Local Probabilistic Sensitivity Measures for Comparing FORM and Monte Carlo Calculations Illustrated with Dike Ring Reliability Calculations

Roger M. Cooke [a,1] Jan M. van Noortwijk [b]

[a] *Department of Mathematics, P.O. Box 5031, 2600 GA Delft, the Netherlands.*

[b] *HKV Consultants, P.O. Box 2120, 8203 AC Lelystad, the Netherlands.*

Abstract

We define local probabilistic sensitivity measures as proportional to $\partial E(X_i|Z = z)/\partial z$, where Z is a function of random variables $X_1, \ldots, X_n$. These measures are local in that they depend only on the neighborhood of $Z = z$, but unlike other local sensitivity measures, the local probabilistic sensitivity of X_i does not depend on values of other input variables. For the independent linear normal model, or indeed for any model for which X_i has linear regression on Z, the above measure equals $\sigma_{X_i}\rho(Z, X_i)/\sigma_Z$. When linear regression does not hold, the new sensitivity measures can be compared with the correlation coefficients to indicate degree of departure from linearity.

We say that Z is probabilistically dissonant in X_i at $Z = z$ if Z is increasing (decreasing) in X_i at z, but probabilistically decreasing (increasing) at z. Probabilisitic dissonance is rather common in complicated models. The new measures are able to pick up this probabilistic dissonance.

These notions are illustrated with data from an ongoing uncertainty analysis of dike ring reliability.

Keywords: Local sensitivity measures; First Order Reliability Methods; Linear models; Dike ring reliability; Uncertainty analysis.
Mathematics Subject Classification: 60E05, 60G50, 62E25, 62J05, 90B25.

[1] Corresponding author: Phone: +31 15 2782548; Fax: +31 15 2787255; Email: r.m.cooke@twi.tudelft.nl.

Preprint submitted to Elsevier Science *10 September 1998*

1 Introduction

The Dutch government is currently undertaking an extensive uncertainty analysis of dike ring reliability. The reliability of dike section i is expressed in terms of the limit-state function

$$Z_i = \text{model factor strength}_i \times \text{Strength}_i(X_{i1}, \ldots X_{in}) - \text{model factor load}_i \times \text{Load}_i(X_{i1}, \ldots X_{in}),$$

where strength and load are functions of uncertain parameters (global and section-specific) $X_{i1}, \ldots, X_{in}$. The model factors are used to express uncertainty in the modelling of strength and load. For dike section i the strength is proportional to

$$\frac{\text{grass factor}_i^3 \times \text{roughness}^{0.25}}{[1 + 0.8 * \log(\text{storm length})]^3 \times \tan(\text{inner slope}_i)^{0.75}}.$$

Roughness and storm length are global variables, although uncertain, which take the same value for each dike section. The grass factor and inner angle are specific for each dike section. In the preliminary 'in house' analysis, the grass factor was treated as a constant. Evidently, Z_i is increasing in model factor strength$_i$ and in roughness, and decreasing in storm length.

The load is a complicated model depending on the river Rhine discharge, the North-Sea water level, wind, and wave attack. This model does not lend itself for presentation here (see Van Der Meer & Janssen [7]).

The limit-state function for a dike ring consisting of k dike sections is

$$Z = \min\{Z_1, \ldots, Z_k\}.$$

The example discussed below involves one failure mechanism, overtopping, and some 300 uncertain parameters. Since all dike sections are exposed to the same sea water levels, the same Rhine discharge and the same winds, there are significant dependencies in the reliabilities of different dike sections. Monte Carlo (MC) and First Order Reliability Methods (FORM) have been used with an 'in house' assessment of uncertainty for the purpose of comparing the dependency modelling and comparing the relative importances of various input parameters.

In the next section, we review the FORM approach to identifying important parameters. In Section 3, we discuss the assumptions underlying the FORM approach in the present case. Section 4 illustrates a phenomenon called 'local probabilistic dissonance'. Section 5 develops a 'Local Probabilistic Sensitivity Measure' (LPSM), Section 6 presents some preliminary results, and a final section gathers conclusions.

2 FORM

Suppose $Z(X_1, \ldots, X_n)$ is a 'deterministic' function of the random vector $\mathbf{X} = (X_1, \ldots, X_n)$. Assuming that Z is analytic, we can linearize it about some point $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$:

$$Z(\mathbf{X}) = Z(\mathbf{x}^*) + \textstyle\sum_{i=1}^n (X_i - x_i^*)\partial_i Z(\mathbf{x}^*) + \ldots \text{HOT (higher order terms)},$$

where ∂_i denotes $\partial/\partial x_i$. The point $\mathbf{x}^*$ is chosen as the 'design point', that is the point with greatest probability density satisfying $Z(\mathbf{x}^*) = 0$.

Let μ_i and σ_i denote the mean and standard deviation of X_i, respectively. Neglecting the HOT's, we have

$$\begin{aligned} Z(\mathbf{X}) &\sim Z(\mathbf{x}^*) + \textstyle\sum_{i=1}^n (X_i - x_i^*)\partial_i Z(\mathbf{x}^*), \\ E(Z) &\sim Z(\mathbf{x}^*) + \textstyle\sum_{i=1}^n (\mu_i - x_i^*)\partial_i Z(\mathbf{x}^*), \\ \mathrm{Var}\,(Z) &\sim \textstyle\sum_{i=1}^n \sigma_i^2(\partial_i Z)^2 + \sum_{i=1}^n \sum_{j=1, j\neq i}^n \partial_i(Z)\partial_j(Z)\,\mathrm{Cov}\,(X_i, X_j), \end{aligned} \tag{1}$$

