Updating piping probabilities with survived loads

Timo Schweckendiek\textsuperscript{1,2}, Wim Kanning\textsuperscript{1}

\textsuperscript{1} Section of Hydraulic Engineering, Delft University of Technology, Delft, Netherlands
\textsuperscript{2} Unit Geo-engineering, Deltares, Delft, Netherlands

Abstract: Piping, also called under-seepage, is an internal erosion mechanism, which can cause the failure of dikes or other flood defence structures. The uncertainty in the resistance of a flood defence against piping is usually large, causing high probabilities of failure for this mechanism. A considerable part of this uncertainty is of epistemic nature, which can be reduced by incorporating extra information. It is shown how the knowledge of historically survived water levels, the main load factor, can be integrated into the probability distribution of the piping resistance variables by means of Bayesian Updating. The effects are demonstrated by means of a realistic numerical example.

1 Introduction

Piping, also called under-seepage, is an internal erosion mechanism, which can cause the failure of dikes or other flood defence structures. This failure mechanism is particularly relevant for river dikes on sandy subsoil, typical conditions in flood protection systems in delta areas. In recent years, the issue of piping received special attention in the Netherlands. Results from the FLORIS (FLOod RIsks and Safety in the Netherlands) project, a nation-wide flood risk analysis applying a reliability analysis to the entire flood defence system, indicate that in some regions the probability of failure due to piping is very high, far greater than acceptable. A large part of the uncertainty causing these high probabilities of failure stem from knowledge (epistemic) uncertainty in the piping resistance and the subsoil conditions. This uncertainty can, in principle, be reduced by incorporating extra information. This paper deals with one specific source of information, the knowledge of historically survived load conditions. A survived load can be interpreted as an incomplete load test. This knowledge can be integrated into the probability distribution of the piping resistance variables by means of Bayesian Updating. The effects are demonstrated by means of a realistic numerical example.
2 Bayesian Updating based on survived loads

2.1 Bayesian Updating

Let $\Theta$ be the random variable for the parameter of a distribution of the quantity of interest $X$, with a prior PDF $f_\Theta(\theta)$. Based non the evidence $\epsilon$ the posterior distribution can be determined by Bayes’ theorem (Bayes [2]):

$$ f_\Theta(\theta|\epsilon) = \frac{P(\epsilon|\theta)f_\Theta(\theta)}{\int_{-\infty}^{\infty} P(\epsilon|\theta)f_\Theta(\theta)} $$

where $P(\epsilon|\theta)$ is the probability of observing the evidence, given parameter $\Theta = \theta$, also called the likelihood $\mathbb{L}(\theta)$. Eq. (1) may also be written in simplified form with the inverse of the denominator expressed as normalizing constant $k$:

$$ f_\Theta(\theta|\epsilon) = k\mathbb{L}(\theta)f_\Theta(\theta) $$

The uncertainty in the estimation of the parameter can be included in the calculation of the probability of the quantity of interest $X$. By “integrating out” the uncertainty in $\theta$, we obtain a composite distribution, the so called Bayesian distribution of $X$:

$$ \tilde{f}_X(x) = \int_{-\infty}^{\infty} f_X(x|\theta)f_\Theta(\theta)d\theta $$

Thus, using $f_\Theta(\theta|\epsilon)$ in Eq. (3) results in the updated Bayesian distribution of $X$. It is emphasized that, in contrast to frequentist approaches, the Bayesian distribution includes both, aleatory and epistemic uncertainties. Therefore, the variance (spread) of such a distribution is in general larger than the variance of the “true” distribution $f_X(x)$, and therefore probably larger than frequentist estimates of the variance. For the remainder of this paper, the Bayesian distribution $\tilde{f}_X(x)$ is used to account for the (total) uncertainty in $X$ in uncertainty, reliability or decision analyses. For sake of simplicity, in the following the tilde is omitted and the distribution of $X$ is just referred to as $f_X(x)$. For a more detailed discussion and examples it is referred to Benjamin and Cornell [3], Ang and Tang [1] and Tang [7].

