6th International Conference on Earthquake Geotechnical Engineering
1-4 November 2015
Christchurch, New Zealand

Spatial Correlation of Seismic Damage for Levee Systems

D.Y. Kwak[1], J.P. Stewart[1], and S.J. Brandenberg[1]

ABSTRACT

Seismic response of levees is typically computed for short segments within which levee geometry, soil conditions, and seismic demands can be assumed to be essentially constant. However, from a flood protection perspective, the performance of the system as a whole is critical because any failure within the system can lead to inundation. An assessment of system risk requires knowledge of individual segment performance in combination with the spatial correlation of damage among segments. We present a correlation model for damage states among levee segments that can be used in combination with segment fragility models and correlated demand functions to assess levee system risk. We compute the autocorrelation of damage states as a function of separation distance for the Shinano River levee system in Japan, based on observational data from two shallow crustal earthquakes. Levee segment damage is found to be spatially correlated up to 4 km separation for small to moderate levels of damage but only to 1 km for severe damage.

Introduction

A levee system is comprised of earth embankments that protect a particular area from flooding. Levee segments are connected in series, so failure of one segment exposes the protected region to possible inundation, and hence comprises system failure. Accordingly, the levee system fragility problem involves analysis of the probability of whether at least one levee segment in the series exceeds a specific damage state, *DS*. The solution of this problem depends strongly on the system length (i.e., number of reaches) and correlation of damage among segments. System probability of failure increases as the number of segments increases and damage correlation decreases.

To illustrate the importance of damage correlation, consider two extreme cases: perfectly correlated and statistically independent. Perfect correlation of damage occurrence requires that all segments in a system are damaged, or not damaged, simultaneously. In this case, the segment having the highest failure probability (highest fragility) will control the system fragility. If the system failure probability is denoted $P(F_S)$ and the failure probability of segment i is $P(f_i)$, we have:

$$P\left(F_S\right) = \max\left(P\left(f_i\right)\right) \tag{1}$$

On the other hand, statistical independence requires $P(F_S)$ to be computed as the complement of system survival, which in turn is the product of each individual segment surviving:

[1]University of California, Los Angeles, CA, USA, kwakdongyoup@gmail.com

$$P\left(F_S\right)=1-\prod_{i=1}^{n}\left(1-P\left(f_i\right)\right) \qquad (2)$$

where n is the total number of segments. Perfect correlation and statistical independence comprise extremes known as uni-modal bounds for a series system (Ang and Tang, 2007). Actual system fragilities are between these extremes:

$$\max\left(P\left(f_i\right)\right)\leq P\left(F_S\right)\leq 1-\prod_{i=1}^{n}\left(1-P\left(f_i\right)\right) \qquad (3)$$

The range of failure probabilities provided by Eq. (3) is often wide. For example, a system of 100 segments each with $P(f_i)$ = 0.05 will have $P(F_S)$ = 0.05 for perfect correlation, and $P(F_S)$ = 0.994 for statistical independence. The actual value of $P(F_S)$ within this large range depends strongly on between-segment correlations.

Previous studies have addressed the system fragility problem for relatively simplified conditions. USACE (2011) and Wolff (2008) compute $P(F_S)$ by dividing the levee system into "reaches" with a characteristic length (typically 100 to 300 m) within which the correlation of damage is assumed perfect, whereas correlations between reaches are taken as zero. In these applications, a "reach" is a length of levee judged to have adequately similar geometry, soil conditions, and loading conditions that the reach can be represented by analysis of a single cross-section. For each reach, $P(f_i)$ is evaluated from geotechnical engineering models and $P(F_S)$ is then computed using Eq. (2).

On the other hand, the Delta Risk Management Strategy (URS and JBA, 2008) computed $P(F_S)$ by summing weighted values of $P(f_i)$, where the weights represent the probability of each reach being the "weakest link". The weight for a particular reach is proportional to $P(f_i)$, and the weights sum to unity. The resulting value of $P(F_S)$ is similar to assuming perfect correlation of damage among reaches, although $P(F_S)$ will actually be less than or equal to the maximum value of $P(f_i)$ using this approach. To account for levee system length, $P(F_S)$ is then multiplied by a correction factor ranging from 0.7 to 1.7, developed from empirical observations of flood events in the Delta.