and if the X_i's are all independent, $\mathrm{Var}\,(Z) \sim \sum_{i=1}^n \sigma_i^2(\partial_i Z)^2$. Now suppose that Z is indeed linear and the X_i's are independent. Then

$$\mathrm{Cov}\,(Z, X_i) = \rho(Z, X_i)\sigma_Z\sigma_i = \partial_i Z\,\mathrm{Cov}\,(X_i, X_i) = \sigma_i^2\partial_i Z$$

so that

$$\rho(Z, X_i)\sigma_Z/\sigma_i = \partial_i Z(\mathbf{x}^*). \tag{2}$$

Note that the LHS involves 'global' parameters, whereas the RHS depends on the design point x_i^*. It is characteristic of linear models that these global and local concepts coincide. $\rho(Z, X_i)$ is taken to represent the importance of X_i for Z. Note that in the FORM model this has both a global and a local interpretation. Note also that the above makes no assumptions regarding the distributions of the X_i except the existence of the first two moments. In order to estimate the probability of $\{Z = 0\}$, the FORM method assumes that the X_i have been transformed to standard normal variables.

Continuing,

$$\sigma_Z^2 = \textstyle\sum_{i=1}^n \sigma_i^2(\partial_i z)^2 = \sum_{i=1}^n \rho^2(Z, X_i)\sigma_Z^2$$

or

$$R^2 = \textstyle\sum_{i=1}^n \rho^2(Z, X_i) = 1.$$

In the terminology of linear models $R^2 = \sum_{i=1}^n \rho^2(Z, X_i)$ is the percentage of the variance of Z that is explained by the linear model (1). If R^2 is less than one, this may be caused *either* by dependencies in the X_i's or by contributions from HOT's in (1).

Several authors (see McKay [3]) propose the correlation ratio CR_i to replace $\rho^2(Z, X_i)$ for cases when Z is not linear:

$$CR_i = \text{Var}\,(E(Z|x_i))/\,\text{Var}\,(Z).$$

Note that CR_i generalizes the global interpretation of importance in (2), but not the local interpretation. Moreover, CR_i cannot be computed in a straightforward way by MC methods.

3 How linear is Z_i?

For dike section i, Z_i is computed from a model involving many cut-offs, edges, and non-linearities. Nonetheless, because of its complexity, the question "how linear is Z_i?" cannot be answered by inspection. Using the MC calculation for section i, we can assess the linearity of Z_i simply by computing R^2. We find $R^2 = 0.977$, with the largest contribution 0.903 coming from one variable (a 'strength model factor'). This does not correspond at all to the partial derivatives computed at the design point, which were dominated by the North-Sea level. On the other hand, performing a conditional R^2 near the region of greatest failure probability we find that (i) the conditional correlations are sensitive to *how* the conditionalization is performed, and (ii) the conditional R^2 is quite small, though still dominated by the globally dominant parameter. This strongly suggests that Z is globally linear, as it is dominated by one variable, but in the region of interest, $Z \sim 0$, which has very low probability mass, Z's behavior is highly non-linear. For this reason it is difficult to interpret the FORM importance parameters in terms of (conditional) correlations from a MC calculation.

4 Local Probabilistic Dissonance

The present data set affords many examples of a curious behavior which underscores the need for local probabilistic sensitivity concepts. It may be the case that Z is strictly increasing in some variable x_i, but for some value z, the conditional distribution $X_i|(Z = z)$ is stochastically *decreasing* in z, in the sense that for all x

$$\Pr\{X_i > x|Z = z + \delta\} < \Pr\{X_i > x|Z = z\}$$

for suitably small positive δ. In this case we speak of local probabilistic dissonance.

A very simple example illustrates how this may arise. Let $Z = X + Y$ with X and Y independent. Let X be uniformly distributed on $[0,1]$ and let Y be uniformly distributed on $[0,1] \cup (2,3]$. Then Z is concentrated on $[0,4]$. As $Z \to 2$ from below, the conditional distribution of $X|Z$ becomes concentrated at 1; however as $Z \to 2$ from above, the conditional distribution of $X|Z$ becomes concentrated at 0. Hence for $2 > \delta > 0$

$$\Pr\{X > x | Z = 2 + \delta\} < \Pr\{X > x | Z = 2\}.$$

In such cases, local sensitivity measures, like a partial derivative; indeed the partial derivatives of Z with respect to X and Y are equal. Global measures will not reveal the local probabilistic influence in the neighborhood of $Z = 2$.

We can illustrate this phenomenon with percentile cobweb plots. The percentile cobweb plot of Figure 1 shows the joint distribution, in percentiles, of Z ('relia') and 10 explanatory variables. From left to right the variables are: roughness ('rough'), storm length ('storm'), model factors for load, strength, significant wave period, significant wave height, and local water level ('mload',

