Degradation of grass revetments: a comparison of field observations and structured expert judgement

Wouter Jan Klerk1,2, Radityo Andjaringrat Adhi3

1 Dept. of Hydraulic Engineering, Faculty of Civil Engineering & Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, Netherlands
2 Deltares, Boussinesqweg 1, 2629 HD, Delft, Netherlands
3 Golder Associates, Jl. Denpasar Raya Block D2, 12940, Jakarta, Indonesia

Abstract. Many flood defences in areas with a mild wave climate and lack of strong currents are covered with a grass revetment, typically on top of an erosion-resistant (clay) layer. The grass revetment increases the erosion-resistance due to its dense root structure, and experimental overtopping and wave impact tests have shown that it significantly increases the resistance to wave impact, run-up and overtopping loads. However, grass is a natural material, and especially when exposed to loads from weather, storms and human influences, assessment of its deterioration rate is difficult. Therefore, a Structured Expert Judgement study was carried out in which experts, mostly dike inspectors, were asked to estimate rates of deterioration for different types of grass revetments. The experts were weighted based on their performance on seed questions. In this paper we compare the results to damage observations in the Rhine delta which have been obtained during routine inspections. From this comparison it is found that degradation rates of grass revetments vary greatly for different locations, which is a probable cause of the considerable variance found from the Structured Expert Judgement study. Using an illustrative case study, we show the implications for inspection and maintenance of grass revetments.

1 Introduction

Flood defences offer protection to many deltas around the world. After construction these are, depending on the country and context, managed and maintained by different types of asset management organizations, which are tasked with keeping the flood defences in the desired state through inspection and maintenance. For protection against waves and strong currents, for instance in coastal areas various types of revetments are used, such as pattern-placed block revetments, rip rap or asphalt revetments. In areas with a mild wave climate and lack of strong currents, the outer protection of flood defences can often be ensured by a grass revetment. However, the daily maintenance and inspection of these revetments is a significant part of the overall maintenance task, as various external factors can damage the revetment and jeopardize its capacity to protect from waves and currents. Examples of such external factors are animal burrowing, drought damage, damage from vehicles and weeds. Each of these factors can result in a degradation of the quality of the grass sod, and with that its capacity to protect from erosion.

In the reliability assessment of flood defences such factors are often not accounted for and a relatively idealized revetment, without major damage is assumed to be present. This might lead to overly optimistic reliability estimates, although it has to be noted that not every damage at every location has an influence on the safety of the flood defence. However, to further assess whether such damages might have a relevant influence on the reliability of flood defences, both knowledge on the effect of damage on the strength of the revetment, and knowledge on the frequency of occurrence is desired.

In this paper we compare two estimates of the frequency of occurrence. First we present results from a Structured Expert Judgement study conducted among several experts and asset managers of grass revetments, where the frequency of damage and the probability of degradation was assessed. Secondly we compare these results to a dataset of 6 years of visual inspection results of 450 kilometres of flood defence in the Dutch Rhine Delta. This enables us to estimate the order of magnitude of the impact of these damages to the reliability of flood defences in areas with a mild wave climate. Section 2 outlines the general approach and background on assessment of damage and grass sod quality as well as for expert judgment. Section 3 presents results from the expert judgement study, and Section 4 presents the comparison with observational data. Implications for risk-based maintenance and reliability assessment are presented in Section 4.4.
2 Approach

2.1 Assessment of the quality of grass revetments

The capacity of grass revetments to withstand waves depends on the density of the root structure of the grass sod. In the Netherlands this is categorized in 3 categories which can be assessed by manually testing (by pulling) a piece of sod. If the sod cannot be pulled apart, the sod is categorized as closed, of the sod can be pulled apart with relatively minor effort, the sod is categorized as open and if the sod falls apart without pulling, the sod is categorized as fragmented. Fragmented sods have limited to no strength to withstand waves and currents, whereas closed sods typically offer sufficient protection for river flood defences.