2.2 Likelihood based on survived loads

Analogue to Bayesian parameter estimation survival analysis, the likelihood function of resistance parameters based on a survived load condition is the probability of survival or non-failure, given the survived load and the parameters. The load may be univariate and known exactly in the simplest case, it may also be a load combination (multivariate) and uncertain, e.g. due to measurement, transformation or model uncertainties.

To give an insightful example, contemplate the simple case of the limit state function $Z = R - S$ ($R =$ resistance, $S =$ solicitation or load). The likelihood function of $R$, given the historically survived load $S^*$ can be formulated as:
\[ \mathbb{L}(\theta) = P(R > S^*) = 1 - F_R(S^*|\theta) \]  \hspace{1cm} (4)

More generally, for a limit state function \( Z = g(\mathbf{X}) \), where \( \mathbf{X} \) is the vector of random variables with parameters \( \theta \), the likelihood function is given by the probability of \( Z \) assuming positive values, given the survived load \( S^* \):

\[ \mathbb{L}(\theta) = P(Z > 0|\theta, S^*) = 1 - P_{f|\theta, S^*} = 1 - \int_{-\infty}^{\infty} I_{Z < 0} f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} \]  \hspace{1cm} (5)

where \( I(\cdot) \) is the indicator function. Thus, determining the likelihood function requires solving a classical structural reliability problem, conditional on the observed load for the feasible range of resistance parameters. The prior information on either \( R \) or \( \mathbf{X} \), respectively on their distributions parameters, may stem from either engineering or expert judgment or from any other source of data, such as measurements of load and strength properties.

### 3 Application to piping

#### 3.1 Simple piping model

To illustrate the concept described above and its effects on reliability considerations, the piping model first established by Bligh [4] is adopted (see Fig. 1).

![Illustration of schematic geometry and parameters for heave and piping](image)

Fig. 1: Illustration of schematic geometry and parameters for heave and piping

The limit state function (LSF) for the adopted model is given by:

\[ Z = L - m_C \ C(h - h_b - 0.3d) \]  \hspace{1cm} (6)

with

- \( L \) piping length [m]
- \( C \) Bligh’s Creep parameter [-] (erosion resistance)
- \( m_C \) factor for model uncertainty associated with \( C \)
- \( h \) water level (main load) [m+NAP]
- \( h_b \) landside water level [m+NAP]
- \( d \) thickness of aquitard at exit point [m]
The formulation is similar to Steenbergen and Vrouwenvelder [6], re-written in a form that is linear in all parameters. Using this piping model, we implicitly make the following assumptions:

- If the maximum head difference exceeds the critical head difference, failure occurs (i.e., time-dependent effects in the piping mechanism are disregarded).
- No cumulative effect, i.e. there is no cumulative damage or “fatigue” effect.
- Piping can occur directly, i.e. there is no dependence on other mechanisms such as uplift. This assumption is realistic in the absence of blanket layers or for thin blanket layers with very low uplift resistance.

3.2 Uncertainties in piping resistance

Tab. 1 shows design values $C_d$ of the piping resistance parameter $C$ for different soil types according to TAW [8].

Tab. 1: Creep factors for piping rules

<table>
<thead>
<tr>
<th>Soil type</th>
<th>Median grain diameter [µm]</th>
<th>$C_d$ (Bligh) [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silt</td>
<td>≤ 105</td>
<td>18</td>
</tr>
<tr>
<td>Very fine sand</td>
<td>105 – 150</td>
<td>18</td>
</tr>
<tr>
<td>Very fine sand (mica)</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Medium-fine sand (quartz)</td>
<td>150 – 210</td>
<td>15</td>
</tr>
<tr>
<td>Medium-coarse sand</td>
<td>210 – 300</td>
<td>12</td>
</tr>
<tr>
<td>Very coarse sand</td>
<td>300 – 2000</td>
<td>9</td>
</tr>
<tr>
<td>Fine gravel</td>
<td>2000 – 5600</td>
<td></td>
</tr>
<tr>
<td>Medium-coarse gravel</td>
<td>5600 – 16000</td>
<td></td>
</tr>
<tr>
<td>Very coarse gravel</td>
<td>&gt; 16000</td>
<td>4</td>
</tr>
</tbody>
</table>