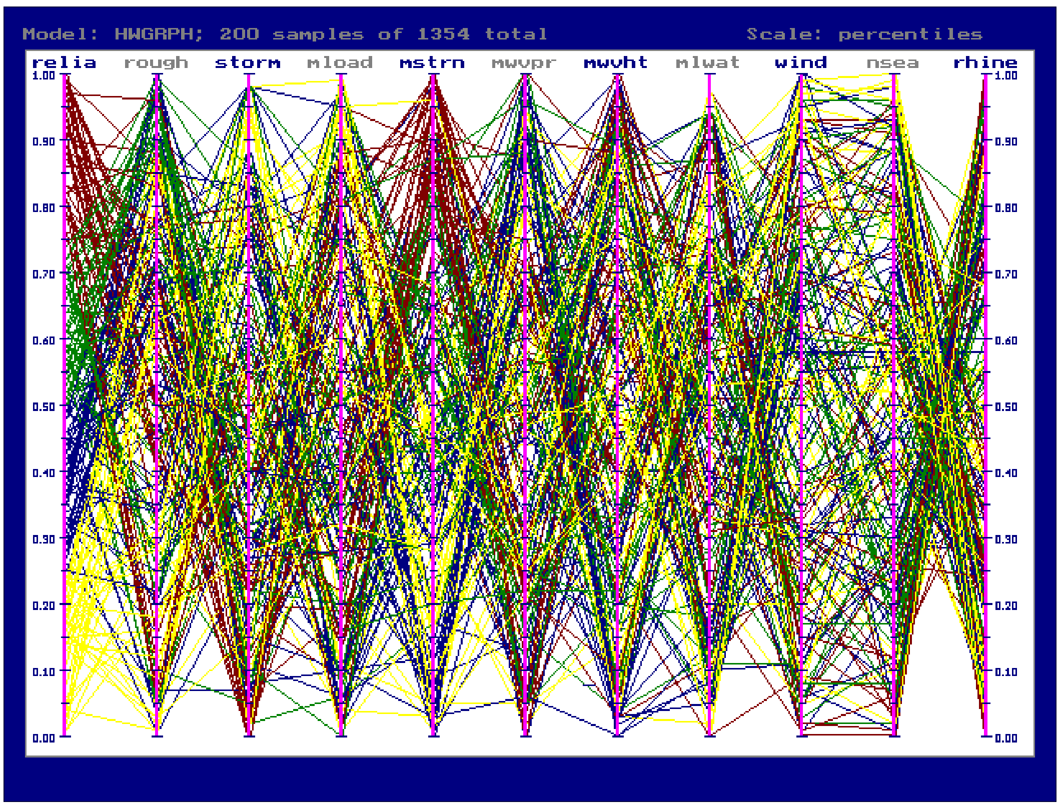

Fig. 1. Percentile cobweb plot of the joint distribution, in percentiles, of the limit-state function 'Z' and 15 explanatory variables.

‘mstrn’, ‘mwvpr’, ‘mwvht’, and ‘mlwat’, respectively), wind (‘wind’), North Sea (‘nsea’) and Rhine discharge (‘rhine’). Each vertical line represents one variable and each broken line represents one sample, intersecting each vertical line in the appropriate percentage point. These data are obtained by first conditionalizing on high, but not critical, sea and river water levels, giving 1354 samples. In 2% of these samples the dike ring actually fails corresponding to the lowest 2% of the variable Z. These 1354 samples are uniformly distributed over all vertical lines (for black and white visualization only 200 samples are shown in Figure 1). Since the water levels are obtained as a sum of contributions from the North Sea and the Rhine, this conditionalization has the effect of negatively correlating the North Sea and the Rhine discharge. These are uncorrelated in the unconditional sample. The negative correlation is shown by the fact that samples with high values for ‘nsea’ tend to have low values for ‘rhine’; ‘wind’ and ‘nsea’ show a strong positive correlation.

Figures 2-3 show four conditional percentile cobweb plots, where conditioning is done on various values of Z. The four cobweb plots correspond to conditionalizing on $Z \geq z_{95}$, $z_{35} \geq Z \geq z_{30}$, $z_{15} \geq Z \geq z_{10}$, and $z_5 \geq Z$. Departure from uniformity indicates that conditionalization affects the distribution of the corresponding variable.

In the top cobweb plot of Figure 2 the variables ‘storm’ and ‘mstrn’ differ most strongly from uniform. ‘mstrn’ is a global model factor to which the variables ‘Strength$_i$’ are positively coupled. We see that very high values of Z are associated with high values of ‘mstrn’ and low values of ‘storm’. From the bottom cobweb plot of Figure 2, we see that values of Z near the 30-th percentile are associated with low values of ‘mstrn’ and high values of ‘storm’. From the bottom cobweb plot of Figure 3, however, we see that values of Z between the 0-th and 5-th percentile are associated with distributions of ‘mstrn’ and ‘storm’ which more resemble the unconditional (uniform) distributions. Hence, in moving from $Z = z_5$ to $Z = z_{30}$ the conditional distribution of ‘mstrn’ is stochastically *decreasing*, and ‘storm’ is stochastically *increasing*. These variables are dissonant in this region. Very low values of Z characteristic of dike ring failure, are strongly associated with very high values of North Sea, and ‘mstrn’ and ‘storm’ regress to their unconditional distributions. In spite of this, Z is strictly increasing in ‘mstrn’ and is strictly decreasing in ‘storm’.

5 Local Probabilistic Sensitivity

In the literature, ‘local sensitivity’ is taken to refer to one point $(x_1, \ldots, x_n)$ in the sample space. Thus $\partial Z / \partial x_i$ is a local measure. For other measures, see e.g. Strozzi et al. [6].

