Structural Safety 60 (2016) 117–129

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

Soil stochastic parameter correlation impact in the piping erosion failure estimation of riverine flood defences

J.P. Aguilar-López [a,*], J.J. Warmink [a], R.M.J. Schielen [a,b], S.J.M.H. Hulscher [a]

[a] University of Twente, Enschede, The Netherlands
[b] Rijkswaterstaat, Arnhem, The Netherlands

ARTICLE INFO

Article history:
Received 6 March 2015
Received in revised form 12 January 2016
Accepted 16 January 2016

Keywords:
Piping erosion
Correlation
Sellmeijer
Failure probability
Copula
Flood defence

ABSTRACT

Piping erosion has been proved to be one of the failure mechanisms that contributes the most to the total probability of failure on the Dutch flood defence systems. The present study aimed to find the impact of correlation and tail dependence between soil parameters present in the Sellmeijer revised limit state equation for piping safety assessment, particularly between the grain size and hydraulic conductivity parameters. A copula based random sampling method was used as a tool to include this effect in the probabilistic estimation of this type of failure. The method was framed in a real case study for a flood defence along the Lek river, in The Netherlands. The results showed that inclusion of correlation between the two parameters reduces the variance of the limit state marginal distribution by almost 10% when compared to the uncorrelated case. This effect changes the tail values sampling frequency and therefore reduces the probability of failure by a factor of 1.7. The omission of correlation between the two parameters for safety assessment based on the Sellmeijer limit state function may result in over dimensioned structures.

1. Introduction

In The Netherlands, large flood risk assessment projects such as the VNK2 [1] have devoted great attention to develop and improve more robust probabilistic estimation methods for the safety assessment of their levee systems. One of the main results from this study was the prioritization of the different failure mechanisms that contribute the most to the total failure probability (Pf) of the levee systems. Reiteratively, piping erosion (PE) was found to be a major threat in most of the components of the system. This type of failure consists in a progressive erosion channel under the flood defence foundation which will eventually start a breaching process due to the loss of stability of the structure. This type of failure can be simulated by the numerical model developed by Sellmeijer [2]. For safety assessment, a revised limit state equation (LSE) [3] was derived based on this same model. This equation describes the safety state of the system given the most sensitive variables involved in the process for the occurrence of this particular failure mechanism. Limit state equations are implemented in probabilistic safety assessments as they can be used to expresses the loads experienced by the flood managing structure as a function of the water level probabilistic distribution. The resistance of the structure against these loads can also be represented as a probabilistic marginal distribution.

It is common in practice to assume that the random variables used for the limit state evaluation are represented by univariate probability density functions. Hence, they are commonly assumed as uncorrelated when no evidence is available. The omission of possible statistical dependence or correlation between different state variables is one major source of error in the failure estimation of reliability of a system when such variables are highly sensitive for the model probabilistic outcome. Correlation analysis is not only concerned about the degree of dependence but also the temporal and spatial distribution of the correlated random variables [1]. Extensive research has been done about the effect of spatial correlation of load and resistance of flood defences in The Netherlands [4–6]. Yet, the correlations were analyzed considering how a variable depends on itself (autocorrelation) along space and time and not within each other. The importance of variable correlation for flood defence structures was demonstrated in the study by Diermanse and Geerse [7] where the influence of bivariate

Abbreviations: PE, piping erosion; Pf, failure probability; LSE, limit state equation.

* Corresponding author at: Drienerlolaan 5, 7522 NB Enschede, The Netherlands.
E-mail addresses: j.p.aguilarlopez@utwente.nl (J.P. Aguilar-López), J.J.Warmink@utwente.nl (J.J. Warmink), R.M.J.Schielen@utwente.nl (R.M.J. Schielen), S.J.M.H.Hulscher@utwente.nl (S.J.M.H. Hulscher).

http://dx.doi.org/10.1016/j.strusafe.2016.01.004
0167-4730/

Sellmeijer revised LSE nomenclature

Zp	residual resistance (limit state) [m]
η	sand drag force factor (0.25) [–]
γ'_{sand}	unitary weight of submerged sand particles [kN/m^3]
γ_w	unitary weight of water [kN/m^3]
θ	bedding angle of sand grains [deg]
d_{70}	70 percent quintile value grain size distribution of sand layer [m]
d_{70m}	calibration reference value (2.08×10^{-4} m) [m]
v	kinematic viscosity of water at 20 °C [m^2/s]
K	hydraulic conductivity of sand [m/s]
g	gravitational acceleration (9.81 m/s^2) [m/s^2]
D	average thickness of sand layer [m]
m_p	modelling uncertainty factor [–]
H_c	critical hydraulic head difference [m]
F_R	resistance factor [–]
F_S	scale factor [–]
F_G	geometric factor [–]
H_c	critical PE resistance head [m]
L	seepage length from entrance point to sand boil water exit [m]
H	water level in the foreside of the flood defence [m]
h_b	water level at hinterside outflow point [m]
d	impermeable layer thickness at the sand boil exit point [m]
Note	The product of the hydraulic conductivity of soil and kinematic viscosity divided by the gravitational acceleration is equal to the intrinsic permeability k [m^2] (noted as lower case k).

correlation modelling between two hydro climatological variables required for dike safety assessment was studied. One case study was done by modelling the inflows from the IJssel River and the water levels in the IJssel Lake, and another one for modelling the wind speeds and water levels in the North Sea coast correlated as well. Both case studies showed the influence that correlation modelling can have in the safety assessment of flood defences. Yet, the correlation impact was only studied in the variables involved for estimating the marginal distributions of the loads applied to the flood defences.

With this study it is intended to quantify the influence of correlation in failure estimation between two parameters present in the erosion model for piping assessment, in particular for a "single" cross section in a riverine flood defence when assessed by the revised Sellmeijer LSE [3]. This equation includes, the representative aquifer grain size parameters ($d70$) and the hydraulic conductivity parameter (K) which can be correlated for different models as presented by [8]. For the design and safety assessment of flood defences in The Netherlands, the dependence of these two parameters is considered by the empirical equation present in [9]. The drawbacks of this equation are that is only valid for sands with $d10 < 0.06$ mm and it also depends on a qualitative factor associated to the packing density of the particles in situ. Hence, the implementation of such equation in a fully probabilistic assessment (correlation inclusion) becomes unreliable. Note that the correlation addressed with this kind of equations (grain size versus hydraulic conductivity) represent the chance that the two variables are dependent disregarding their location (spatial correlation) in time and space (non-stationary process).

Despite including the dependence of these two parameters, the correlation degree between the two of them is not constant for all quantiles either. Base on the sample distribution, a higher correlation is expected for sands with larger percentages of smaller grains. Such variability of the correlation is known as tail dependence. This is also not included in the actual probabilistic assessment methods for flood defence reliability of PE and can have an important effect in the structure reliability.

During the PE process, only the most upper part of the aquifer is eroded which means that the $d70$ statistical distribution should be representative of mainly that zone. It is common to find finer grain distributions in the upper layer of the aquifers which will imply a lower correlation degree between the $d70$ and the K parameters. When the grain distribution of the most upper layer of the aquifer is significantly different and finer with respect to the one associated to the whole aquifer average distribution, the measured representative conductivity values for the whole aquifer can be assumed as uncorrelated. However, in the actual practice a detailed sampling procedure of only the upper layer of grain size and permeability is not practical for such longitudinal structures. Hence, the $d70$ and K statistical descriptors are estimated by indirect methods.

The fact that the upper layer may have a distinct granulometric distribution with respect to the lower layers, does not imply that correlation and/or tail dependence between $d70$ and K are negligible. It will only mean that the degree of dependence between the two parameters is lower than expected. Yet, this degree of correlation and tail dependence might change the probabilistic outcome if found significant enough. It is also important to state that not all variables involved in this process necessarily should be considered as correlated, despite the fact that significant correlation can be estimated from their dataset. Sufficient physical evidence of the origin of the correlation should be proven before deciding to include its effect in a structural reliability assessment. In other words correlation does not necessarily implies causation.

For the data of an existent river flood defence located in The Netherlands along the Lek river, the correlation and its physical origin were studied.

In order to structure the research, three main questions were addressed:

1. Is there considerable correlation between the representative grain size ($d70$) and the hydraulic conductivity (K)?
2. How to select and validate a correlation bivariate model (copula family) for the failure estimation due to piping?
3. How important is the impact of correlation between $d70$ and K in the failure probability estimation due PE when estimated by Sellmeijer revised limit state equation?

The outline of the paper consist in the physical process of the PE failure mechanism and its limit state function which are explained in detail in Section 2, plus the implementation of the copula functions for generating the correlated random samples. In Section 3 the case study and the input data used for the failure estimation are described. Section 4 describes the results obtained from estimating correlation from the field collected data. In Section 5 the selection and validation of a model that describes best the soil behaviour is presented. In Section 6, the results of the correlation effect on the limit state function marginal distribution and failure probabilities are presented. In Section 7, the results of each research question are discussed and finally the main conclusions of the study are presented in Section 8.