In inspections, inspectors do not explicitly assess the sod quality, but assess damages using the Dutch guideline ‘Digigids’, which enlists various damage parameters (burrows, weeds, tracks, bare spots etc.), which can be categorized as good, reasonable, mediocre and bad (Het Waterschapshuis 2016; Klerk et al. 2019). Typically, inspectors do not register points that are categorized as good, but spots with a lower category have to be registered in the inspection application on their phone or tablet, and are thus part of the dataset. The dataset that we consider contains georeferenced points with a specified parameter, categorization, dimensions of the damage and further remarks, as well as pictures made of the damage spot (not in all cases). This dataset has been gathered in the annual spring inspection, so for each year we have 1 point of measurement, usually somewhere in February or March.

From a reliability perspective, in a safety assessment the sod quality for a certain dike section (e.g., of 1 kilometre length) is used, while in inspections both small spots (e.g., a single animal burrow) and larger damages (e.g., weeds over a longer stretch) are registered. However, from a reliability perspective, assuming that loads from waves and currents are equal along the section, the weakest spots determine the reliability.

Henceforth, in this analysis we assume that a damage for a single spot, determines the sod quality for a larger section.

2.2 Approach for expert judgement

The aim of Structured Expert Judgment (SEJ) is to obtain subjective probabilities, and has been derived from the theory of rational decision making in Savage (1954). The subjective probability of a person is an actual judgment, typically representing the way a person thinks, in view of their knowledge and information (Jeffrey 2004). The fundamental assumption of the classical model of expert judgment as formulated by Cooke (Cooke 1991) is that reliability or statistical accuracy of expert’s opinions for unknown variables, can be measured by the reliability or statistical accuracy of their opinions on known situations. In this method, the weight of the experts is calibrated based on their response to seed questions with known answers. Based on these answers their competence and their way of expressing uncertainties can be quantified.

In this study the degradation of grass sod quality is the main target variable. The challenge in this case is to derive seed variables that are related to the target variable. Unfortunately, at the time of this study there was no quantitative data available on other cases of grass degradation. Nevertheless, the capacity of assessing uncertainty can also be determined using other related questions. In this study we use questions on the quality of a sod cover in a specific picture (see Figure 1 for an example), and the degradation time of grass revetments in wave overtopping experiments.

In the first type of seed question, experts are asked to evaluate a sod condition through a picture using a range 1-9 (where 1 indicates fragmented sod and 9 indicates a closed sod). In the second group of seed questions, experts are given information about wave overtopping experiments that evaluate grass conditions through a series of wave attacks (Steendam et al. 2012). Experts are given three pieces of information; a photo of the grass condition at the beginning of the experiment, the wave loads during the experiment, and a photo of the grass condition at a certain point of the experiment. Experts are asked to estimate the point in time at which the experiment the second photo was taken.

![Figure 1 Example of an open sod. Similar pictures were used for the seed questions for the expert judgement study](image)

The goal of the expert judgment study is to elicit the degradation rate of a grass revetment. Thus, the target variables focus on the degradation of a particular grass revetment in a period of time. As typically revetments are inspected both in autumn and in spring, a time period of 6 months was used. So the general target is to determine the probability that a closed sod has degraded to an open or fragmented sod after a 6-month period.

In order to allow for cross-validation between answers, different units of reference, that can each be translated to a probability have been used for expressing this probability in different questions:

- The time period it takes for a grass cover to degrade.
- The probability that a grass cover degrades in a six-month winter period.
- The number of sections out of 100 sections that will be degraded after 6 winter months.
Most of the questions were asked for a flood defence in a rural area with a mild wave climate. To assess the difference with urban areas, part of the questions were also posed for a densely populated area.

By using Cooke’s Classical Method, experts are asked to express their uncertainty with a representation of 5th, 50th, and 95th percentile. The elicitation is measured through the calibration and information score. The calibration score measures the statistical likelihood that the actual values of the seed variables correspond to the expert assessment. A low calibration score means that expert estimations deviate from the actual value. Information measures the degree to which the uncertainty distributions of an expert are concentrated compared to other experts. A narrow bound refers to a confident estimation where a larger bound refers to a larger bound of uncertainty. Both calibration and information scores are combined into a combined expert assessment called Decision Maker (DM) (Aspinall 2008).