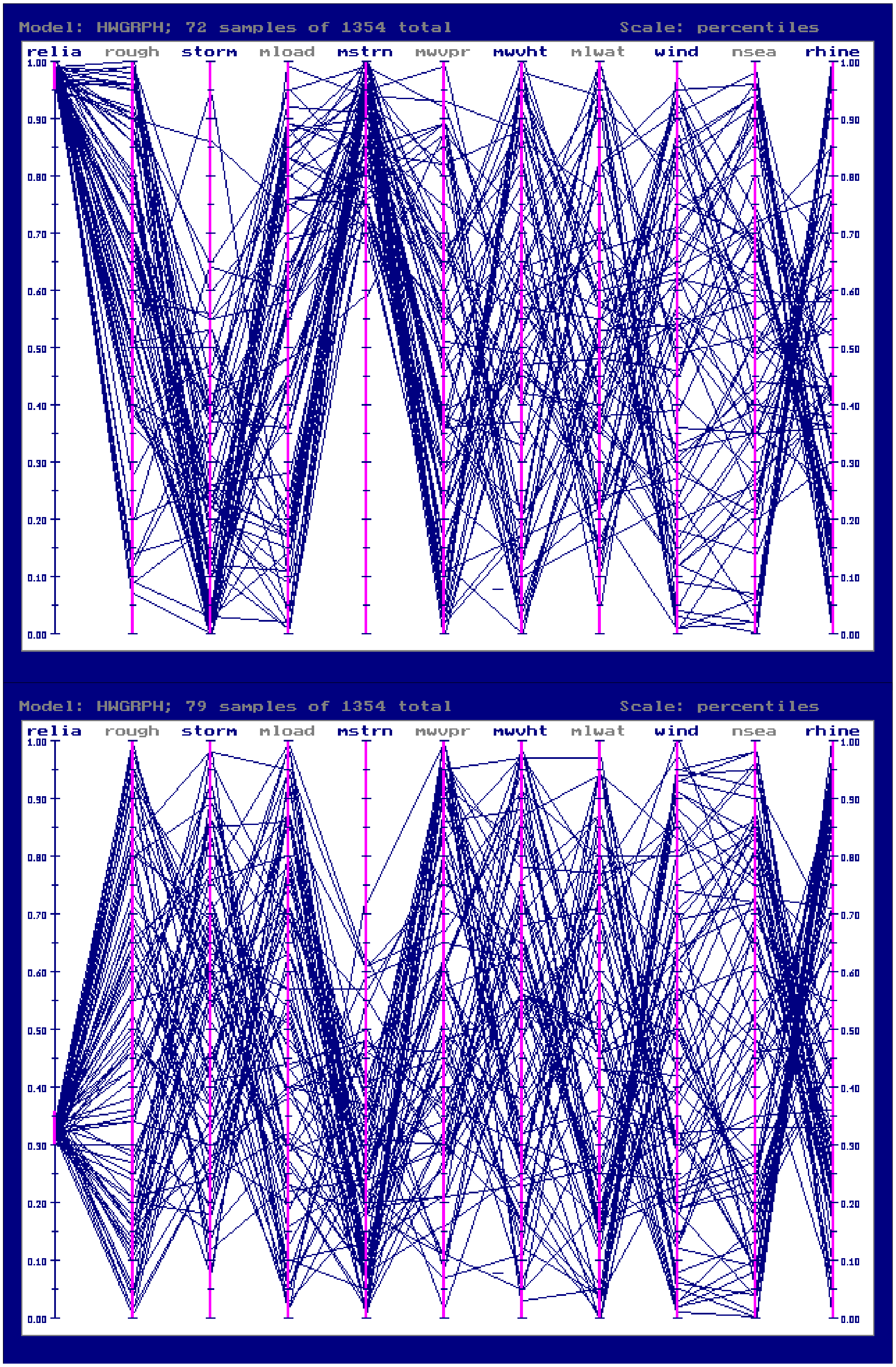

Fig. 2. Two conditional percentile cobweb plots, where conditioning is done on the limit-state function $Z =$'relia' for $Z \geq z_{95}$ and $z_{35} \geq Z \geq z_{30}$.