2. Limit state safety assessment method

For the present study, the random variables involved in the Sellmeijer revised LSE are sampled and propagated by a Monte Carlo reliability method. As a proposed tool to assess the effect of correlation between the sensitive random variables in PE estimation, the statistical bivariate joint distribution method for correlated sampling known as "copula" [10] was implemented. Copula joint distribution models are capable of inducing correlation between two univariate marginal distributions while maintaining their statistical moments fixed. Different copula functions can be found in the literature according to which topological behaviour is desired when tail dependence representation is required [11]. The degree of correlation is one of the copula main input parameters. Therefore, it also allows to generate n random samples with a particular desired degree of dependence. The final product consists of a table of failure probabilities as function of the degree of correlation estimated for each of the chosen copula models.

2.1. PE Sellmeijer revised limit state equation

PE is also known as backward erosion, and consists in the loss of stability of the flood defence structure due to the erosion of the granular foundation stratum (Fig. 1). In order for PE to happen, a previous failure mechanism called "uplift" must have occurred as well.

This mechanism consists in the lifting and breakage of the impervious layer between the body and the foundation of the dike, due to a high hydrostatic pressure. Afterwards, the water movement inside the aquifer from the river side towards the inland side, transports fine grains which originate a longitudinal void also referred as "pipe". Once the pipe has developed for a length equal or greater than the width (L) of the flood defence (Fig. 1), it is assumed to be in failure state. Nevertheless, the breaching of the flood defence may not necessarily occur in that exact instant. Several empirical and numerical models (equations) for estimating the critical head of PE have been developed since the early 20th century, such as Bligh and Lane [12]. A more robust conceptual numerical model and LSE was developed in The Netherlands by Sellmeijer [2].

Recently, the LSE was re-calibrated with the obtained results of several experiments at different scales [3]. The LSE is presented in Eqs. (1)–(5).

$$Z_p = H_c - (h - h_b - 0.3d) \tag{1}$$

$$H_c = m_p(F_G)(F_R)(F_S)L \tag{2}$$

$$F_S = \frac{d_{70m}}{\sqrt[3]{\left(\frac{\nu K}{g}\right)L}}\left(\frac{d_{70}}{d_{70m}}\right)^{0.4} \tag{3}$$

$$F_R = \eta\frac{\gamma''_{sand}}{\gamma_w}\tan(\theta) \tag{4}$$

$$F_G = 0.91\left(\frac{D}{L}\right)^{\left(\frac{0.28}{\left(\frac{D}{L}\right)^{2.8}-1}+0.04\right)} \tag{5}$$

In many cases, a small water ditch is located behind the dike (Fig. 1) which serves as a drainage and irrigation control structure. Then, the hydraulic head exerted by the flood defence is calculated as the difference between the riverside withstanding water level (H) and the landside ditch water level (h_b) plus an additional resistance margin estimated as 30% of the aquitard layer thickness. This margin takes into account the additional flow resistance from the vertical flow through the crack in the impermeable layer [13].

2.2. Hydraulic conductivity by Kozeny–Carman equation

The Kozeny–Carman equation estimates the hydraulic conductivity based on the soil representative diameter and the porosity, which is commonly used in studies where in situ measurements are scarce. It is applicable to large range of grain and several types of soil as presented by Chapuis [14], which makes it suitable for stochastic sample generation. Mazzoleni et al. [15] also used the Kozeny–Carman equation for PE evaluation along the Po river where the porosity showed to be an important parameter which can be also used for the failure prediction.

$$K = \frac{\rho g}{\mu}\frac{n^3}{(1-n)^2}\frac{d_m^2}{180} \tag{6}$$

where:

K	[m/s^2]	Hydraulic conductivity
ρ	[kg/m^3]	Water density
g	[m/s^2]	Gravitational acceleration (9.81 m/s^2)
μ	[kg s/m]	Water dynamic viscosity
n	[–]	Porosity
d_m	[m]	Representative diameter

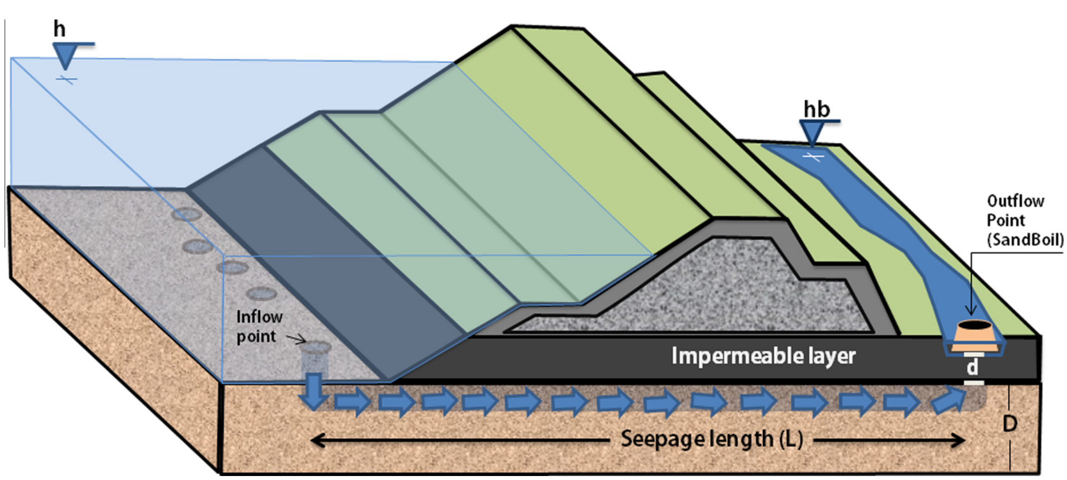

Fig. 1. PE process under flood defence.

When measured values of porosity are not available, Vukovic and Soro [16] present an empirical formula to calculate this parameter as a function of the ratio between the *d60* and *d10*, according to:

$$n = 0.255\left(1 + 0.83^{\frac{d60}{d10}}\right) \tag{7}$$

The Kozeny–Carman equation it can be implied that the porosity term has a significant influence in the hydraulic conductivity estimation. Hence, its uncertainty is also an important factor in the accuracy of the hydraulic conductivity estimation. This uncertainty will be reflected in the measured values or can be represented as an additional term affecting Eq. (7). However, information of these uncertainty related to the empirical regression is limited and is beyond the scope of the present study. Therefore it was not included in the calculations.

2.3. Copula correlation models

Copula functions can be used for generating correlated values during a random sampling process. They allow to build joint distributions from two or more variables while maintaining the statistical properties of their marginal distributions [17]. All types of possible copulas (families) are derived from the Sklar theorem which states that every probability function can be written as a copula multivariate function of the uniformly transformed marginal values [10]. The most common families are the "Gaussian" and the "Archimedean" functions (e.g. Gumbel, Frank or Clayton). For the bivariate case, the general copula cumulative probability function can be written as Eq. (8):

$$F_{X,Y}(x,y) = C[F_X(x), F_Y(y)|\theta] \tag{8}$$

where $F_X(x)$ and $F_Y(y)$ represent the cumulative distribution functions of the random variables x and y. The θ symbol represents the general correlation copula parameter which describes the degree of dependence between variables x and y and $F_{X,Y}$ is the Joint probability function given x and y. For continuous marginal distributions, there is always a unique C matrix that relates x and y. Their probability marginal are always contained inside the unit square $[0,1]^2$ for the case of bivariate joint distributions.

Tail dependence is an additional feature of correlated variables which is defined as the probability that extreme values of the variable x are achieved given large values of the variable y. Full correlation means that variable x has a one to one relation with variable y and that the probability of any other value to be associated with x is equal to 0. Conversely, fully uncorrelated variables have probability of any other value to be associated with x equal to 1. In general, bivariate variables can have the same degree of correlation with different tail dependencies. The Gaussian copulas for example are built as a function of the normal distribution and allow equal degrees of positive and negative correlation. However, they will not induce any tail dependence. Tail dependence is important for reliability assessment as the occurrence of two extreme values at the same time might either increase or reduce the chances of failure to occur in comparison to the cases where no tail dependence and no correlation is present.

The Archimedean copulas such as "Gumbel" or "Clayton" on the other hand allow to generate the samples correlated with a stronger dependence in either of the two tails of the bivariate joint distribution.

Copula families are built based on mathematical descriptions that relate the different marginal distributions of the univariate x and y distributions via different functions known as "generators" (Table 1).

Every generator function will originate a copula type based on its convexity and monotonic behaviour. The general sampling algorithm for all copula models consist in generating two uncorrelated sets of uniformly distributed samples as (0,1]. Each of these sets are assumed to be equivalent to the random cumulative probability values (u and v). Afterwards, the copula value C for each pair of samples u and v is calculated with its correspondent generator function and the chosen degree of correlation. As a result, a third set of samples is generated, which represents the copula bivariate function. Finally this Copula function which is also bounded in (0,1] can be inverse sampled with the help of two additional auxiliary random uniformly generated sets bounded and the original generator function. Note that a larger data set of the initial sampling procedure will ensure a "smoother" sampling. In the present study, the random copula generation tools included in the MATLAB program were implemented as they are optimized for efficient sampling of large random value sets.

Copulas are generated taking into account the dependence represented by the correlation parameters (α) for the Archimedean family and "ρ" Pearson's coefficient for the Gaussian family (Table 1). The "α" and "ρ" parameters at the same time, can be expressed by the Kendall's correlation coefficient (τ). Kendall's rank correlation measures the degree of dependence based on how many data points are concordant compared with the ones that are not concordant when ordered from smallest to greatest. This condition ensures that the dependence degree will not be affected by any transformation in the original dataset, especially when any of the variables is not normally distributed [21]. Regardless of the correlation coefficient selection, the dependence degree can be expressed in terms of the generator correlation parameter [18] as shown in Table 2.