The combination of all expert assessments can be combined into one combined estimate through different weighting schemes: Equal weight DM, Global weight DM, Item weight DM and Optimized DM. (Cooke and Goossens 2000). Equal weight DM assigns an equal weight to all experts regardless of their performance on the seed questions. Global weight uses non-normalized weights depending on expert performance and assumes that all seed questions have the same weight. Item weight is a weighting scheme where different seed questions are weighted differently. Optimized DM maximizes the global non-normalized weight of the DM.

3 Expert judgement results

3.1 Results of seed questions

The expert judgement questionnaire was handed out to 7 experts in the Netherlands. These experts are experienced in both operational asset management of grass revetments and research and consultancy in the field of grass revetments. All experts have at least a year of experience in the related subjects.

The expert uncertainty is processed by using a toolbox for structured expert judgment in MATLAB called ANDURIL (Leontaris and Morales-Nápoles 2018). ANDURIL is developed to support quantification of uncertainty by applying Cooke’s classical model for structured expert judgment. Table 1 shows calibration and relative information score for each expert as well as the different DM’s. Expert 3 is found to be the best expert with high calibration and information scores. Relatively speaking it is found that most experts are more confident on the target questions compared to the target and seed questions together (i.e., total information score is higher). Based on the feedback from experts, there were no problems with the seed questions nor the target questions. In conclusion, the seed questions do distinguish which experts perform well both in the accuracy of their estimate (calibration) and in their uncertainty estimates (information). After scoring and combing expert opinions, the two best DM are the DM Equal (Calibration: 0.493), and DM Optimized (Calibration: 0.707) due to their relatively high calibration scores. It must be noted that DM Optimized is equal to Expert 3, i.e. Expert 3 has a very high weight. Further, DM Equal has a relatively low information score which indicates that equally weighted experts result in a lot of uncertainty in weighted answers to the seed questions.

<table>
<thead>
<tr>
<th>Expert ID</th>
<th>Calibration</th>
<th>Relative Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Seed questions</td>
</tr>
<tr>
<td>Exp. 1</td>
<td>8.92E-06</td>
<td>1.314</td>
</tr>
<tr>
<td>Exp. 2</td>
<td>3.22E-05</td>
<td>2.699</td>
</tr>
<tr>
<td>Exp. 3/DM Opt</td>
<td>0.707</td>
<td>2.190</td>
</tr>
<tr>
<td>Exp. 4</td>
<td>1.17E-07</td>
<td>1.785</td>
</tr>
<tr>
<td>Exp. 5</td>
<td>0.014</td>
<td>1.114</td>
</tr>
<tr>
<td>Exp. 6</td>
<td>0.006</td>
<td>0.752</td>
</tr>
<tr>
<td>Exp. 7</td>
<td>0.006</td>
<td>0.890</td>
</tr>
<tr>
<td>Equal Weight</td>
<td>0.493</td>
<td>0.243</td>
</tr>
<tr>
<td>Global Weight</td>
<td>0.314</td>
<td>2.144</td>
</tr>
<tr>
<td>Item Weight</td>
<td>0.314</td>
<td>2.144</td>
</tr>
</tbody>
</table>

Table 1 Expert performance for each expert including calibration and relative information score.

In order to ensure that the selection of seed questions and weight factors for experts are robust, both itemwise and expertwise robustness were investigated. For itemwise robustness the influence of excluding individual seed questions on calibration and information scores was investigated. It was found that only excluding seed question 8 had some positive effect on calibration and information scores. For expertwise robustness excluding either experts 3 or 5 has a large effect on the overall score as both these experts contribute heavily to the overall scores. Hence, the elicitation is robust against the choice of seed items but has significant dependence on experts 3 and 5.

3.2 Expert estimates for degradation

The target variables are also evaluated using ANDURIL. In order to cross-reference and assess the consistency of expert answers, different units of reference were used in the formulation of the target questions, namely:

- Time (number of months) until a revetment has been degraded.
- Probability that a revetment degrades in a 6-month winter season.
- Number of sections at a flood defence consisting of 100 sections that degrades in a 6-month winter season.
Some of these questions focused on degradation of grass revetments due to spot damage with and without maintenance, some focused on general degradation without maintenance. The latter are not considered any further as these do not match the data to which the findings will be compared. For each unit of reference, the values for degradation from closed to open and closed to fragmented sod were elicited. Additionally, for the number of sections, expert estimates were obtained for both urban and rural areas.