The principal limitation of these methods is that they incorporate the spatial correlations of damage in an arbitrary manner without empirical justification. Defining the characteristic length of a reach is difficult without formally considering spatial correlation of demand and resistance. Furthermore, the characteristic length may be different for earthquakes than for floods due to the spatial variability of the loading condition, and because soil properties that resist floods (hydraulic conductivity, erodibility) are different from those that resist earthquakes (liquefaction susceptibility, undrained shear strength). Furthermore, high water demands posed by floods are likely more spatially correlated than ground shaking demands. A more robust solution that accounts for spatial correlation is therefore needed.

In this manuscript, we present a correlation coefficient model of damage states using a dataset from Japan. We first briefly introduce segment fragility functions that define segment damage probability as a function of ground shaking, geological conditions, and groundwater elevation (Kwak et al., 2015). A "segment" in the context of our study is 50 m long. Like a "reach" in prior work, our segments are assumed to be described by a single fragility model (akin to perfect within-segment correlation). What makes the present work distinct, however, is that we quantify spatial correlations of damage states for adjacent segments (prior work assumed zero between-reach correlation). We quantify segment damage using Boolean variables representing damage states, define variables related to the correlation of damage states, and develop estimates of spatial correlation using auto-correlation analyses of levee damage data. To our knowledge, this is the first quantification of spatial correlation of seismic damage derived from field performance data.

Description of Japanese Case Studies

Kwak et al. (2015) developed fragility functions for 50 m long levee segments in which probability of exceeding a damage threshold is expressed as a function of peak ground velocity (*PGV*), surface geology, and groundwater elevation. The data set used for fragility development, illustrated in Figure 1, consists of detailed damage observations along the Shinano River levee system in Japan following earthquakes in 2004 and 2007. There are three sub-rivers in this system, denoted SH1, SH2, for the downstream and upstream portions of the Shinano River, and UO for the Uono River that is a tributary of the Shinano River. The system consists of 3318 segments.

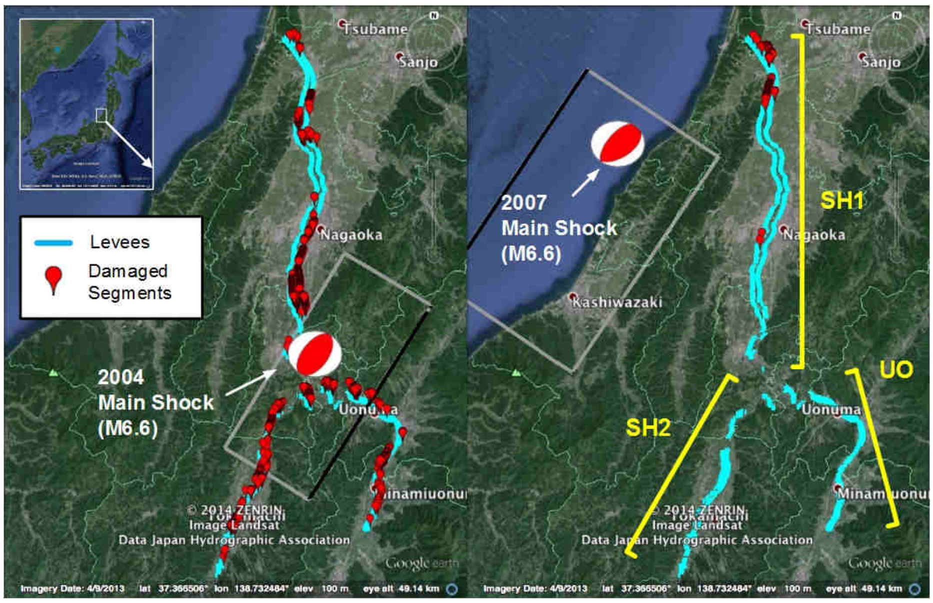

Figure 1. Levees along the Shinano River system (SH1, SH2, and UO) on Google Earth map. Locations are shown of levee damage, epicenters (shown with the moment tensor solution), and finite fault planes (black line at top) (after Kwak et al., 2015).