Table 2
Equivalence expressions between dependence copula generator parameters and Kendall's ranking correlation coefficient (τ).

Gaussian	Gumbel	Clayton
$\tau = \frac{2\arcsin(\rho)}{\pi}$	$\tau = 1 - \frac{1}{\alpha}$	$\tau = \frac{\alpha}{2+\alpha}$

This equivalence allows to model the correlation for three different type of copulas while inducing the same degree of dependence during the random sampling process. Hence it is possible to compare the different results as a function of the correlation degree for each copula type. The tail dependency between variables will also change according to a correlation parameter.

2.4. Goodness of fit tests

In order to validate the (synthetical) type of copula and its selection criteria, several types of goodness of fit tests were

Table 1
Copula families and generator functions.

Family	Gaussian	Archimedean	
Type	Gauss	Gumbel	Clayton
Copula $C(u,v)$	$\Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$	$\exp\left(-[(-\ln u)^\alpha + (-\ln v)^\alpha]\right)^{1/\alpha}$	$\max\left[(u^{-\alpha} + v^{-\alpha} - 1), 0\right]^{1/\alpha}$
Correlation parameter	"ρ" Pearson correlation coefficient normal distribution Φ^{-1}	$\alpha \in [1, \infty)$	$\alpha \in [0, \infty)$

performed. First, one joint probability surface copula was constructed from the available dataset, as a benchmark for validation. This kind of bivariate models are also known as "empirical copula" models [11]. They are built based on the univariate marginal probabilities of each of the correlated variables. The accuracy of the empirical copula cannot be judged unless additional data from the location becomes available for validation. Afterwards, a joint probability distribution surface was generated for each of the different copula models. Bivariate distributions can be represented as three dimensional surfaces where each point has as coordinates its parameter value for the x and y axis and its joint cumulative probability in the z axis. Finally the empirical surface was compared with each of the copula surfaces in order to determine the goodness of fit of each of the copula families. The available methods range from graphical, over different residual error estimators to more formal statistical methods as it was stated by Vandenberghe et al. [19]. Any of these three different type of methods might result in different conclusions if analyzed solely. Therefore, a copula selection should only be done after performing more than one of the methods.

The graphical method consisted in generating contours from the four pre built surfaces and plotting them on top of each other. Afterwards a visual examination was performed in order to evaluate which synthetic copula looks more similar to the empirical one. As an additional check, the SAPE method is proposed in order to explain the fitting error based on the amount of available data. A second way to estimate the goodness of fit of a model is by measuring the difference between the observed data and the predicted one. For an average value the rooted mean squared error (Eq. (9)) is a suitable approach for interpolated surfaces.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_{emp}(u_{i,}v_i) - C_{family}(u_{i,}v_i)\right)^2} \tag{9}$$

The third goodness of fit test corresponds to an statistical one. Different studies [19–21] have recommended to use the Anderson Darling goodness of fit method for copula selection over methods such Kolmorogv-Smirnoff or Cramér von-Mises. This test is rank based and therefore gives more importance to the tail dependency. It is also non-parametric. Initially the test was developed to prove normality of sample populations. The two sample Anderson Darling test statistic (AD^2) is calculated as a function of the marginal probabilities of the sampled dataset with n samples. Afterwards AD^2 calculated statistic value (Eq. (10)) is compared to the AD^2 equivalent value of the standardized normal distribution with a desired degree of significance.

$$AD^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left(\ln(F(x_{(i)})) + \ln(1 - (F(x_{(n+1-i)})))\right) \tag{10}$$

In 1987 Scholz and Stevens [22] developed a generalization of the two sample Anderson–Darling test ($AD^2{}_{kN}$) applicable for any type of statistical distribution (Eq. (11)). The major drawback of this generalization is that the threshold statistic value is sample size dependent. Therefore, it needs to be calculated for every time the test is performed with the formula:

$$AD^2_{kN} = \frac{1}{N}\sum_{i=1}^{k}\frac{1}{n_i}\sum_{j=1}^{N-1}\frac{(NM_{ij} - jn_i)^2}{j(N-j)} \tag{11}$$

The main goals of this test is to try to reject the null hypothesis that states that a certain observed sample comes or behaves like a pre-defined theoretical sample. M_{ij} represents the number of observations in the ith sample which are smaller than K_{jth} value. In our case the null hypothesis will try to answer if the empirical copula behaves like the other three synthetic copula models.

2.5. Reliability method: Monte Carlo with copula random sampling

Monte Carlo method was chosen as the reliability estimation method for this study, as it allows to implement copula models in an easier way compared to more common reliability methods such as SORM or FORM [23]. The first phase of the method consists in generating the random samples of each of the parameters. The parameters are divided in two groups. The first group random samples is generated by a classical inverse CDF transform method. The second group is also generated with a CDF inverse transform method but instead of generating random uncorrelated values from a uniform distribution between zero and one, they are generated with a random copula sampling method [17] to a chosen "tau" degree of correlation. The correlated sets were generated for 10 different correlation coefficient degrees (τ) ranging from almost 0 until 0.99. This process was repeated for the three copula types. The second phase of the procedure consists in evaluating the correlated and uncorrelated sample set groups in the Sellmeijer LSE. If the critical resistance head (H_c) is smaller than the sampled hydraulic head (h–h_b–0.3d), then it is accounted as failed sample. In the third and last phase after evaluating all the generated combination samples, the probability of failure is estimated as the ratio between the number of failed samples and the total number of evaluations. The probability density function of the limit state can also be constructed from all the results obtained by each Zpi evaluation, represented afterwards a probability density function.

3. Case study: Lekdijk

The riverine flood defence called the "Lekdijk" located along the Lek river in the province of Utrecht was used as a case study. For this particular location, the VNK2 project [24] recommended a strengthening measure as the defence did not comply with the safety standard. The field sampled "Lek" dataset is composed of 76 soil gradation samples which include d10, d60 and d70 measured in the lab. No hydraulic conductivity measurements were available. Hence the K values were estimated from d60 and d10 by the Kozeny–Carman indirect method (Section 2.2). The estimated uncertainties (mean, standard deviation and type of distribution) of d70 and K are presented in Table 3.

3.1. Stochastic parameters

The prior marginal distributions used as input data are presented in Table 3. Most of the soil derived mean and variation coefficients were obtained from field samples which were further

Table 3
Input data for stochastic failure estimation of PE.

Variable	Unit	Distribution	Mean	COV	Source
η	[–]	Constant	0.25	–	VNK [24]
γ_{sand}	[kN/m^3]	Normal	26.5	0.1	VNK [24]
γ_w	[kN/m^3]	Constant	9.81	–	–
Θ	[deg]	Constant	37	–	VNK [24]
d_{70}	[m]	Log-normal	0.000333	0.15	Field
d_{70m}	[m]	Constant	0.000208	–	VNK [24]
v	[m^2/s]	Constant			–
K	[m/s]	Log-normal	0.000302	1	Kozeny–Carman
g	[m/s^2]	Constant	9.81	–	–
D	[m]	Log-normal	65	0.1	Field
m_p	[–]	Log-normal	1	0.12	VNK [24]
L	[m]	Log-normal	70	0.1	Field
H^*	[m]	Gumbel	a = 4.357	b = 0.288	HR2006 [25]
h_b	[m]	Normal	0.5	0.1	Field
d	[m]	Log-normal	4.3	0.3	Field

* The load term H is assumed to fit a Gumbel extreme distribution with shape and location parameters equal to a and b.

analyzed in the lab. For the non-measured parameters and their associated distribution type functions, recommended values were obtained from VNK [24]. For the extreme distribution fitting of the water levels (loads), information contained in the boundary conditions report for primary water defences assessment was used [25].

3.2. Uplift/heave probability estimation

For the estimation of PE failure probability it is also required to calculate the probability of uplift/heave [26]. Both failure mechanisms need to occur in order to ensure the progression of PE as explained in Section 2.1. For this particular study, the correlation between $d70$ and K does not influence the estimated probability of uplift and therefore remains constant for all simulations ($Pf_{uplift} = 0.621$). Note that from the system reliability point of view, uplift and PE are treated as a fully independent parallel system which means that the failure probability is calculated as:

$$Pf_{PE/uplift} = (Pf_{PE}) \cap (Pf_{uplift}) \tag{12}$$

3.3. Complementary data sets

For benchmarking the results obtained from the Lekdijk dataset, the samples and measurements presented in the studies of the datasets presented by Vienken and Dietrich [27] and by van Beek et al. [28] were used. The first one corresponds to a field campaign of an aquifer in Bitterfield (Germany). The dataset include borehole sampling of $d10$, $d60$ and its correspondent slug test measurements of hydraulic conductivity. The second one corresponds to different Dutch sand samples used for different scale experiments. The data set includes $d70$ grain sizes and their correspondent hydraulic conductivity lab measurements. This data set is mostly composed of heterogeneous sands.