1indications in accordance with NEN 5104 (September 1989)

The modeling of uncertainties in piping resistance for this study is adopted from the FLORIS project (VNL) in the Netherlands (Rijkswaterstaat [5]). Since the tabulated design values $C_d$ are conservative by definition, it is assumed that the underlying distribution of the uncertain variable has a lognormal distribution with moments $C \sim LN(\mu_C, \sigma_C)$ with $\mu_C = C_d/1.5$ and $\sigma_C = 0.1$. $C_d = 0.15\mu_C$ (c.o.v. of 15 %).

In order to account for the model uncertainty related to $C$, a model factor $m_C$ is used (multiplicative model). It expresses the imperfection of $C$ to represent the resistance against piping; even knowing the values of all resistance variables including $C$ exactly, one could not predict the critical head difference with certainty. Its distribution is assumed to be $m_C \sim N(1, 0.15)$.

The other resistance variable, the piping length $L$, is commonly modeled by means of a Normal distribution, too. The mean value is based on measurements or “educated guesses”, best on the data available of geometry and subsurface. A realistic value to account for the uncertainty in the piping length is a c.o.v. of 10%, resulting in $L \sim N(\mu_L, 0.1\mu_L)$. One could also count $d$ (thickness of the blanket layer) to the resistance variables in the piping performance function. In the examples treated in this paper $d$ is treated deterministically.
4 Numerical Example

The concept described above is illustrated by means of a numerical example. The chosen parameters are realistic and very similar to the real data of a location in the Eastern part of the Netherlands. The a-priori parameters are shown Tab. 2. In the subsequent sections, modifications of this case are presented and the varying assumptions are defined at the beginning of each section.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Distribution</th>
<th>Parameters / moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$ [-]</td>
<td>Lognormal</td>
<td>$\mu_C = 15.0$ $\sigma_C = 2.0$</td>
</tr>
<tr>
<td>$m_C$ [-]</td>
<td>Normal</td>
<td>$\mu_{m_C} = 1.0$ $\sigma_{m_C} = 0.15$</td>
</tr>
<tr>
<td>$L$ [m]</td>
<td>Normal</td>
<td>$\mu_L = 30.0$ $\sigma_L = 3.0$</td>
</tr>
<tr>
<td>$h$ [m+NAP]</td>
<td>Gumbel</td>
<td>$\alpha = 8.27$ $\beta = 0.406$</td>
</tr>
<tr>
<td>$h_b$ [m+NAP]</td>
<td>deterministic</td>
<td>$h_b = 7.5$</td>
</tr>
<tr>
<td>$d$ [m]</td>
<td>deterministic</td>
<td>$d = 0.8$</td>
</tr>
</tbody>
</table>

In the subsequent examples it is assumed that a water level with a 100 year return period of $h^* = 10.14$ [m+NAP] has been observed and no piping failure has occurred.

4.1 One uncertain resistance variable (case 1)

For sake of illustration, the first case is simplified with respect to the general case. Only $C$ is considered uncertain on the resistance side, $L$ and $m_C$ are considered deterministic with their respective mean values. In order to be able to compare the cases, the standard deviation of $C$ is chosen as $\sigma_C = 3.0$, which can be shown is approximately equivalent to the combined distribution of $m_C \cdot C$ with their original parameters from Tab. 2. The uncertainty in $C$ is assumed to be of epistemic nature and can, in principle be reduced. It is noted that these assumptions were only made for sake of illustration and are not realistic.