Crack depth, crack width, and crest subsidence was measured for each segment, and a discrete damage state was assigned based on these observations. Damage states range from zero for no damage to four for severe damage (e.g., levee collapse). Figure 2 shows seismic levee fragility expressed as the probability of exceeding a specific *DS* versus *PGV*, conditioned on surface geomorphology (G_N) and relative ground water elevation (D_W) (defined below). The fragility functions are statistically lower than average for G_N category 1 (relatively firm materials in mountainous regions or gravel terrace deposits). Parameter D_W in Figure 2 is the relative depth of the ground water elevation to the levee base at the time of the earthquake, therefore high D_W corresponds to shallow ground water and presumably increased liquefaction risk. Figure 2 shows that higher groundwater elevation causes increased fragility.

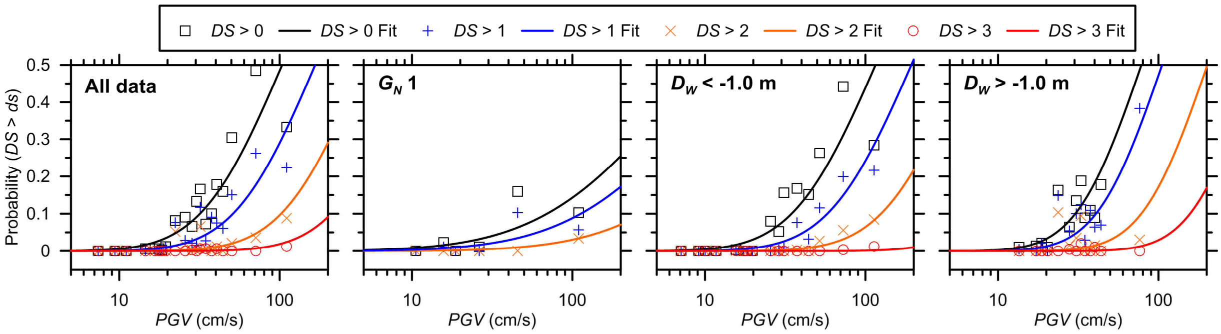

Figure 2. Probabilities of exceeding damage states for (a) all of the segments combined, (b) segments with geomorphic (G_N) category 1 corresponding to mountain or gravel terrace deposits, (c) groundwater depth relative to levee base greater than 1m, and (d) groundwater depth less than 1 m below levee base (after Kwak et al., 2015).

Correlation Coefficient of Damage States

The damage states assigned by Kwak et al. (2015) can be represented as Boolean variables (i.e., zero or one) depending on whether each segment exceeded a particular damage state. If we define 'survival' as the damage for segment i (s_i) falling below a particular DS, then, by definition, the correlation coefficient of survival ($\rho_{DS,s}$) between segments i and j is (Kutner et al., 2004):

$$\rho_{DS,s} = \mathrm{cov}\left(s_i s_j\right) / \sqrt{\mathrm{var}\left(s_i\right)\mathrm{var}\left(s_j\right)} \tag{4}$$

where $\mathrm{cov}(s_i s_j)$ is the covariance of survival events for segments i and j, and $\mathrm{var}(s_i)$ is the variance of s_i. Based on the definition of covariance (Ang and Tang, 2007), the $\mathrm{cov}(s_i s_j)$ can be expanded as:

$$\mathrm{cov}\left(s_i s_j\right) = E\left[\left(s_i - \mu_{s,i}\right)\left(s_j - \mu_{s,j}\right)\right] = E\left(s_i s_j\right) - E\left(\mu_{s,i} s_j\right) - E\left(\mu_{s,j} s_i\right) + E\left(\mu_{s,i}\mu_{s,j}\right) = E\left(s_i s_j\right) - \mu_{s,i}\mu_{s,j} \tag{5}$$

where $\mu_{s,i}$ represents the mean of s_i, which is analogous to the expected value $E(s_i)$ or the probability $P(s_i)$. Note that the last three terms in the third expression in Eq. (5) are all equal to $\mu_{s,i}\mu_{s,j}$. The $\mathrm{var}(s_i)$ can be expanded as:

$$\mathrm{var}\left(s_i\right) = E\left(s_i^2 - \mu_{s,i}^2\right) = E\left(s_i^2\right) - E\left(\mu_{s,i}^2\right) = \mu_{s,i} - \mu_{s,i}^2 \tag{6}$$