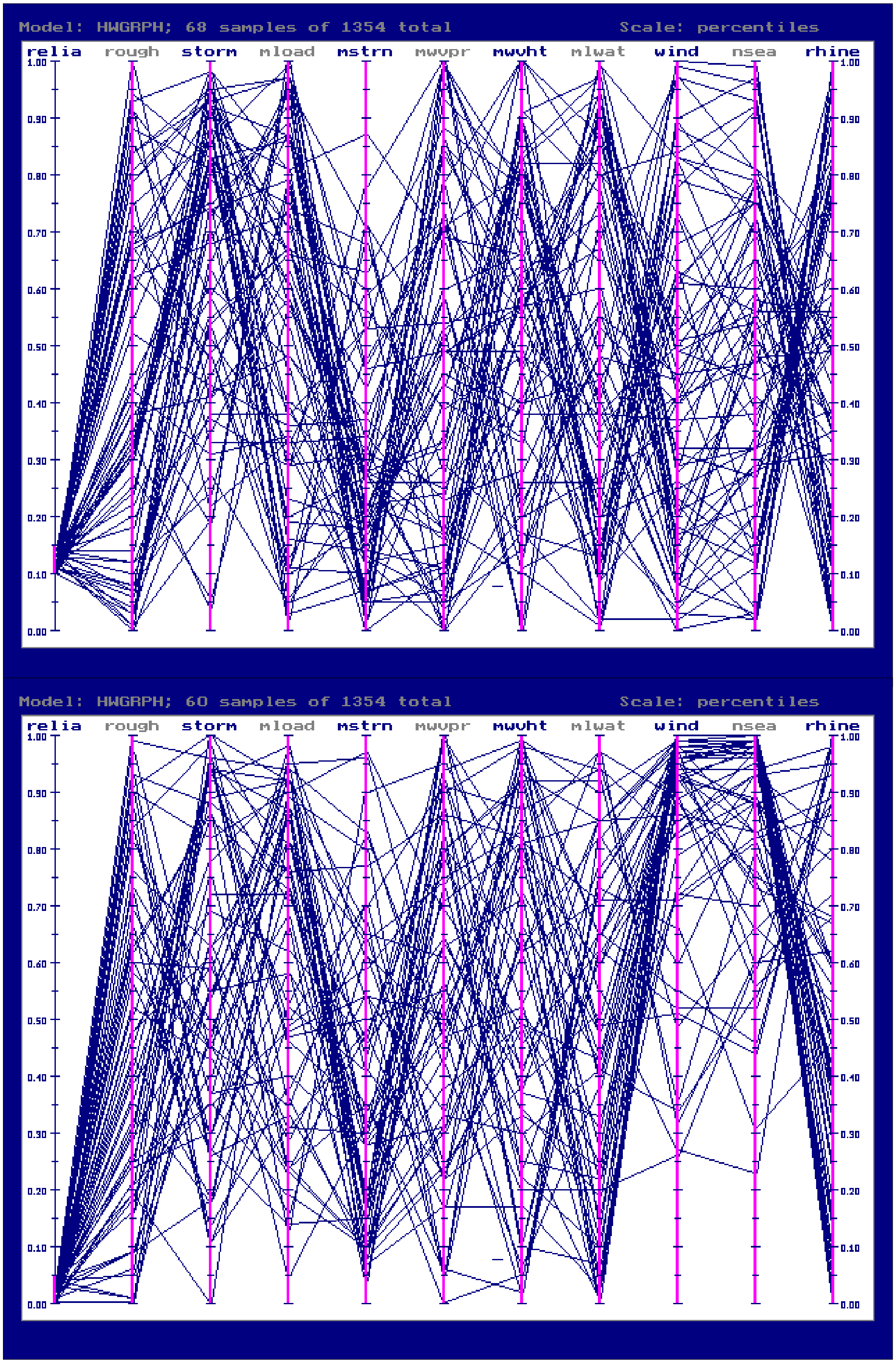

Fig. 3. Two conditional percentile cobweb plots, where conditioning is done on the limit-state function $Z =$'relia' for $z_{15} \geq Z \geq z_{10}$ and $z_5 \geq Z$.

In the present case we wish to focus on a unique value of a function of the uncertain parameters, and hence we focus on a manifold of lower dimension, typically a hypersurface. For convenience in the following discussion, we denote this manifold as the set

$$\{Z = 0\} = \{\mathbf{x} \in \mathbb{R}^n | Z(\mathbf{x}) = 0\}.$$

We wish to identify those components of $\mathbf{X}$ which are important for $\{Z = 0\}$. For this purpose we focus on the conditional random vector $\mathbf{X}|(Z = 0)$, and two approaches suggest themselves:

- compare $\mathbf{X}|(Z = 0)$ with $\mathbf{X}$, and
- consider the rate of change with respect to Z of (some function of) $\mathbf{X}|Z$ at $Z = 0$.

The first approach may miss important local behavior at $\{Z = 0\}$. Thus the conditional distribution of the storm length given $\{Z = 0\}$ resembles its unconditional distribution, but it is *not* independent of Z. Using the conditional expectation in the second approach affords good results which will coincide with FORM results when the linear model (1) holds.

Consider

$$E(X_i | Z = 0).$$

If X_i were independent of Z, then this conditional expectation would be simply $E(X_i)$. If $X_i = Z$, then clearly $E(X_i | Z = z) = z$. This suggests a local probabilistic sensitivity measure proportional to

$$\left. \frac{\partial E(X_i | Z = z)}{\partial z} \right|_{z=0}.$$

By requiring agreement with the FORM measure in the case the FORM assumptions hold, we can determine the appropriate proportionality constants. Proposition 1 relates this measure to the partial derivatives of a FORM linear approximation. This proposition is normally proved using properties of the joint normal distribution; however the more general proof method sketched in the Appendix can be used to obtain results under other distributional and functional assumptions, examples of which are given in Propositions 3 and 4.

Proposition 1 *Let the random vector* $\mathbf{X} = (X_1, \ldots, X_n)$ *have independent standard normal coordinates and let*

$$Z = \sum_{i=1}^{n} \alpha_i X_i,$$

then $\mathbf{X}|Z$ *is normal with*

$$
\begin{aligned}
E(X_i|Z=z) &= z\alpha_i/\ell^2, \\
\mathrm{Var}\,(X_i|Z=z) &= (\bar{\ell}_i/\ell)^2, \\
\rho(X_i,X_j) &= -\,\alpha_i\alpha_j/[\bar{\ell}_i\bar{\ell}_j],
\end{aligned}
$$

where $\ell^2 = \sum_{i=1}^n \alpha_i^2$ *and* $\bar{\ell}_j^2 = \sum_{i=1, i\neq j}^n \alpha_i^2$.

Assume now that for $X_i \sim \mathrm{N}(\mu_i, \sigma_i)$:

$$Z = \sum_{i=1}^n \alpha_i X_i = \sum_{i=1}^n [\alpha_i\sigma_i(X_i - \mu_i)/\sigma_i + \alpha_i\mu_i] = \sum_{i=1}^n [\alpha_i\sigma_i Y_i + \alpha_i\mu_i]$$

for $Y_i = (X_i - \mu_i)/\sigma_i \sim \mathrm{N}(0,1)$. Applying the above proposition, we find

$$\frac{\partial E(X_i|Z=z)}{\partial z} = \frac{\partial E(\sigma_i Y_i + \mu_i|Z=z)}{\partial z} = \frac{\sigma_i \partial E(Y_i|Z=z)}{\partial z} = \frac{\sigma_i^2\alpha_i}{\sum_{j=1}^n \sigma_j^2\alpha_j^2}.$$

From (2) we have $\sigma_i\alpha_i = \rho(X_i, Z)\sigma_z$, and since $\sum_{j=1}^n \sigma_j^2\alpha_j^2 = \sigma_Z^2$, we have:

$$\frac{\partial E(X_i|Z=z)}{\partial z} = \frac{\rho(X_i,Z)\sigma_i}{\sigma_z}.$$

We therefore take

$$\mathrm{LPSM}_i = \frac{\sigma_z \partial E(X_i|Z=z)}{\sigma_i \partial z} \tag{3}$$

as a local probabilistic sensitivity measure.