The likelihood function of $C$ can be formulated as:

$$\mathbb{L}(C) = P(Z > 0|e) = P(L - m_C \cdot C (h^* - h_b - 0.3d) > 0|C)$$

(7)

Since $C$ was the only variable considered random in this case, the likelihood function is deterministic may be re-written as:

$$\mathbb{L}(C) = \begin{cases} P(Z > 0|C) = 1 & \text{if } C < C_c \\ P(Z > 0|C) = 0 & \text{if } C \geq C_c \end{cases}$$

with $C_c = L/(h^* - h_b - 0.3 \cdot d) =$ critical Bligh parameter [-]

In other words, the survival of $h^*$ has proven that $C$ must be smaller than $C_c$. Fig. 2 illustrates that the observed water level had a low yearly exceedance probability (1 %) and that the critical (i.e. “proven”) resistance value was relatively high (low $C$ values mean high resistance). That implies that a large part of the probability mass of $C$ can be redistributed, in fact the part with values greater than $C_c = 12.51$ [-].
The effect of updating can be appreciated in the posterior distribution of $C$ (see Fig. 3 (a)) and the updated fragility curve ($P\{Z < 0|h\}$). In fact, the posterior distribution of $C$ is the prior distribution truncated at $C_c$ (and normalized).

In Fig. 3 (b), an example of a short term prediction with expectation $E[h|c] = 9.5$ [m+NAP] and standard prediction error of $\sigma_{h|c} = 0.3$ [m] (typical for about 3 days lead time on rivers).
was plotted together with the fragility curves in order to appreciate the impact of the updating procedure on short-term predictions.

There is also an effect of the updated resistance on the reliability. As indicated in Tab. 3, the probability of failure using the water level distribution of the yearly maxima (long-term) decreases by a factor of 5. The effect on the short-term reliability with the scenario of a water level prediction as in Fig. 3 is even much more significant, though for water level predictions higher than the observed survived level the effect may vanish.

**Tab. 3: Prior and posterior probabilities of failure (case 1)**

|                      | with \( f(C) \)   | with \( f(C|\epsilon) \) |
|----------------------|-------------------|----------------------------|
| Probability of failure \( P_f \) | 3.1 E-2           | 6.1 E-3                    |
| Reliability index \( \beta \) | 1.86              | 2.51                       |
| Probability of failure \( P_f \) | 2.8 E-1           | 4.1 E-3                    |
| Reliability index \( \beta \) | 0.59              | 2.63                       |

**4.2 One uncertain resistance variable and model uncertainty (case 2)**

In the second case, model uncertainty is added as uncertainty, whilst the piping length \( l \) is still treated deterministically. As mentioned, with the parameters in Tab. 2, the aggregate uncertainty of \( m_C \cdot C \) is practically equivalent to the uncertainty of \( C \) in the previous case. This way the results can be compared well.

When several uncertain variables or parameters are involved, one has to be careful with the definition of the likelihood function. The basic question is, which variables can be updated and which ones cannot. E.g. epistemic uncertainties can be updated, whilst intrinsic uncertainties cannot. In this particular example, the fact of survival provides information about the actual value of \( h \), whilst it does not prove information on the model uncertainty. In fact, the model uncertainty weakens the effect of updating \( C \). The likelihood function is again:

\[
\mathbb{L}(C) = P(Z > 0|\epsilon) = P(L - m_C \cdot C \cdot (h^* - h_b - 0.3d) > 0|C)
\]  

However, in this case, it is a stochastic function, since \( m_C \) is uncertain. The likelihood function and the posterior distribution are shown in Fig. 4 (a). The same figure also shows that the posterior distribution is well described by a lognormal distribution with moments \( \mu_{C|\epsilon} = 13.4 \) and \( \sigma_{C|\epsilon} = 1.54 \) (Fig. 4). The impact on the probability of failure much less compared to case 1 (see Tab. 4), this is mainly due to the (irreducible model uncertainty).