Since s_i is a Boolean variable, s_i^2 is equal to s_i resulting in $E(s_i^2) = E(s_i)$. Substituting $\mathrm{cov}(s_i s_j)$, var(s_i), and var(s_j) from Eqs (5)-(6) into Eq. (4), we obtain:

$$\rho_{DS,s} = \frac{E\left(s_i s_j\right) - \mu_{s,i}\mu_{s,j}}{\sqrt{\mu_{s,i}\left(1-\mu_{s,i}\right)\mu_{s,j}\left(1-\mu_{s,j}\right)}} = \frac{P\left(s_i \cap s_j\right) - P\left(s_i\right)P\left(s_j\right)}{\sqrt{P\left(s_i\right)\left(1-P\left(s_i\right)\right)P\left(s_j\right)\left(1-P\left(s_j\right)\right)}} \tag{7}$$

where $P(s_i \cap s_j)$ is the probability of the intersection of s_i and s_j. Note that Eq. (7) can be re-written in terms of failure states (f_i; one for damage and zero for no damage):

$$\begin{aligned}\rho_{DS,s} &= \frac{E\left(s_i s_j\right) - \mu_{s,i}\mu_{s,j}}{\sqrt{\mu_{s,i}\left(1-\mu_{s,i}\right)\mu_{s,j}\left(1-\mu_{s,j}\right)}} = \frac{E\left[\left(1-f_i\right)\left(1-f_j\right)\right] - \left(1-\mu_{f,i}\right)\left(1-\mu_{f,j}\right)}{\sqrt{\left(1-\mu_{f,i}\right)\left(1-\left(1-\mu_{f,i}\right)\right)\left(1-\mu_{f,j}\right)\left(1-\left(1-\mu_{f,j}\right)\right)}} \\ &= \frac{1-E\left(f_i\right)-E\left(f_j\right)+E\left(f_i f_j\right)-1+\mu_{f,i}+\mu_{f,j}-\mu_{f,i}\mu_{f,j}}{\sqrt{\mu_{f,i}\left(1-\mu_{f,i}\right)\mu_{f,j}\left(1-\mu_{f,j}\right)}} \\ &= \frac{E\left(f_i f_j\right) - \mu_{f,i}\mu_{f,j}}{\sqrt{\mu_{f,i}\left(1-\mu_{f,i}\right)\mu_{f,j}\left(1-\mu_{f,j}\right)}} = \rho_{DS,f}\end{aligned} \tag{8}$$

where $\mu_{f,i}$ is the mean of f_i, which is equivalent to $E(f_i)$. Eq. (8) shows that the correlation coefficients for survival and failure Boolean damage states are equivalent. Accordingly, we drop the '*s*' and '*f*' from the subscripts and refer to the correlation coefficient of damage states as ρ_{DS}.

Direct computation of ρ_{DS} would require observations of the same levee system exposed to many earthquakes. The probability of survival for a segment *i* [i.e., $P(s_i)$] is the mean of s_i from many samples, whose reliability is highly dependent on the number of samples, which must be from events that produce shaking that is strong enough to have the potential for causing damage. The joint distribution $P(s_i \cap s_j)$, which is the probability of survival of both segments *i* and *j*, similarly requires a large number of samples for a reliable estimate. In practice, data will seldom be available with which to compute ρ_{DS} from observed damage states. In the following section, we present an autocorrelation coefficient approach that relies on a large volume of data for a few events. This approach is investigated as a means by which to approximate ρ_{DS}.

Autocorrelation Coefficient of Damage States

Autocorrelation represents the cross-correlation between a data vector and an offset, or lagged, version of the same vector in which the values in the vector appear in the same order but are shifted by a prescribed distance. In our case the lag represents a separation distance between levee segments. The correlation is computed between the original and shifted data vectors, and the process is repeated for all possible shifts. The resulting correlation values are then plotted as a function of the lag distance to develop an autocorrelation function. The autocorrelation function is equal to the damage state correlation if damage state correlation is stationary in space (i.e., if the correlation of damage states is a function only of spatial separation distance). We lack adequate observations to empirically verify whether damage state correlation is stationary, and therefore adopt the autocorrelation terminology ρ_{ac}.