In the case of the linear model (1) we have $\mathrm{LPSM}_i = \rho(X_i, Z)$ in agreement with the FORM measure. However, if linearity does not hold, LPSM can be used to capture the local interpretation of (2) and it can be easily computed in MC calculations.

More generally, the above result holds whenever the regression of X_i on Z is linear:

Proposition 2 *Let* $E(X_i|Z) = kZ$ *for some constant* k*, then*

$$\mathrm{LPSM}_i = \frac{\sigma_z \partial E(X_i|Z=z)}{\sigma_i \partial z} = \rho(X_i, Z).$$

Propositions 3 and 4 present the local probabilistic sensitivity measures for sums of independent gamma variates (with equal scale parameters) and sums of independent exponential variates (with different means), respectively.

Proposition 3 *Let* $X_1, \ldots, X_n$ *be independent gamma variates with mean* α_i/β *and variance* α_i/β^2*, where* $\alpha_i > 0$ *and* $\beta > 0$ *for* $i = 1, \ldots, n$*. Suppose further that* $W = \sum_{i=1}^n X_i$ *and* $Z = W - \alpha_0$ *for* $\alpha_0 > 0$*. It is well-known that* *the conditional probability distribution of X_i when $W = w$ is a transformed beta distribution, i.e. that*

$$f_{X_i|W}(x_i|\,w) = \frac{\Gamma(\sum_{j=1}^{n} \alpha_j)}{\Gamma(\alpha_i)\Gamma(\sum_{j=1,\,j\neq i}^{n} \alpha_j)} \left[\frac{x_i}{w}\right]^{\alpha_i - 1} \left[1 - \frac{x_i}{w}\right]^{\left(\sum_{j=1,\,j\neq i}^{n} \alpha_j\right) - 1} \frac{1}{w}$$

for $0 \leq x_i \leq w$ and zero otherwise. The conditional expectation of X_i when $W = w$ equals

$$E(X_i|\,W = w) = w \cdot \alpha_i / [\textstyle\sum_{j=1}^{n} \alpha_j].$$

The local probabilistic sensitivity measure for the sum of independent gamma distributed random quantities can be written as

$$\mathrm{LPSM}_i = \frac{\sqrt{\sum_{j=1}^{n} \alpha_j/\beta^2}}{\sqrt{\alpha_i/\beta^2}} \cdot \left.\frac{\partial E(X_i|\,W = z + \alpha_0)}{\partial z}\right|_{z=0} = \sqrt{\frac{\alpha_i}{\sum_{j=1}^{n} \alpha_j}}.$$

Note that the LPSM does not depend on α_0.

Proposition 4 *Let $X_1, \ldots, X_n$ be independent exponential variates with mean α_i and variance α_i^2, where $\alpha_i > 0$ and $\alpha_i \neq \alpha_j$ unless $i = j$, $i = 1, \ldots, n$. Suppose further that $W = \sum_{i=1}^{n} X_i$ and $Z = W - \alpha_0$ for $\alpha_0 > 0$. The conditional expectation of X_i when $W = w = z + \alpha_0$ equals*

$$E(X_i|\,W = w) = g_i(w)/f_W(w),$$

where the function $g_i(w)$ can be found in Eq. (12) and the probability density function of the sum $W = \sum_{i=1}^{n} X_i$ is called the general Erlang or general gamma distribution:

$$f_W(w) = \sum_{i=1}^{n} \frac{1}{\prod_{j=1,\,j\neq i}^{n} [1 - \alpha_j/\alpha_i]} \times \frac{1}{\alpha_i} \exp\left\{-\frac{w}{\alpha_i}\right\}. \tag{4}$$

The general Erlang distribution has been used in theories of radioactive decay, queuing, psychology, and reliability (see e.g. Jensen [1], McGill & Gibbon [2], and Speijker et al. [5]). The local probabilistic sensitivity measure for the sum of independent exponentially distributed random quantities can be written as

$$\mathrm{LPSM}_i = \frac{\sqrt{\sum_{j=1}^{n} \alpha_j^2}}{\alpha_i} \cdot \frac{g_i'(\alpha_0) f_W(\alpha_0) - g_i(\alpha_0) f_W'(\alpha_0)}{[f_W(\alpha_0)]^2}, \tag{5}$$

where $g_i'(\alpha_0)$ can be found in Eq. (13). A special case of the general Erlang distribution, suggested by Jensen [1], arises when $\alpha_i = a/(b + i - 1)$ for $i = 1, \ldots, n$:

$$f_W(w) = \frac{\Gamma(b+n)}{\Gamma(b)\Gamma(n)} \left[\exp\left\{-\frac{w}{a}\right\}\right]^{b-1} \left[1 - \exp\left\{-\frac{w}{a}\right\}\right]^{n-1} \frac{1}{a} \exp\left\{-\frac{w}{a}\right\}.$$

The latter probability density function can be recognized as a transformed beta distribution.

6 Results

Table 1 shows the LPSM results for (some of) the variables in Figure 1. FORM results for the dike ring are not available at present. Both the FORM and the Monte Carlo calculations identify dike section 11 as the most critical dike section. Results for dike section 11 can be compared.

The low LSPM contribution of the Rhine discharge may be exaggerated by the negative correlation with the North Sea induced by the sampling technique. Other variables would not be affected in this way. We see that there are significant differences between the LPSM and the FORM measures. In light of Proposition 1, this is most likely explained by non-linearities in the Z function. Note that the variable 'rough' is quite dissonant for dike section 11. This indicates that low values of roughness are strongly associated with values of Z somewhat above zero, but near zero, roughness regresses to its unconditional

Table 1
Probabilistic sensitivity in terms of LPSM and FORM.

