Thilo. 2013. "Failure of the Breitenhagen levee (2013)." Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt.

Zhang, J., W. H. Tang, and L. M. Zhang. 2010. "Efficient Probabilistic Back-Analysis of Slope Stability Model Parameters." *Journal of Geotechnical and Geoenvironmental Engineering* 136: 99–109.