Values of ρ_{ac} were computed for DS > 0, 1, and 2 using observations along the Shinano River system. Levee damage occurred along SH1 in the 2004 and 2007 earthquakes and along SH2 and UO only in the 2004 earthquake. Levees are discontinuous along the rivers in some cases, and autocorrelation is computed by combining all continuous sections greater than 0.5 km in cumulative length. For a given separation distance x, a continuous section of levee only contributes data to the autocorrelation calculation if its cumulative length exceeds x. Accordingly, for large x, only a subset of the data having long continuous stretches of levee are used. The longest stretches contain 798 segments (39.9 km) for SH1, 221 segments (11.05 km) for SH2, and 362 (18.1 km) segments for UO. For DS > 0 and DS > 1, this provides 1381 segments per earthquake (or 2762 for both earthquakes) for evaluation of model parameters in Eq. (9). For DS > 2, there are 798 segments (SH1 only) since no occurrences of damage in this range occurred for SH2 and UO during the 2007 earthquake. The data are inadequate to compute ρ_{ac} for DS > 3.

The resulting values of ρ_{ac} are plotted versus separation distance in Figure 3. The values of ρ_{ac} are near unity at a separation distance near zero, and decrease approximately exponentially with separation distance. Variation of ρ_{ac} with distance is regressed as follows:

$$\rho_{ac} = \begin{cases} 1 & \text{if} \quad x = 0 \\ c_{DS} \exp\left(-3 \times x / \alpha_{DS}\right) + \varepsilon_x & \text{if} \quad x > 0 \end{cases} \tag{9}$$

where x is the lagged distance, c_{DS} and α_{DS} are regression coefficients, and ε_x is an error term. The regression coefficient α_{DS} is equal to the 'range' in a semi-variogram, which is the lag where ρ_{ac} becomes practically zero. Eq. (9) is divided into different equations for x = 0 and x > 0 to facilitate a more accurate fit to the data than would be afforded by forcing ρ_{ac} to be unity at x = 0 in the functional form for the regression (i.e., by making c_{DS} = 1). Accordingly, there is a step from 1.0 to c_{DS} as x becomes finite. The coefficients were regressed using separation distances in the range $x \leq 1.0$ km in order to best fit the data in that critical range.

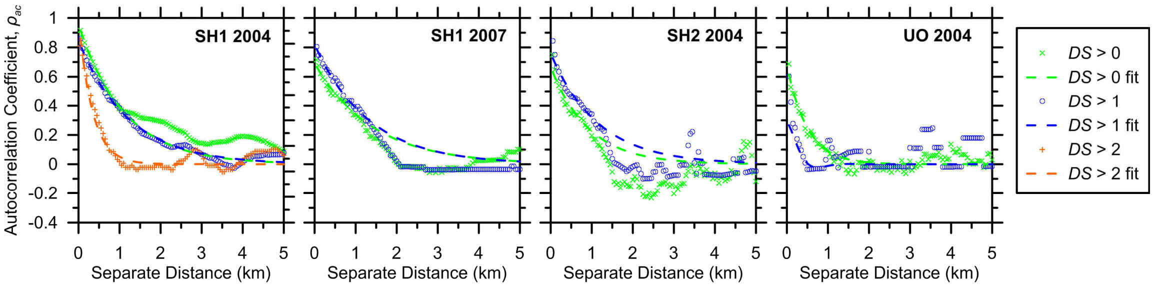

Figure 3. Autocorrelation coefficients of damage states for levee systems for the SH1, SH2, and UO rivers from the 2004 and 2007 Niigata earthquakes.

The correlation model for DS > 0 and DS > 1 is well constrained by the data because the results in Figure 3 are very similar for different river systems and different earthquakes. Figure 4 synthesizes the data from Figure 3 for DS > 0 and 1, and the resulting regression coefficients are c_{DS} = 0.77 and α_{DS} = 3.7 km. Regression for the DS > 2 case results in c_{DS} = 0.8 and α_{DS} = 1 km, though this case is not constrained by as many observations as the *DS* > 0 and 1 cases. The recommended regression coefficients are summarized in Table 1.

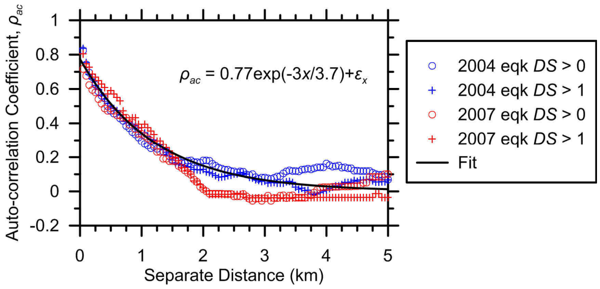

Figure 4. Auto-correlation coefficient ρ_{ac} for *DS* > 0 and 1 damage states combining all river systems for 2004 and 2007 earthquakes and a regressed fit line considering combined data set.