Results for DM Optimized and DM Equal are shown in Table 2. Here it is found that different units of reference result in significantly different estimates. For instance, for DM Optimized, the time reference results in a significantly larger degradation probability than the probability and dike section references. This is most likely a reflection of the uncertainty of the experts in providing these estimates. This can be caused both by the fact that it is very hard to estimate degradation rates, and by the fact that these can vary significantly among different areas. It should also be noted that for ‘degradation time’ slightly different results are to be expected as the other questions were limited to degradation in winter months, whereas the time reference also covers summer months (which have a different degradation behaviour, but typically also lower rates). From the comparison between values for rural and urban areas it is found that experts estimate that in urban areas flood defences will degrade faster than in rural areas. In the next chapter we will compare this to the observation data.

For the comparison with data in the following section we will mostly use the results for DM Equal as the relative uncertainty among experts is better reflected than for DM Optimized, which is entirely dominated by expert 3.

### 4 Comparison with inspection data

#### 4.1 Description of available data

Throughout the year several types of flood defence inspections are carried out. These can typically be distinguished in: spring and autumn inspections to assess the condition before and after the winter season, emergency inspections during or after extreme water levels and storms, and theme inspections aimed at specific types of damage (e.g., detection of drought cracks) (CIRIA 2013). Additionally, flood defence asset managers do relatively general inspections (e.g., by car) almost every week. In this study we only consider data from spring inspections, as this is the inspection that is mostly focused on determining the quality of the grass cover.

For our analysis we use inspection records from 472 km of primary flood defences in the Dutch Rhine area. These records have been obtained in the years 2013-2019, data for 2016 is not available. The inspections have been conducted using the Digigids as a guideline (Het Waterschapshuis 2016).

In this study we focus on damage to the grass revetment and relate these to the quality of the grass sod in time. In total there are 3658 registrations of damage to the grass revetment, where 1411 have bad as classification, 1108 mediocre, 1118 reasonable, and 21 have good or unknown classification. It has to be noted that in some areas inspection data can be missing as no inspections were carried out due to execution of dike reinforcement works. However, these areas are relatively small compared to the total portfolio of primary flood defences considered.

Table 3 provides an overview of the different parameters considered, including a description and the number of records in the dataset.

<table>
<thead>
<tr>
<th>Damage parameter</th>
<th>Description</th>
<th>Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal burrowing</td>
<td>Small and large damage due to burrowing by various animals</td>
<td>1189</td>
</tr>
<tr>
<td>Rutting</td>
<td>Damage due to for instance vehicles and preferential walking routes along the slope</td>
<td>844</td>
</tr>
<tr>
<td>Bare spots</td>
<td>Open spots in the grass cover</td>
<td>494</td>
</tr>
<tr>
<td>Weeds</td>
<td>Growth of weeds with a non-erosion resistant root structure</td>
<td>432</td>
</tr>
<tr>
<td>Rough/woody vegetation</td>
<td>Growth of bushes/woody vegetation potentially causing discontinuity of grass cover</td>
<td>421</td>
</tr>
<tr>
<td>Damage by livestock</td>
<td>Grass cover damage caused by (grazing) livestock</td>
<td>93</td>
</tr>
<tr>
<td>Covering</td>
<td>Covering of the grass by other material potentially causing asphyxiation</td>
<td>68</td>
</tr>
<tr>
<td>Sod strength &amp; degree of coverage</td>
<td>Assessment of the strength of the grass sod or degree of coverage</td>
<td>39</td>
</tr>
</tbody>
</table>

Table 3 Overview of damage parameters in the dataset, including a description of the type of damage and the number of records.
4.2 Approach for deriving degradation from data

The first step is to translate the classification made in the inspection to a sod quality. Based on the definitions given in the Digigids for all but 1 parameter (erosion damage) reasonable and good align with a closed sod, mediocre aligns with an open sod, and bad with a fragmented sod. For pragmatic reasons it was decided to use this translation for all parameters.