4. Results: correlation degree between *K* and *d*70

Sellmeijer et al. [3] also showed that the hydraulic conductivity and the $d70$ representative diameter have the highest relative predictive importance for the Sellmeijer revised limit state equation. Therefore, the present study only considered the effect of correlation of these two parameters. For the present case study, samples from the aquifer stratum were analyzed and characterized in order to determine their *d10*, *d60* and *d70* representative values. Ideally, conductivity in situ measurements will ensure that the estimated correlation is closest to the real one. However, for the present study there is not sufficient measured data in terms of hydraulic conductivity. By the use of the Kozeny–Carman equation, it was possible to estimate the hydraulic conductivity for each soil sample. From the obtained results two major aspects can be highlighted. The first one is that the value for the mean univariate marginal distribution of the generated hydraulic conductivity (2.87×10^{-4} m/s), is similar to the one used in the VNK2 study (2.5×10^{-4} m/s). Their coefficient of variation (0.98 and 1.0) are very similar as well. This shows that the results by the VNK and the ones obtained from this study are comparable. The second aspect is concerned about the possible inducement of spurious correlation from the fact that the hydraulic conductivity was also calculated from the grain sampled data. The correlation between the grain size quantiles and the estimated conductivities estimated by Kendall's method for the Lek dataset are presented in Table 4.

Table 4
Kendall's correlation coefficient between variables.

Kendall's τ	*d10* [mm]	*d60* [mm]	*d70* [mm]	*K* [m/s]
d10 [mm]	1.00	0.786	0.764	0.930
d60 [mm]		1.00	0.968	0.714
d70 [mm]			1.00	0.692
K [m/s]				1.00

The Kozeny–Carman model estimates the conductivity values as a function of the $d10$ and the porosity. At the same time, the porosity is estimated from the $d10$ and $d60$. Nevertheless, $d10$, $d60$ and $d70$ represent three independent lab measurements. For the Lekdijk dataset, the obtained Kendall coefficient of correlation between $d70$ and K is 0.692 (Table 4). The three different datasets (Lekdijk, Bitterfield and Dutch Sands) correspond to heterogeneous, highly heterogeneous and homogeneous (lab and real scale experiments) aquifers measured in different conditions (saturation, temperature, compaction) which can have a significant effect in the correlation analysis of each of the sets. Nevertheless, the main goal of analyzing the three datasets is to find out the order of magnitude of potential correlation while using three different methods of conductivity measurement (indirect Kozeny–Carman computed values, in situ slug test and direct flow measurement from the experimental set up). The first dataset contained grain size values for a highly heterogeneous unconsolidated aquifer located in Germany. From their dataset, $d10$ and $d60$ values were measured for 108 samples where a Kendall's τ of 0.723 was obtained. This value can also be validated from the statistical significance point of view by the p-value test. This test represents the probability of obtaining this same correlation coefficient while the null hypothesis remains true ($d70$ and K are uncorrelated). The p-value with significance 0.05 was 9.9×10^{-19} for this dataset. Additionally, 17 samples were presented in the same study with their correspondent in situ hydraulic conductivity measurement. From this set, a Kendall's τ of 0.5396 with a p-value of 0.0014 is obtained. The second dataset from the study of van Beek et al. [28] is composed by values of $d70$ and K of 50 samples of different Dutch sands used in small, medium and real scale experiments for PE. The Kendall's τ obtained for this dataset is 0.522 with a p-value of 5.1e−7 which shows that the estimated tau value is statistically significant.

The tail dependence is another important issue to be studied in reliability assessment of structures as extreme tail values become more important when correlated with other variable extreme values during failure estimation. For the Lekdijk dataset, one may think that based on the plotting of the actual variable values (Fig. 2a), a stronger left tail correlation might be evident.

While this might be true, it is also important to stress out that correlation of variables is measured between its probability marginal and not by the correlation between the actual variable values. In that sense it can be observed in Fig. 2b, the dataset presents strong tail dependence in the left tail and mild tail dependence in the right side when plotted in terms of the variable marginal probabilities. There is also higher scattering of the data points around the mean value area of both variables compared to their tails, which implies that a tail dependent model is suitable for representing the relationship. For the datasets from the previously mentioned studies presented by Vienken and Dietrich [27] and by van Beek et al. [28], only the last one showed potential left tail dependence (visual inspection).

5. Results: copula model selection and validation

5.1. Graphical method

From the obtained empirical surface, contours were extracted, each spaced by 0.1 joint probability units. Next, the three copula

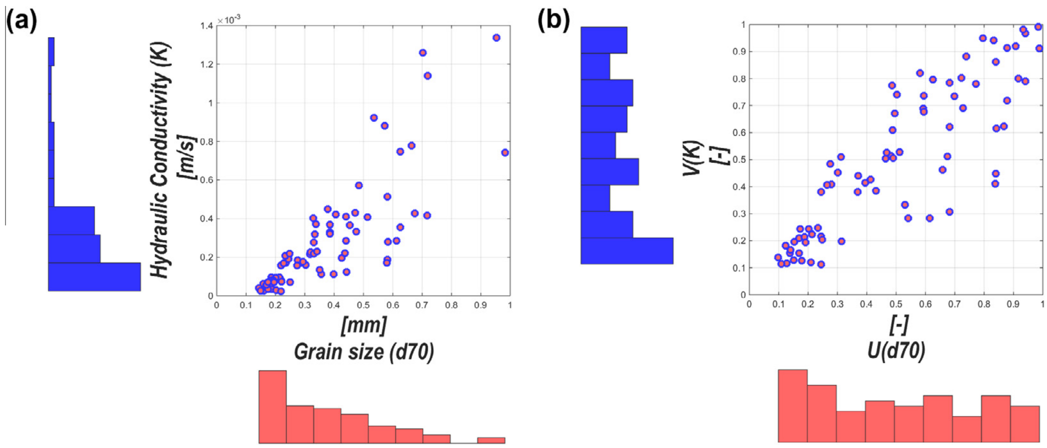

Fig. 2. (a) $d70$ vs K and (b) marginal probabilities $U(d70)$ vs $V(K)$.

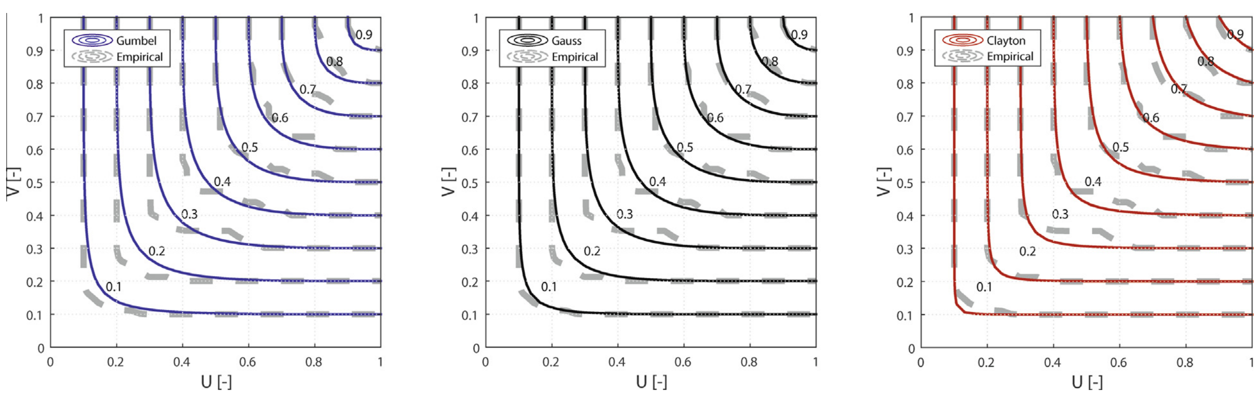

Fig. 3. Generated empirical probability contours (dashed) vs Gumbel, Gauss and Clayton copula probability contours with rank correlation coefficient $\tau = 0.692$.

surfaces (Gumbel, Gauss and Clayton) were generated while maintaining the same Kendall's correlation coefficient from the Lekdijk dataset ($\tau = 0.692$).

After plotting the generated contours (Fig. 3) it is evident that the three copulas do not outstand between each other in their general performance. As the sampled data does not present a smooth behaviour, the points located in the vertex of the empirical copula are not smoothly curved as the theoretical ones. The available dataset does not cover uniformly and sufficiently the unit space. Therefore, the observed rough edges are a result of the linear interpolation method. From the results of this method, the Gumbel copula can be discarded as the "empirical" contours are always further from the fitted ones.

As these contours were generated from interpolated surfaces, it is possible to visualize the regions which present higher errors by estimating the spatial absolute percentual error (SAPE):

$$\mathrm{SAPE}_{ij} = \left(\frac{|C_{\mathrm{family}}(u_i, v_i)|}{C_{\mathrm{emp}}(u_i, v_i)} - 1\right) \times 100\% \qquad (13)$$

This type of error is an indicator of the spatial distribution and goodness of fit of the copula models. It is interpreted as the percentage of discrepancy between the empirical copula and the other copula types associated in the space (Fig. 4). The marginal probabilities u and v are correspondent to the variables $d70$ and K respectively.

From the SAPE results shown in Fig. 4, it can be observed that the Gumbel copula presents errors in almost 30% around the mean value area ($U = 0.5$, $V = 0.5$), whereas the Gauss and Clayton copulas present errors that are lower than 10%. In the case of the right tail dependence, the two models with higher scattering (Fig. 2a) for right tail values represent better the soil behaviour as they present a lower SAPE.