Table 1. Recommended regression coefficients for defining correlation structure.

Damage State	c_{DS}	α_{DS} (km)
DS > 0 and 1	0.77	3.7
DS > 2	0.8	1

To explore the influence of the number of observations on the accuracy of the computed values of ρ_{ac}, we conducted a mathematical study in which ρ_{DS} is known [ρ_{DS}=0.77exp(-3*x*/3.7)], and random Boolean variables are computed for 798 and 2762 segments. Correlation among random samples was achieved using Cholesky decomposition (Baecher and Christian, 2003). The process is repeated 5 times to characterize the repeatability of the random process. The results are illustrated in Figure 5, and indicate that errors in the computed values of ρ_{ac} become smaller as the number of segments increases, and that correlation tends to be under-predicted when the number of segments is small. The number of segments available for the Shinano River system is total 2762 for *DS* > 0 and 1, and these cases are therefore well-constrained, particularly in the region for $x < 1$ km for which we computed the regression coefficients. However, the number of segments for *DS* > 2 is 798, and correlation may be under-predicted for that case. More observations will be required in the future to ascertain the significance of this potential error.

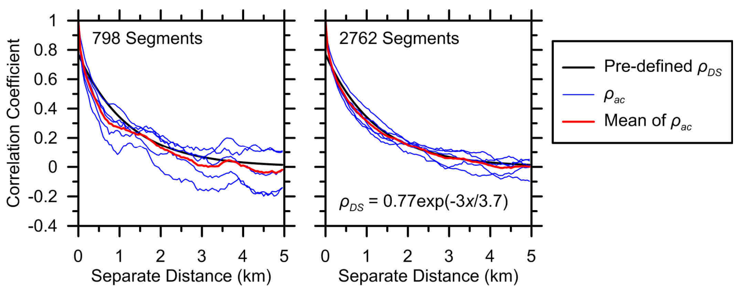

Figure 5. Comparison of pre-defined ρ_{DS} with five randomly generated realizations of auto-correlation ρ_{ac} with their mean.

Conclusions

Spatial correlation of levee damage is important for system fragility, but has not been considered in a rigorous, defensible manner in previous studies. We believe this oversight has resulted in part from the difficulty associated with quantifying the correlation structure for levee damage. This study quantifies autocorrelation of damage observations from a levee system in Niigata, Japan that was strongly shaken by two shallow crustal earthquakes. We regress autocorrelation data of discrete levee damage states and find correlation for separation distances between segments of less than one to four km, depending on the considered level of damage severity.

Formally accounting for the correlation structure is an improvement over arbitrarily selecting a characteristic length for a levee reach. While we recognize that the results presented herein are specific to a particular region in Japan, they nonetheless facilitate more rational consideration of correlation structure in future analyses of seismic levee system risk.

Acknowledgments

This work was supported by California Department of Water and Resource (CDWR) under contract number 4600008849. These funding sources are gratefully acknowledged.

References

Ang AHS, Tang HT. *Probability Concepts in Engineering: Emphasis on Applications in Civil & Environmental Engineering*. 2nd ed. Wiley: New York, 2007.

Baecher GB, Christian JT. *Reliability and Statistics in Geotechnical Engineering*. John Wiley & Sons Ltd: England, 2003.

Kutner M, Christopher JN, Neter J, Li W. *Applied Linear Statistical Models*. 5th ed. McGraw-Hill: New York, 2004.

Kwak DY, Stewart JP, Brandenberg SJ, Mikami A. Characterization of Seismic Levee Fragility Using Field Performance Data. *Earthquake Spectra* 2015. doi:10.1193/030414EQS035M.

U.S. Army Corps of Engineers (USACE). *Design of I-Walls, Appendix I: Length Effects in Levee System Reliability*. Washington DC, 2011.

URS Corporation (URS), Jack R. Benjamin & Associates Inc. (JBA). Levee Vulnerability, In *Technical Memorandum of Delta Risk Management Strategy Final Phase 1*. Oakland, 2008.

Wolff TF. Reliability of Levee Systems, In *Reliability-Based Design in Geotechnical Engineering*, edited by K. Phoon. Taylor & Francis: New York, 2008.