Variable	**Dike Ring LPSM**	**Dike Section 11 LPSM**	**Dike Section 11 FORM**
North Sea Location	0.29	-0.18	-0.043
North Sea Shape	-0.046	0.23	-0.0084
North Sea Scale	-0.083	0.083	-0.054
Rhine Location	-0.060	0.12	-0.024
Rhine Scale	-0.16	0.54	-0.078
Wind Angle Variability	0.30	0.13	0.063
Roughness	-0.22	-0.24	0.012
Storm Length	0.48	-0.10	-0.013
Independent Wind Location	0	0	-0.0038
Independent Wind Scale	-0.27	-0.068	0.051
Model factor Strength	-0.47	-0.11	na
North Sea	-1.57	-0.70	-0.72
Dependent Wind	-1.41	-0.78	na
Rhine discharge	-0.041	-0.12	-0.56

distribution. Evidently, numerical measures like LPSM or indeed the FORM measures must not be used uncritically. While they can be used to focus attention on interesting variables, a full understanding of their role should be based on graphical inspection of the joint distribution, as in Figures 2-3.

7 Summary and conclusions

For large models whose input parameters are uncertain, we are often confronted with the problem of choosing a small set of 'important' parameters. Global measures like (rank) correlations, or correlation ratio's may not be appropriate when we are interested in a specific region of the output variable(s). Thus, in modelling dike ring reliability, we are not interested in the variables driving dike ring reliability on normal days with low wind and water. Rather we are interested in the driving variables when the dike is near failure. Of course, on most days the dike integrity is not threatened, and global measures will be predominantly influenced by what happens on 'most days'.

Sensitivity measures such as partial derivatives, which are local in the input variables may miss important local probabilistic behavior. 'Probabilistic dissonance' is said to arise when variables are 'deterministically increasing' and yet 'probabilistic decreasing' (or vice versa) in certain regions. Such behavior is not at all uncommon in complex models, and cannot be discerned by 'deterministic' local sensitivity measures.

The local probabilistic sensitivity measures proposed here (3) are intended to identify variables which are active in a submanifold defined by a given point of the output variable. In the case of linear regression of input on output, these measures coincide with the product moment correlation. In other cases, comparing the local probabilistic sensitivity measures with correlations may be used to assess the departure from linearity.

Appendix

Proof of Proposition 1 The proposition can be proved using the conditionalization formula for the joint normal distribution applied to $(X_1, \dots, X_n, Z)$, even though the covariance matrix is singular (Rao [4]). However, a more flat-footed proof gives more insight. The basic steps are as follows.

(i) Express X_1 as a function $g(Z, X_2, \dots, X_n)$.
(ii) Write the conditional probability density function of $X_1, \dots, X_n$ given Z

as proportional to

$$f_{X_1}(g) \prod_{i=2}^{n} f_{X_i}(x_i),$$

where $f_{X_i}(x_i)$ is the probability density function of X_i.

(iii) Reduce the above, and set the coefficients of x_i^2, x_i, and $x_i x_j$ equal to the corresponding terms in the joint normal density for $n-1$ variables. This involves solving a system of simultaneous equations. Terms without any x's are absorbed into the proportionality constant.

(iv) Since the system of simultaneous equations has a solution, it follows that the conditional distribution is joint normal, and the parameters can be obtained from the appropriate coefficients.

Rather than giving a general proof, we illustrate the proof with an index-free version for $n = 3$. Let

$$Z = aW + bX + cY.$$

The conditonal probability density function given $Z = z$ is proportional to

$$\exp\left\{-\frac{1}{2}\left[\left(\frac{z - bx - cy}{a}\right)^2 + x^2 + y^2\right]\right\}.$$

If the probability density function is joint normal, this must be proportional to

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]\right\}.$$

Expanding these expressions, and dropping terms not involving x or y, leads to the equations

$$x^2[1 + (b/a)^2] = x^2/[\sigma_x^2(1-\rho^2)], \quad (6)$$
$$y^2[1 + (c/a)^2] = y^2/[\sigma_y^2(1-\rho^2)], \quad (7)$$
$$zb(1-\rho^2)/a^2 = \mu_x/\sigma_x^2 - \rho\mu_y/[\sigma_x\sigma_y], \quad (8)$$
$$zc(1-\rho^2)/a^2 = \mu_y/\sigma_y^2 - \rho\mu_x/[\sigma_x\sigma_y], \quad (9)$$
$$bc/a^2 = -\rho/[\sigma_x\sigma_y(1-\rho^2)]. \quad (10)$$

Solving these equations for μ_x, μ_y, σ_x, σ_y, and ρ leads to the desired expressions.

Some hints in solving these are as follows. Put $D = \sqrt{(a^2+b^2)(a^2+c^2)}$. From Eq. (6) follows

$$a^2/[(1-\rho^2)D] = \sigma_x\sigma_y. \quad (11)$$

Eqs. (10) and (11) give $\rho = -bc/D$. Putting this into (6) and (7) gives

$$\sigma_x^2 = (a^2 + c^2)/\ell^2, \quad \sigma_y^2 = (a^2 + b^2)/\ell^2,$$

where $\ell^2 = a^2 + b^2 + c^2$. Solving (8) and (9) for μ_x and μ_y proves the proposition for $n = 3$. □

Proof of Proposition 2 Since

$$E(X_i Z) = E(E(X_i|Z)Z) = kE(Z^2)$$

and

$$E(X_i)E(Z) = kE^2(Z),$$

it follows that

$$\rho(X_i, Z) = \frac{k[E(Z^2) - E^2(Z)]}{\sigma_Z \sigma_i} = \frac{k\sigma_Z}{\sigma_i}.$$