From the left tail dependence, the Clayton copula performs better despite the fact that the extreme value representation is less accurate compared with the other two. This can be explained as the empirical copula surface does not have sufficient data for representing (interpolating) correctly the left tail dependence. In addition, it can be observed (Fig. 2b) that the left lower corner area ($U \leqslant 0.1$, $V \leqslant 0.1$) has no available points that represent extreme left tail values in the dataset. This is the main reason why all copulas present higher errors in that area. Furthermore, the Clayton copula surface is steeper around that area. Then, the difference between the empirical and the Clayton surfaces is higher around that area as it is shown in Fig. 4.

5.2. Root mean square error (RMSE)

The RMSE is calculated for each of the 76 sampled points used for the copula fitting. The best performance was obtained for the actual calculated tau correlation coefficient ($\tau = 0.692$) with a Gauss copula (Fig. 5). For low τ values, the Clayton copula always performs better. For higher values, both Gumbel and Gaussian copulas present almost the same value of RMSE.

The results presented in Fig. 5 show that for all copula families, $\tau = 0.692$ results in the minimum RMSE value for each copula fitting. This value is chosen based on the results presented in Table 4 obtained from the Lek dataset. Note that the trend was represented by evaluating values every 0.1 measures of τ. Nevertheless the

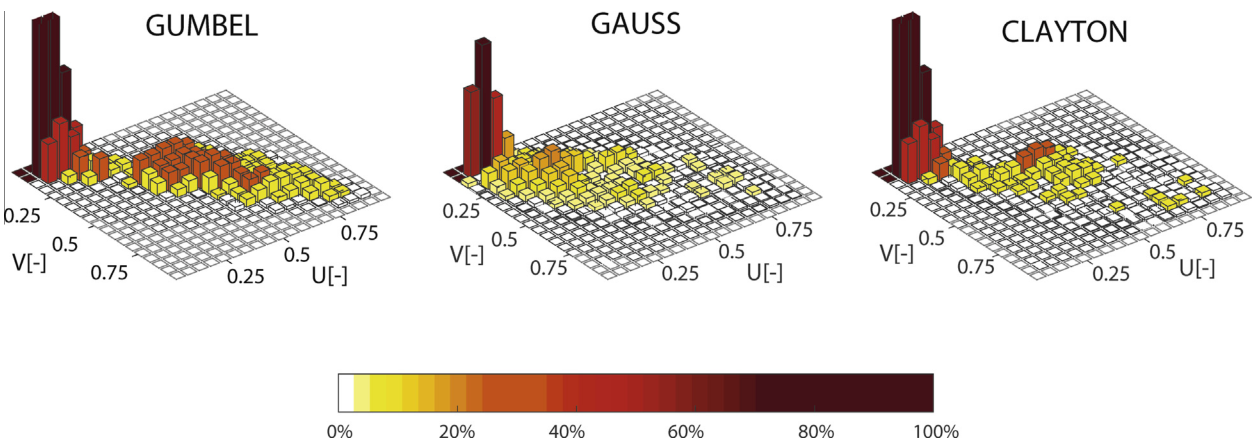

Fig. 4. SAPE of the three different copula functions.

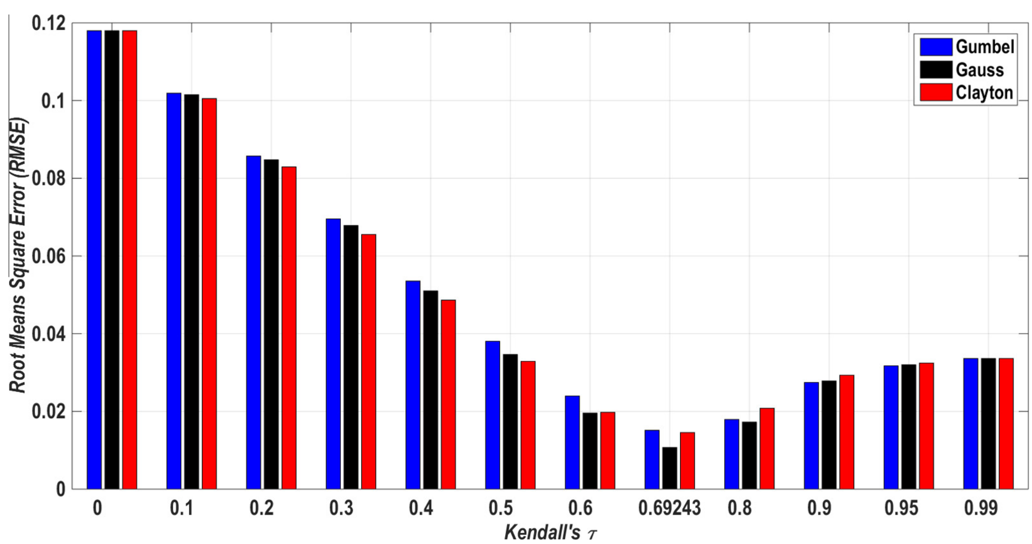

Fig. 5. Root mean square errors for different Kendall's rank correlation coefficients (τ) estimated from the 76 samples.

Table 5
Anderson Darling test results for the three copula fittings.

C_{family}	Null hypothesis H_0 $C_{emp} = C_{family\ type}$	Probability of p 5% significance	AD^2_{kN} statistic	AD^2_{kN} critical value
Gumbel	Failed to reject	0.3801	0.9574	2.4948
Gaussian	Failed to reject	0.3479	1.0176	2.4948
Clayton	Failed to reject	0.2304	1.3059	2.4948

lowest RMSE for each copula might be achieved in different but very close values to 0.692. Hence it was replaced instead of the 0.7 in Fig. 5.

From this plot it can also be concluded that neglecting correlation ($\tau = 0$) may result in higher errors despite the exact representation of the tail dependence. For example when comparing the RMSE value found for the correlation value estimated for the Lek dataset (Table 4, $d70$ versus K is $\tau = 0.692$) the error is 0.0107 with respect to the 0.118 obtained for the 100% uncorrelated case ($\tau = 0.0$).

5.3. Formal statistical goodness of fit test

For the present study, the generalized Anderson–Darling test (AD^2_{kN}) was implemented using the 76 empirical copula values of U and V coordinates, against the copula values of the three copula families for the exact same marginal probability coordinates (U, V). The results in Table 5 show that the three synthetic copula families are capable of representing the empirical surface since AD^2_{kN} statistic value is lower than the critical value. However, the p-value of the test shows that the probability of achieving a better distribution fitting than the Clayton copula is the lowest.

6. Results: correlation impact in the reliability assessment

When including correlation of variables in the limit state function evaluation, a change in the marginal distribution variance is expected. However, this amount of change is directly related to how sensitive the limit state function is with respect to the correlated variables.

6.1. Impact in the limit state marginal distribution

Left tailed correlated joint distributions will represent higher chances of sampling a lower conductivity value and low grain size $d70$ diameter in comparison with an uncorrelated joint distribution (Fig. 6). As a first guess, it is obvious to imply that bigger $d70$ grains are more difficult to drag and consequently structures founded in sand with greater diameters are less prone to suffer PE processes. Yet according to the Sellmeijer PE failure model, for lower

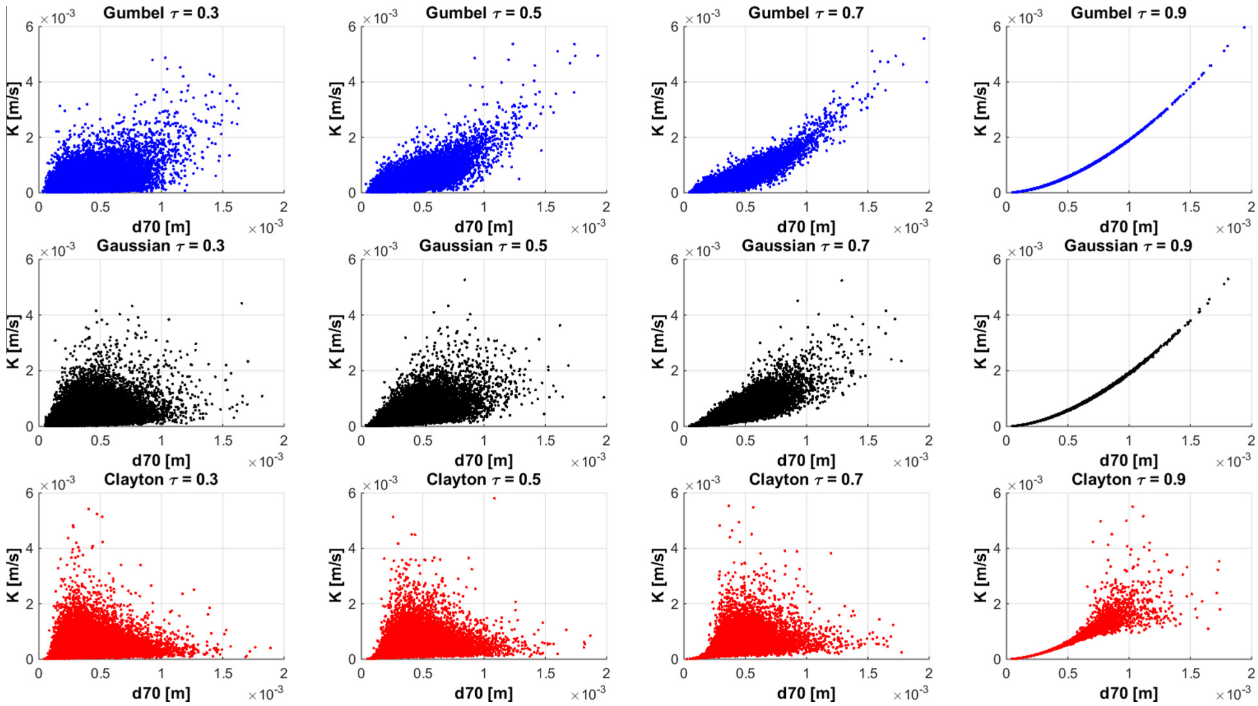

Fig. 6. Generated copula samples of *d70* representative grain size diameter versus hydraulic conductivity for different correlation degrees (Kendall's tau coefficient).

conductivity values a lower probability of structural failure is expected. As the rolling resistance of the grains is less probable to be exceeded for lower conductivity values (less flow inside the cavity), correlation will imply also that a higher frequency of smaller *d*70's (Fig. 6) is a sign of a safer structure. Therefore, correlation of these two variables will always represent a safer combination state of the structure comparted to the uncorrelated case. Note that the variance of the *K* parameter in Sellmeijer model has a much greater impact in the total variance of the PE limit state equation than the variance of the *d*70 according to the multivariate analysis results used to fine tune the Sellmeijer model [3].