Next our goal is to derive the degradation rate for a section of a certain length. Therefore we split the flood defences in sections of 200 meters (total of 2379 sections) and couple all damage records to these sections. Note that this does not mean that it is exactly the same damage (spot), but we assume that due to the different damages the sod quality for the section should be considered the same as what follows from the damage classification. In that context the location of the damage spot is not important. It should be noted that from a risk perspective the location of the damage spot is not important. It

Additionally, two separate datasets have been made for dike sections that are in urban areas (i.e., vicinity of cities or villages) and sections that are in a rural area. This allows for analysis of the difference between urban and rural flood defences. This has been done by distinguishing sections based on whether they are within a 100 meter radius of population centres (cities/villages) according to the 2011 dataset of Statistics Netherlands (CBS, , obtainable via https://geodata.nationaalgeoregister.nl/bevolkingskernen2011/wfs). In our dataset 673 sections are located in urban areas, 1706 are in rural areas.

4.3 Comparison of data and expert judgement results

Based on the data we can derive degradation rates per section per year. These are shown in Figure 2 in the left pane. Bars denote the overall average probability of a closed sod degrading to an open or fragmented sod. Dots indicate the probability for the 4 main damage causes. It is found that burrowing, rutting and bare spots contribute almost evenly to overall degradation. Comparison with expert estimates shows that typically expert estimates are slightly higher, but that for the section reference the estimates align quite well. However, the variations indicated by the expert, as well as the encountered differences between experts might be explained by spatial differences in the Rhine area (i.e., areas with high rates of damage and low rates of damage). It should be noted that experts estimate the rate of degrading from closed to open as higher than from closed to fragmented, while in practice it is the opposite.

The variation between different subsets of dike sections is further investigated in Figure 3. Here the violin plot in the left pane indicates the distribution of the degradation probability for 1000 randomly sampled subsets of 100 sections. The right pane shows the ranges provided by individual experts and for the different decision makers. It is found that for degradation to open sod expert estimates are slightly conservative, but for degradation to fragmented sod most experts align well with the distributions obtained from data. When comparing the estimates to the data, expert 7 is found to be quite accurate, while in the calibration of the expert judgement expert 7 was assigned a very low weight based on the seed questions. This might indicate that the seed questions might not have completely reflected the quality of expert estimates in this study.

Based on the subsets of data for urban and rural areas, we can also compare the probability of damage through different causes for each area. This is shown in Figure 4 for the 4 main causes of damage. It is found that rutting and burrowing occur more frequently in rural areas. It is expected that burrowing occurs more frequently in more remote and peaceful areas, but for rutting this is a bit counterintuitive as for instance traffic density is typically higher in urban areas. On the other hand, structural counter measures might prevent damage in such areas (e.g., better
protection of the grass next to the road). Bare spots however can also be caused by increased human activity (e.g., pedestrians) and these do occur more often in urban areas. Overall the transition probability for closed to open sod is approximately the same for both areas, for closed to fragmented urban areas have a slightly lower probability of degrading: 5.2% versus 6.0% for rural areas, mostly due to the large increase in burrowing.

The analysis presented here uses a section length of 200 meters as basis. However, longer sections have a larger probability of damage occurring in some spot, as well as some damage spot not being repaired properly. Thus, the choice of section length has a direct impact on the probabilities obtained. **Figure 5** shows the probability of degradation per section per year relative to the length of the section. Colored lines indicate the values based on the dataset. We find that the probability of degrading from closed to fragmented increases with section length, although the relative increase becomes smaller for larger sections. For degradation to open sod the probability of degradation becomes smaller, which is likely caused by the fact that the probability of not having an inspection registration that translates to a fragmented sod becomes smaller for longer sections. In such cases the weakest spot counts. For even longer sections the probability of having an open sod can even decrease.

The dashed black line in **Figure 5** is a theoretical approximation for degradation to a fragmented sod if we assume that a section consists of statistically independent sections of $L_{ind} = 200$ meters. We can then compute the probability for an arbitrary length $L$ through:
\[ P(C \rightarrow F, L) = 1 - (1 - P(C \rightarrow F, L_{ind}))^{L_{ind}}. \]

where \( P(C \rightarrow F, L) \) is the annual probability of degrading from closed to fragmented sod for a section of length \( L \). For shorter sections this aligns well, but for longer sections there is a significant difference with the values obtained from the dataset. There are a couple of explanations for this, but one of the most important ones might be that damage can occur due to causes that have stronger spatial correlation. An example would be damage to the sod due to a high river water level or drought. If correlation between sections is taken into account, the theoretical curve will be closer to the one obtained from data. This can be further investigated by investigating the spatial correlation for different types of damage.

\[ P_i = \sum_s P(s) \cdot P(F|s). \]

where \( P(s) \) is the probability of sod quality \( s \) being present, and \( P(F|s) \) the failure probability given sod quality \( s \). Probabilities \( P(s) \) are obtained from the data analysed in previous chapters. Note that these values are conditional on a closed sod being present in the year before.