**Tab. 4: Prior and posterior probabilities of failure (case 2)**

|                      | with \( f(C) \)   | with \( f(C|\epsilon) \) |
|----------------------|-------------------|----------------------------|
| Probability of failure \( P_f \) | 3.1 E-2           | 1.8 E-2                    |
| Reliability index \( \beta \) | 1.86              | 2.09                       |
| Probability of failure \( P_f \) | 2.8 E-1           | 1.4 E-1                    |
| Reliability index \( \beta \) | 0.59              | 1.08                       |
The effect on the fragility curve in Fig. 4 is much less for the present case with model uncertainty, which is also reflected in the lesser impact on the short-term reliability in Tab. 4.

4.3 Two uncertain resistance variables (case 3)

In the third case, two uncertain resistance variables, $C$ and $I$, are considered, whilst model uncertainty is disregarded ($m_C = 1$, deterministic). The aggregate uncertainty in $C$ and $I$ is again comparable to the previous cases, so the impact on the reliability can be compared, too. The uncertainty in $C$ and $I$ is assumed to be purely epistemic, similar to case 1, however in this case the likelihood function is two-dimensional:

$$
\mathbb{L}(C, L) = P(Z > 0 | \epsilon) = P(L - m_C C(h^* - h_b - 0.3d) > 0 | C, L)
$$  \hspace{1cm} (10)

Without model uncertainty, a line can be determined, given the survived water level, that separates the combinations of $C$ and $L$ that should have failed under this loading from those that are still possible. This is illustrated in Fig. 5 (a). The parameter combinations on the upper left hand side of the limit state line ($Z = 0 | h^*$) have practically been proven to be impossible, given the observation of survival. In the Bayesian Updating process this probability mass is re-distributed to the lower right hand side (see Fig. 5 (b)), as becomes also clear by rewriting the deterministic likelihood function in the following form:

$$
\mathbb{L}(C, L) = \begin{cases} 
P(Z > 0 | C) = 1 & \text{if } L/C < (h^* - h_b - 0.3d) \\ 
P(Z > 0 | C) = 0 & \text{if } L/C \geq (h^* - h_b - 0.3d) \end{cases}
$$  \hspace{1cm} (11)
Having determined the posterior joint probability density, we can determine the posterior marginal distributions of $C$ and $L$ (in general: $f_X(x) = \int_y f_{X,Y}(x,y)dy$), see Fig. 6. The posterior moments differ significantly from the prior ones.

Using the posterior distributions, one has to be very careful, since the posterior marginal distributions are not independent anymore, even though the prior marginal distributions were assumed to be independent. The safest way to use the updated information in further analyses is using the joint pdf directly, whether it has been determined numerically or by simulation (e.g. Markov Chain Monte Carlo). If that is not possible, the correlation structure has to be taken care of in any other suitable manner.
In the current analysis, the posterior reliability analyses have been carried out using numerical integration of the joint pdf $f(C, L|\varepsilon)$ (fourth column in Tab. 5). In order to demonstrate the error using independent marginal distributions, these results are given in the last column of the same table ($f(C|\varepsilon)f(L|\varepsilon)$).

Tab. 5: Prior and posterior probabilities of failure (case 3)

|                      | with $f(C, L)$ | with $f(C, L|\varepsilon)$ | with $f(C|\varepsilon)f(L|\varepsilon)$ |
|----------------------|----------------|-----------------------------|-----------------------------------------|
| Probability of failure $P_f$ | 3.1 E-2        | 6.4 E-3                     | 7.5 E-3                                 |
| Reliability index $\beta$   | 1.86           | 2.49                        | 2.43                                    |
| Probability of failure $P_f$ | 2.8 E-1        | 5.5 E-3                     |                                          |
| Reliability index $\beta$   | 0.59           | 2.54                        |                                          |

It is remarkable that the posterior probabilities of failure, i.e. the effects of updating the resistance variables, are practically the same as in case 1. This can be explained by the fact that, on the one hand, the aggregate uncertainty in the resistance was practically equal, and, on the other hand, despite of having two uncertain variables in the absence of model uncertainty we could assign probability zero to some regions of the parameter space of $C$ and $I$, just as in case 1.