This additional safety obtained from correlation is explained from a probabilistic point of view by a combined effect of a mean and variance reduction of the limit state marginal distribution. The first one is very mild and will make the structure less safe in average terms. However, the variance reduction has a much stronger positive effect making the distribution less spread (Table 6) and consequently safer. The overall effect is translated in a safer structure as it can be concluded from the reduction in the coefficient of variation no matter the copula correlation model used. It can also be observed from the variation coefficients for each copula family, how the marginal effect of the variance reduction is greater for low correlation values.

6.2. Impact in the tail located events

For failure, the zone bounded between $-\infty$ and 0 (tails zoom Fig. 7) is the one of more interest as its integral represents the probability of the flood defence to fail due to PE. The bars in Fig. 7 represent the histogram of the limit state function obtained from an uncorrelated stochastic estimation. The other three lines represent the obtained limit state probability density function for the three copula families with different degree of rank correlation.

From the tail zoom, it can be observed that the rate of change of probability of failure is also different. For example, the Clayton copula gives a higher probability of failure for a low rank correlation. It can also be observed that every time the induced correlation is increased, the frequency peak of the pdfs increases as well. This is expected as the reduction of the variance due to correlation will redistribute the area as the density function becomes steeper in the tails.

Even though there is a significant change in the total variance of the model, it cannot be concluded that this change is attributed solely to these two variables (*d*70 and *K*). The other variables included in the Sellmeijer limit state equation are also fluctuating along their uncertainty ranges and therefore the variance of the LSE will also be affected by them in a minor scale.

Table 6
Obtained standard deviation of Zp for different rank correlation copulas.

	GUMBEL			GAUSS			CLAYTON		
Tau	Mean [m]	Std [m]	COV [%]	Mean [m]	Std [m]	COV [%]	Mean [m]	Std [m]	COV [%]
0.00	4.710	2.490	53	4.710	2.490	53	4.710	2.490	53
0.30	4.661	2.308	50	4.656	2.284	49	4.650	2.240	48
0.50	4.628	2.191	47	4.628	2.166	47	4.620	2.128	46
0.70	4.606	2.095	45	4.602	2.075	45	4.605	2.061	45
0.90	4.592	2.031	44	4.590	2.029	44	4.599	2.030	44
0.95	4.589	2.024	44	4.589	2.021	44	4.594	2.028	44
0.99	4.591	2.019	44	4.591	2.022	44	4.589	2.023	44

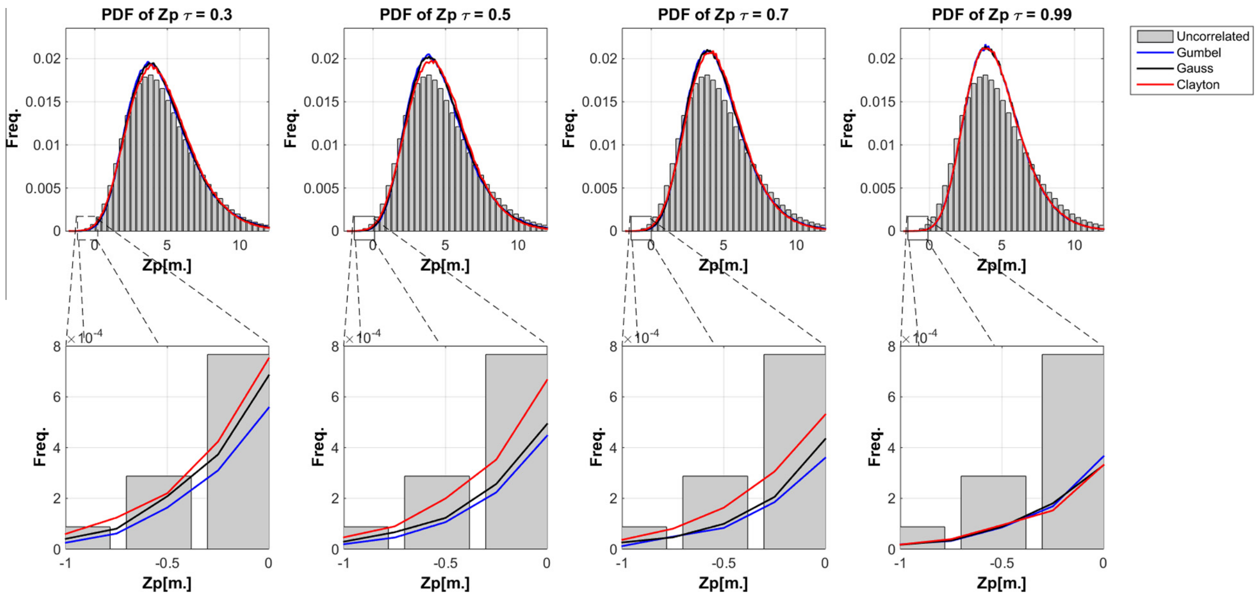

Fig. 7. Probability density functions of LSE for different copulas with different correlation.

6.3. Reliability Index (β)

In the Dutch regulation, the minimum return periods that flood defences need to have are clearly defined for each of the main flood defence systems inside the country. In the case of the Lekdijk, a minimum return period of 2000 years is required for river flood defences located in the dike ring 16. This value is equivalent to the total probability of all possible types of failure mechanisms combined for all the flood system components (e.g. representative cross sections). For each copula function, the total (PE/Uplift) failure probability was estimated by inducing different degrees of correlation between $d70$ and K. The results with their correspondent 95% confidence intervals (dashed lines) are presented in Fig. 8. These intervals are calculated from the different probability estimations from the Monte Carlo simulation for each degree of correlation.

For reliability assessment of structures it is common practice to refer to the Hasofer–Lind reliability index [29] to define the structure safety instead of the failure probability. This is a more comprehensible measure in terms of failure for designers who are guided by legislative design codes.

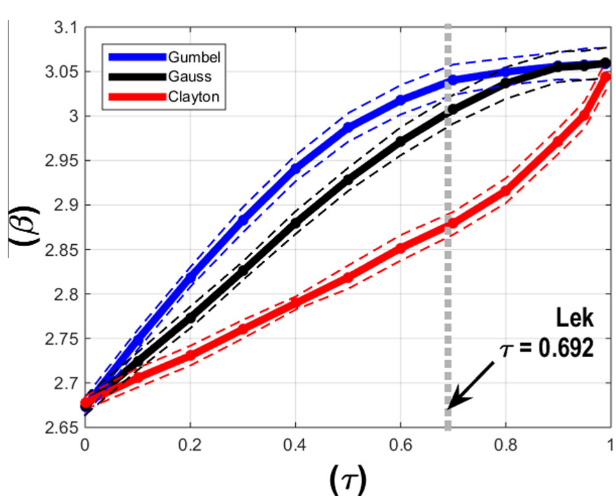

Fig. 8. PE/Uplift β index of the Lekdijk flood defence as a function of rank correlation coefficient (τ) with 95% confidence bounds.

The VNK2 study concluded that for the Lekdijk flood defence in Vianen. the return period associated solely to PE (PE/Uplift) in the specific location for this study is 280 years ($Pf = 3.5 \times 10^{-3}$/year). This is equivalent to a reliability index ($\boldsymbol{\beta}$) of 2.69. For our study, for the uncorrelated case ($\tau = 0$) the β reliability index obtained is equal to 2.675 (Fig. 8). This is equivalent to a return period of 267 years ($Pf = 3.74 \times 10^{-3}$/year). The small difference in reliability indexes between the VNK value and the present study can be attributed to the different LSE used (non revised versus revised Sellmeijer) and the more detailed field values used for representing the parameter uncertainties in this study. Also the analysis is done for a single cross section where the seepage length is select based on possible inflow and outflow points observed in the field. The dashed line represents the best correlation estimate between $d70$ and K (Table 4, $\tau = 0.692$) obtained from the soil sampled dataset (Fig. 8). In this case the β index will differ significantly with respect to the uncorrelated case no matter which copula model is chosen.