□

Proof of Proposition 4 Let $X_1, \ldots, X_n$ be independent exponential variates with mean α_i and variance α_i^2, where $\alpha_i > 0$ and $\alpha_i \neq \alpha_j$ unless $i = j$, $i = 1, \ldots, n$. Suppose further that $W = \sum_{i=1}^n X_i$ and $Z = W - \alpha_0$ for $\alpha_0 > 0$. By using the general Erlang distribution given in Eq. (4) twice, the conditional probability distribution of X_i when $W = w$ can be written as

$$f_{X_i|W}(x_i|\, w) =$$
$$= \frac{\frac{1}{\alpha_i} \exp\left\{-\frac{x_i}{\alpha_i}\right\} \times \sum_{k=2}^{n} \frac{1}{\prod_{j=2,\, j\neq k}^{n} [1 - \alpha_j/\alpha_k]} \times \frac{1}{\alpha_k} \exp\left\{-\frac{w - x_i}{\alpha_k}\right\}}{\sum_{k=1}^{n} \frac{1}{\prod_{j=1,\, j\neq k}^{n} [1 - \alpha_j/\alpha_k]} \times \frac{1}{\alpha_k} \exp\left\{-\frac{w}{\alpha_k}\right\}}$$

for $0 \leq x_i \leq w$ and zero otherwise. The conditional expectation of X_i when $W = w = z + \alpha_0$ equals

$$E(X_i|\, W = w) =$$
$$= \sum_{k=2}^{n} \frac{\exp\{-w/\alpha_k\}/\alpha_k}{\prod_{j=2,\, j\neq k}^{n} [1 - \alpha_j/\alpha_k]} \times \int_{x_i=0}^{w} \frac{x_i}{\alpha_i} \exp\left\{-\left[\frac{1}{\alpha_i} - \frac{1}{\alpha_k}\right] x_i\right\} dx_i \times \frac{1}{f_W(w)} =$$
$$= \sum_{k=2}^{n} \frac{\exp\{-w/\alpha_k\}}{\prod_{j=2,\, j\neq k}^{n} [1 - \alpha_j/\alpha_k]} \times \frac{1}{[1/\alpha_i - 1/\alpha_k]^2} \times \frac{1}{\alpha_i \alpha_k} \times$$
$$\times \left[1 - \left(1 + \left[\frac{1}{\alpha_i} - \frac{1}{\alpha_k}\right] w\right) \exp\left\{-\left[\frac{1}{\alpha_i} - \frac{1}{\alpha_k}\right] w\right\}\right] \times \frac{1}{f_W(w)} =$$

$$= \left(\sum_{k=2}^{n} \frac{1}{\prod_{j=1,\, j \neq k}^{n} [1 - \alpha_j/\alpha_k]} \times \frac{1}{[\alpha_k/\alpha_i - 1]} \left[\exp\left\{ -\frac{w}{\alpha_k} \right\} - \exp\left\{ -\frac{w}{\alpha_i} \right\} \right] + \right.$$

$$\left. + \frac{1}{\prod_{j=1,\, j \neq i}^{n} [1 - \alpha_j/\alpha_i]} \times \frac{w}{\alpha_i} \exp\left\{ -\frac{w}{\alpha_i} \right\} \right) \times \frac{1}{f_W(w)} =$$

$$= \frac{g_i(w)}{f_W(w)}. \tag{12}$$

Taking the partial derivative of $E(X_i | W = z + \alpha_0)$ with respect to z proves Eq. (5), where

$$g_i'(\alpha_0) =$$

$$= \sum_{k=2}^{n} \frac{1}{\prod_{j=1,\, j \neq k}^{n} [1 - \alpha_j/\alpha_k]} \times \frac{1/\alpha_k}{[\alpha_k/\alpha_i - 1]} \left[\exp\left\{ -\frac{\alpha_0}{\alpha_i} \right\} - \exp\left\{ -\frac{\alpha_0}{\alpha_k} \right\} \right] -$$

$$- \frac{1/\alpha_i}{\prod_{j=1,\, j \neq i}^{n} [1 - \alpha_j/\alpha_i]} \times \frac{\alpha_0}{\alpha_i} \exp\left\{ -\frac{\alpha_0}{\alpha_i} \right\}. \tag{13}$$

□

References

[1] A. Jensen. An elucidation of Erlang's statistical works through the theory of stochastic processes. In E. Brockmeyer, H.L. Halstrøm, and Arne Jensen, editors, *The Life and Works of A.K. Erlang*, pages 23–100. Copenhagen: Transactions of the Danish Academy of Technical Science, No. 2, 1948.

[2] W.J. McGill and J. Gibbon. The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, 2:1–18, 1965.

[3] M.D. McKay. Nonparametric variance-based methods of assessing uncertainty importance. *Reliability Engineering and System Safety*, 57:267–279, 1997.

[4] C.R. Rao. *Linear Statistical Inference and Its Applications*. New York: John Wiley & Sons, 1973.

[5] L.J.P. Speijker, J.M. van Noortwijk, M. Kok, and R.M. Cooke. Optimal maintenance decisions for dykes. Technical Report 96-39, Faculty of Mathematics and Computer Science, Delft University of Technology, The Netherlands, 1996.

[6] F. Strozzi, V. Calenbuhr, M.A. Alos, and J.M. Zaldivar. Sensitivity analysis using Lyapunov exponents: application to chemical reactors. In A. Saltelli and H. von Maravić, editors, *SAMO 95: Theory and applications of sensitivity analysis of model output in computer simulation*, pages 55–57. Luxembourg: EUR 16331, 1995.

[7] J.W. van der Meer and J.P.F.M. Janssen. Wave run-up and wave overtopping at dikes and revetments. Publication 485, Delft Hydraulics, The Netherlands, 1994.