Also the effect on the fragility curve (see Fig. 7) is very much similar to case 1 due to the same reasons. This is also reflected in the posterior values for short term reliability (water level prediction) in Tab. 5.

![Fig. 7: Prior and posterior fragility curves (case 3)](image)

**4.4 Two uncertain resistance variables and model uncertainty (case 4)**

In the last case, all random variables are treated as uncertain as specified at the beginning of this section. Notice that the total uncertainty is different and that the results cannot be compared to the previous cases directly anymore.
The likelihood function is the same as in the previous case 3 (eq. 10), however, in this case it is a stochastic function. In the previous case, the likelihood function distinguished sharply between zero and one probability, in the current case it is a smooth function (see Fig. 8 (a)).

![Likelihood and Posterior joint pdf](image)

**Fig. 8: Contours of likelihood function and posterior distribution (case 4)**

The posterior marginal distributions and their statistical moments are displayed in Fig. 9. Neither the shift in mean value nor the decrease in variance is as significant as in case 3. The model uncertainty weakens the updating effect, as already shown in case 2.

![Posterior marginal distributions of C and L](image)

**Fig. 9: Posterior marginal distributions of C and L (case 4)**

The posterior marginal distributions again follow well their original distribution types.
Tab. 6: Prior and posterior probabilities of failure (case 4)

<table>
<thead>
<tr>
<th>with $f(h)$ (long-term)</th>
<th>prior</th>
<th>posterior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of failure $P_f$</td>
<td>3.7 E-2</td>
<td>1.6 E-2</td>
</tr>
<tr>
<td>Reliability index $\beta$</td>
<td>1.79</td>
<td>2.13</td>
</tr>
</tbody>
</table>

| with $f(h|\varepsilon)$ (short-term) | prior | posterior |
|-------------------------|-------|-----------|
| Probability of failure $P_f$ | 2.9 E-1 | 1.1 E-1 |
| Reliability index $\beta$ | 0.54 | 1.20 |

The effect on the reliability is similar to case 2 (see Tab. 6), even though the two cases are not perfectly comparable anymore due to the difference in total (aggregate) prior uncertainty. Differences could be due to the accuracy of the numerical solution procedure. The same holds for the fragility curves presented in Fig. 10.

![Prior and posterior fragility curves (case 4)](image)

Fig. 10: Prior and posterior fragility curves (case 4)

Notice that so far all calculations have been carried out mainly by means of numerical integration techniques. The last case is a four-dimensional problem and requires already considerable calculation effort and computer memory. Higher dimensional problems might require the use of alternative techniques, such as Markov Chain Monte Carlo sampling.

5 Conclusions and outlook

Bayesian statistics have been applied to update the probability distributions of resistance variables based on survived load conditions in a structural reliability problem. The key to the problem is establishing a suitable likelihood function. This is usually the probability of survival or non-failure, given the observed load, which itself may be uncertain e.g. due to measurement errors. From the simplified numerical example we can conclude that in the most simplified and rather artificial form (case 1: only epistemic uncertainty in one resistance variable, certain load observation) the impact of updating is significant. In the more realistic cases including model uncertainty, the effect of updating is much less pronounced. It is interesting to see that the number of variables to be updated itself does not influence
the effect of updating, if the total variance is kept the same, only intrinsic or non-reducible uncertainties like the model uncertainty in cases 2 and 4 do. The more uncertainty of the latter type is present in a problem, the lesser the value of information of survival of observed load. The updated fragility curves show that the effect may still be interesting for operational forecasting purposes, when the short-term expectation of the load is still below the observed survived load.

The presented approach is currently being extended to come to a more comprehensive and realistic assessment and updating for the piping mechanism. Aspects to be included in the near future are combined failure due to uplift and piping, geological anomalies (probability of weak spots), spatial variability or more specific observations regarding the occurrence of uplift and piping mechanisms (e.g. water and sand boils). Furthermore, the approach is not restricted to survival observations, is also suitable for any kind of measurement or observation, for which we can establish a likelihood function of the resistance variables.

6 References