From three models, the Clayton copula is the most conservative as it will always result in lower reliability indexes with respect to the Gaussian and the Gumbel copulas. The explanation for this behaviour can be deduced from the combined inclusion of the $d70$ and K parameters in the FS term (Eq. (3)) of the Sellmeijer limit state function. The hydraulic conductivity is powered to 1/3 whereas the $d70$ is powered to 2/5. A smaller power of a small number will always be greater than a larger power of the same small number. This means that the denominator will be greater than the numerator in average. The Clayton copula induces higher tail dependency for left tail values (low permeability and low representative diameter) and consequently higher chances of sampling them at the same time. In conclusion, the FS term probability distribution of a Clayton copula is more likely to have a smaller mean value than the other two and consequently a lower critical head mean value (H_c) with respect to the other two.

7. Discussion

7.1. Research question 1

Is there any considerable correlation between the representative grain size (d70) and the hydraulic conductivity?

First, it's important to note that the hydraulic conductivity values for this study were calculated and not measured. For the

Lekdijk soil investigation, three different independent values were used to generate the marginal distribution of the hydraulic conductivity. More explicitly, the *d70, d60 and d10* values are obtained from the gradation curves of each sample which are measured by passing the collected soil trough different incremental sieves. While one sample contains all grain sizes and knowing that $d60$ and $d70$ are quite similar for well graded soils, they are still measured as independent values. Kozeny–Carman equation only uses $d10$ and $d60$ values for the hydraulic conductivity estimation. Hence, the correlation found between $d10$ and $d60$ is independent of the correlation found between K and $d70$. This can be observed in the Kendall's correlation values presented in Table 6. Nevertheless, the obtained values must also be analyzed by comparing their results with other studies in order to clarify if they are feasible.

The obtained values for mean and standard deviation from the sampled data for *d70* and K in the present study are in fully agreement with the ones used for the VNK study [24]. This suggest that implementing the Kozeny–Carman equation gives a good approximation value in terms of the order of magnitude and statistical characterization when compared to the measured values in the field. Consequently, it proves to be a powerful tool when no field measurements are available. The correlation between $d60$ and *d70* is normally quite high for granular soils and even more for this dataset in particular. Therefore when calculating K as a function of *d10* and *d60* and then estimating the correlation degree from the *d70* values obtained from the same samples, the expected degree of dependence is higher than calculated from independent measured values. Nevertheless, the actual Dutch methodology even accounts for this correlation degree by including an "alpha" coefficient in the method used for estimating the hydraulic conductivity [30]. The optimal situation would be if for each borehole sample, an in situ measurement of the local hydraulic conductivity was performed instead of calculating them via Kozeny–Carman model. Yet, this kind of data was not available for the present study. In addition, the aleatory uncertainty associated to the measurement of the hydraulic conductivity is large in comparison to the one associated to the $d70$. In the present reliability method this issue is addressed by the selection of the coefficient of a variation (COV, Table 3) for the two variables during the random sampling. The influence of this uncertainty is case dependent and cannot be generalized. Still, a sensitivity analysis of the coefficient of variance is a helpful tool to estimate its influence in the failure probability value for a fixed correlation degree. In contrast, the epistemic uncertainty derived from the simplification of assuming an equivalent hydraulic conductivity value for the whole aquifer (Sellmeijer limit state function) can be quantified by implementing the complete Sellmeijer numerical solution in which the hydraulic conductivity can be represented in more detail with a layered aquifer for example.

In order to determine if the high correlation value obtained between $d60$ and K in the present study ($\tau = 0.714$, Table 4) is a feasible value, the sample dataset from the research paper from Vienken and Dietrich [27] was used. Their study is presented with a dataset that contains $d60$ values with their correspondent in situ measurements of slug testing for hydraulic conductivity. The Kendall's rank correlation between $d60$ and K obtained for their dataset is $\tau = 0.5396$ ($\rho = 0.64$). These samples were obtained from a highly heterogeneous unconsolidated aquifer in Bitterfield, Germany. Measurements from less heterogeneous sands presented by van Beek et al. in [28] were also analyzed resulting in a $d60$ versus K correlation $\tau = 0.522$ ($\rho = 0.671$). This results show that the correlations obtained for the Lekdijk (Table 4, $d60$ versus K $\tau = 0.714$ and $d70$ versus K $\tau = 0.692$) is not impossible to obtain but is significantly different to the one obtained for this research case study. Both studies can also be assumed as highly reliable given the sampling techniques employed for the conductivity measurement. The Bitterfield dataset presented Vianken et al. is composed of in situ measurements only which makes it more realistic but the significance can be quite low as only 22 slug tests were done in 4 different boreholes. Then the same dataset contains 108 core samples which were characterized for this sampling campaign as well. The rank correlation coefficient obtained for this larger dataset between *d10* and *d60* is equal to $\tau = 0.608$. This value is closer to the one found for Lekdijk case (Table 4, $\tau = 0.786$). Despite the differences found for all Bitterfield, Dutch sands and Vianen aquifer, all show either left tail dependence (Section 5) and/or significant correlation.

7.2. Research question 2

How to select and validate a correlation bivariate model (Copula family) to correctly include its effect in the failure estimation due to PE?

In principle, if no data was available for deriving any quantitative conclusion, porous media theory suggests that the soils with smaller representative diameter values tend to present higher correlation in the resistance of water to flow inside them, Bear and Buchlin [31] explain how these two parameters are correlated according to Darcy's law, if inertial effects are included in its differential form. If so, a new quadratic term appears in the equation which expresses the exponent relation between the grain size and the flow velocity in presence of a porous matrix. The study presented by Chapuis [14] shows that most relevant empirical models for predicting conductivity from grain sizes are based on this exponent relation. In fact, it can also be concluded that models are majorly built based on values of the smaller representative particle sizes from the sand samples while their representative hydraulic conductivity values range between 1×10^{-2} and 1×10^{-14} m/s. Hence it can be concluded that the bivariate joint distribution should have a stronger left tail dependence (smaller grain sizes) which allows to discard the "Gumbel" copula function from the start. However, it was included in the present study as a measure of probabilistic bounding for understanding the consequences of assuming the wrong model. From the point of view of the obtained results of the different goodness of fit of the copula models, the Clayton performed better than the Gaussian in two of the three tests. Nevertheless the results do not reflect an extreme over performance between the two remaining models. In addition, all three models had difficulty representing the left tail dependence. Most likely to be originated from the fact that the empirical copula surface lacks information in the left tail corner area. Therefore the SAPE test is a more reliable performance method for assessing copula goodness of fit when low tail data coverage is observed.

From all the results presented in this study, the Clayton copula better represent the soil behaviour based on the literature and the available soil data. Nevertheless, the results of this study cannot be used to recommend one over the other one for the general case. Yet for the Lekdijk case study, the Clayton copula is recommended as a first choice if no additional information is available.

7.3. Research question 3

How important is the impact of correlation between d70 and K in the reliability assessment against PE?

The two main observed effects are the reduction of the variance and the "no effect" in the average marginal resistance value of the structure due to PE. As it was stated, the mean resistance of the structure is not affected at all. However, the variance reduction originated from the correlation and tail dependence represents a more reliable structure with respect to the uncorrelated case. This also be explained from the physical point of view as a greater conductivity allows the water to flow easier in the aquifer which makes it less safer for erosion to occur. However, for conductivity to increase there has to be a larger porosity which can only be achieved by the increasing the percentage of bigger grains in the grain distribution. These larger grains are more difficult to drag and therefore they make the aquifer less prone to be eroded. Consequently both extreme tail values are counteracting with each other, reducing their importance reflected in the estimated resistance value in the limit state marginal distribution evaluation.

The results for the Lekdijk field data suggest that *d*70 and *K* are correlated with $\tau = 0.692$ (Table 4) with left tail dependence which is better represented by a Clayton copula function. With these characteristics, the probability of having PE is 2.01×10^{-3}/year. This is equivalent to have a return period of 498 years or a $\beta = 2.877$. For assessing the impact in the safety assessment of PE, the error is estimated as the difference in failure probabilities between the resultant reliability indexes of any correlation assumption compared to these results. Therefore, three different scenarios can be derived:

*Scenario 1: d*70 and *K* are 100% correlated.

This is a highly unrealistic scenario but it can be used to understand what assumption impacts the failure estimation the most. The reliability index for any 100% correlated copula model is equal to 3.05 which is equivalent to a Pf of 1.14×10^{-3}/year. The error that one can incur by assuming the model to be completely correlated is 43% of overestimation. This means that the flood defence will be assumed as 43% safer.

*Scenario 2: d*70 and *K* are correlated in $\tau = 0.692$ while assuming a wrong copula model.

This scenario will be equivalent to choose the results obtained for a Gumbel copula correlated with the "correct" degree of dependence. The estimation of failure probability due to PE would be equal to 1.19×10^{-3}/year ($\beta = 3.04$). Therefore the assessment of the flood defence is 40.7% safer against PE.

*Scenario 3: d*70 and *K* are 100% uncorrelated (Actual assumption in these type of assessments).

In the actual procedure for PE failure estimation, *d*70 and *K* are assumed as 100% uncorrelated which means that the Lekdijk probability of failing because of PE is 3.71×10^{-3}/year ($\beta = 2.677$). That would represent an underestimation of 85% in the reliability of the flood defence towards PE. In other words, the flood defence is assumed as 85% less safe than what it can actually be. Such results may drive decision makers towards strengthening policies which might not be required. Or at least not until a more robust soil investigation is performed.

For the system that includes the Lekdijk (Dike ring 16) it was found that according to the VNK safety assessment results, 79.4% of the total failure probability could be attributed solely to the PE failure mechanism. According to the Dutch regulation, the Lekdijk defence must have at most a total failure probability of 5×10^{-4}/year. Therefore, the minimum allowable failure probability of the Lekdijk due to PE is 0.794^{*} (5×10^{-4}/year) which equals 3.97×10^{-4}/year ($\beta = 3.35$). For design purposes, the percentage of failure budget (maximum allowable probability of failure due to an specific failure mechanism) is even more strict as presented in [32]. In their study, the failure budget allocated for estimating the maximum allowable failure probability is 35% of the total failure probability (3.97×10^{-4}/year or $\beta = 3.58$). Consequently no matter the correlation degree, the Lekdijk must be strengthen as its estimated reliability index obtained by including the effect of correlation with the best fitting copula model is $\beta = 2.877$. Nevertheless, if correlation is included in the strengthening measures design (flood defence width, berm or sheet pile), a less expensive design will be obtained.

As a final remark, it was observed that this method can also be used for probability bounding. More kinds of copulas are available but for the present study, only the ones that showed dependence degrees in each of the tails and non tail dependence (Gaussian) were selected. It is also acknowledged by the authors that large amounts of samples are required in order to have a reliable estimate of the degree of correlation between *d70* and *K* and best copula model selection.

8. Conclusions and recommendations

- Strong evidence that significant correlation degree between *d70* and *K* is feasible is concluded from the results obtained from the Lekdijk case study and from the complementary datasets provided by other authors. Nevertheless, it is also acknowledged that for the present case study, the high correlation degree originates from the high intrinsic correlation structure present between the *d10* and *d70* values for the Lek dataset in particular.
- Based on the Lekdijk dataset, the Clayton copula model is capable of describing the bivariate behaviour and the physics of the soil in this location better than the Gumbel or the Gauss copula. Therefore, the Clayton copula or any other copula model with stronger left tail dependence is recommended as it can represent statistically and physically better the behaviour of the correlation between the *d*70 representative grain size and hydraulic conductivity.
- No matter which copula type is chosen, the correlation inclusion in the Sellmeijer LSE shows always a reduction in the failure probability of PE. This is a result of the monotonic variance reduction in the limit state marginal density function that results from joint effect of including *K* and *d70* as correlated in the LSE method.
- The present study showed that the assumption of uncorrelated variables implemented in the actual safety assessment for PE performed in The Netherlands might be conservative. The assumption of no correlation between *d*70 and *K* will be translated to higher probabilities of failure when the structure is assessed with the Sellmeijer revised equation.
- The assumption of any tail dependent model with any correlation degree will result in a smaller error in the estimation of the failure probability due to PE when compared to the 100% uncorrelated case.
- Correlation assessment for the Sellmeijer PE model is recommended as for the Lekdijk case study the failure probability is overestimated by a factor of 1.84 ($Pf_{correlated} = 2.01 \times 10^{-3}$/year, $Pf_{un\text{-}correlated} = 3.71 \times 10^{-3}$/year) when assuming the two parameters to be 100% uncorrelated.
- Solutions like wider cross sections, sheet piling or berms are common ways to cope with PE. For probabilistic design of flood defences, the omission of correlation can result in less cost effective designs by adding this kind of measures when they might be not needed. Therefore, a more detailed soil investigation is recommended in locations where failure is expected or in locations where the historical performance of the structure differs significantly with the expected limit state.
- All goodness of fit methods are derived from the interpolation of copula surfaces. Therefore the discretization criteria used for such interpolation should be optimized.

Acknowledgements

"This research is supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs." The authors would like to thank Raymond van der Meij and the flood risk department from Deltares, Doctor

Mireia Romaguera from ITC Twente Faculty, the anonymous reviewers of this journal and the Integral and sustainable design of Multi-functional flood defences program for the help and support.

References

[1] Jongejan R, Maaskant B, Ter Horst W, Havinga F, Roode N, Stefess H. The VNK2 project a fully probabilistic risk analysis for the major levee systems in The Netherlands. In: Takeuchi ACK, editor. 5th international conference on flood management (ICFM5). Tokyo, Japan: IAHS; 2013. p. 480.

[2] Sellmeijer JB, Koenders MA. A mathematical model for piping. Appl Math Model 1991;15:646–51.

[3] Sellmeijer H, de la Cruz JL, van Beek VM, Knoeff H. Fine-tuning of the backward erosion piping model through small-scale, medium-scale and IJkdijk experiments. Eur J Environ Civ Eng 2011;15:1139–54.

[4] van Noortwijk J, Vrouwenvelder A, Calle E, Slijkhuis K. Probability of dike failure due to uplifting and piping. In: Kafka GISaP, editor. The 10th European conference on safety and reliability (ESREL). Munich-Garching, German: Taylor & Francis; 1999. p. 1616.

[5] Vrouwenvelder ACWM. Spatial effects in reliability analysis of flood protection systems Second International Forum on Engineering Decision Making. Lake Louise Canada; 2006.

[6] Kanning W. The weakest link: spatial variability in the piping failure mechanism of dikes. The Netherlands: TU Delft; 2012.

[7] Diermanse FLM, Geerse CPM. Correlation models in flood risk analysis. Reliab Eng Syst Saf 2012;105:64–72.

[8] Rosas J, Lopez O, Missimer TM, Coulibaly KM, Dehwah AH, Sesler K, et al. Determination of hydraulic conductivity from grain-size distribution for different depositional environments. Ground Water 2014;52:399–413.

[9] Förster U, Van den Ham G, Calle E, Kruse G. Onderzoeksrapport: Zandmeevoerende Wellen. Report; 2012.

[10] Sklar A. Fonctions de Répartition à n Dimensions et Leurs Marges. Publications de l'Institut de Statistique de L'Université de Paris 1959;8:229–31.

[11] Nelsen RB. An introduction to copulas. 2nd ed. New York: Springer; 2006.

[12] Ojha CSP, Singh VP, Adrian DD. Assessment of the role of slit as a safety valve in failure of levees. Int J Sedim Res 2008;23:361–75.

[13] Koelewijn AR. SBW Hervalidatie piping E. Evaluatie 0,3d rekenregel. SBW. Netherlands: Deltares; 2009.

[14] Chapuis R. Predicting the saturated hydraulic conductivity of soils: a review. Bull Eng Geol Environ 2012;71:401–34.

[15] Mazzoleni M, Barontini S, Ranzi R, Brandimarte L. Innovative probabilistic methodology for evaluating the reliability of discrete levee reaches owing to piping. J Hydrol Eng 2014.

[16] Vukovic M, Soro A. Determination of hydraulic conductivity of porous media from grain-size composition. Littleton, Colo: Water Resources Publications; 1992.

[17] Biller B, Gunes Corlu C. Copula-based multivariate input modeling. Surv Oper Res Manage Sci 2012;17:69–84.

[18] Attoh-Okine NO. Pair-copulas in infrastructure multivariate dependence modeling. Constr Build Mater 2013;49:903–11.

[19] Vandenberghe S, Verhoest NEC, De Baets B. Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall. Water Resourc Res 2010;46.

[20] Genest C, Rémillard B, Beaudoin D. Goodness-of-fit tests for copulas: a review and a power study. Insurance: Math Econ 2009;44:199–213.

[21] Dobrić J, Schmid F. A goodness of fit test for copulas based on Rosenblatt's transformation. Comput Stat Data Anal 2007;51:4633–42.

[22] Scholz FW, Stephens MA. *K*-sample Anderson–Darling tests. J Am Stat Assoc 1987;82:918–24.

[23] Steenberghen HM, Lassing BL, Vrouwenvelder ACWM, Waarts P. Reliability analysis of flood defense systems HERON: Special issue – Risk based water defence. 2004;49.

[24] Vergouwe R, van den Berg MCJ, van der Scheer P. VNK2: Overstromingsrisico dijkring 16 Alblasserwaard en de Vijfheerenlanden. The Netherlands Rijkswaterstaat WVl; 2014.

[25] Rijkswaterstaat. HR2006 Hydraulische Randvoorwaarden primaire waterkeringen voor de derde toetsronde 2006–2011. Delft, The Netherlands: Ministerie van Verkeer en Waterstaat, Directoraat-Generaal Rijkswaterstaat; 2007.

[26] Schweckendiek T, Vrouwenvelder ACWM, Calle EOF. Updating piping reliability with field performance observations. Struct Saf 2014;47:13–23.

[27] Vienken T, Dietrich P. Field evaluation of methods for determining hydraulic conductivity from grain size data. J Hydrol 2011;400:58–71.

[28] van Beek VM, Knoeff H, Sellmeijer H. Observations on the process of backward erosion piping in small-, medium- and full-scale experiments. Eur J Environ Civ Eng 2011;15:1115–37.

[29] Tichý M. On the reliability measure. Struct Saf 1988;5:227–32.

[30] Calle E, Weijers J. Technisch rapport voor controle op het mechanisme piping bij rivierdijken. Deltares (Geodelft); 1994.

[31] Bear J, Buchlin JM, Von Karman Institute for Fluid Dynamics. Modelling and applications of transport phenomena in porous media. Dordrecht, Boston: Kluwer Academic Publishers; 1991.

[32] Jongejan R, Calle E. Calibrating semi-probabilistic safety assessments rules for flood defences. Georisk: Assess Manage Risk Eng Syst Geohazards 2013;7:88–98.