To illustrate this, we consider a section of 200 meters, for which the target failure probability is 1/1000. Typically, an assessment of this section would be based on the failure probability under the assumption that the sod is closed. We assume two cases: where the failure probabilities with closed sod are 1/2000 and 1/3000 per year. We assume that the failure probability increases by a factor 5 if the sod becomes open, and we vary the factor of increase for a fragmented sod between 5 and 100, this is shown in Figure 6. We see that for the case with failure probability of 1/2000 for a closed sod, the requirement is no longer satisfied if the failure probability increase for a fragmented sod is larger than 15. For the case with 1/3000 this value is approximately 33. This shows that whether damage is of relevance for flood defence reliability strongly depends on the relative strength decrease due to damage. This can depend strongly on the type of damage and the location along the slope (van Bergeijk et al. 2021).

Therefore, this is an important aspect to take into account in registering damage in the context of risk-based maintenance and inspection. Overall, in many places damage might not really be a safety issue, whereas in other places more accurate and frequent inspections might be required. In such cases, even if in a certain year it is determined that the sod is closed, for estimates of the failure probability in the future degradation should be considered.

\[ \text{Figure 5 Relation between probability of degradation and length of the considered sections.} \]

4.4 Relevance of findings for risk-based maintenance

In this analysis we have explored the probability that sod quality of a grass revetment along the Dutch Rhine river degrades in a year. The data that was used consists of a large dataset of various damages to the grass cover layer. It has to be noted that not all damages have equal consequences for the failure probability of these flood defences. For instance: rutting damage at the crest of a flood defence with a relatively large residual height will not have much effect on overall failure probability, while it might have if the height is already barely sufficient. However, typically the consequences of such damage are not included in failure probability assessments.

If the sod is on top of an erosion-resistant clay layer, the consequences might be minor, especially if the damage is shallow (e.g., no major burrowing), however if the grass cover is on top of a sand layer, rapid erosion might occur if the sod is not providing proper protection. The actual failure probability given the probabilities of different sod qualities \( s \) (i.e., closed, open and fragmented) can be obtained using the following formula:

\[ \text{Figure 6 Graph showing the relation between the actual failure probability in relation to the relative increase in failure probability if the sod quality becomes fragmented. Two lines indicate different cases with different initial failure probabilities assuming a closed sod.} \]

5 Conclusions

Often reliability estimates for flood defence structures do not consider most of the damage types that might occur
in the field. In this paper we compared expert estimates for the deterioration of grass revetments in a riverine area with a dataset obtained from field inspections. The expert estimates were obtained from a questionnaire among several experts and flood defence asset managers, using Cooke’s method. It was found that expert estimates have large uncertainty, but average estimates align rather well with the data considered. The data analysis showed that the most frequently registered types of damage are burrowing, rutting, bare spots and weeds. From a comparison between urban and rural flood defences it was found that in urban areas more bare spots occur, possibly due to the more intensive use of the slope in general. In rural areas rutting and burrowing are encountered more frequently. For rutting, this might be explained by the fact that in urban areas protection of the roadside and stairs on the slope are typically present, which is not the case in rural areas.

An important aspect in the analysis of damage is the spatial distribution and correlation. It was found that there is a large dependence of degradation rates on the length of the flood defence sections considered. However, assuming independent sections results in a significant overestimation of the degradation rate, which suggests that there is spatial correlation between damages. Further investigating spatial correlation for different types of damage can be a promising avenue for further research which can shed light on why damages occur in certain areas, and be used to derive more specific degradation rates. Such specific degradation rates can then be used as input for determining at which locations degradation is problematic for flood defence safety, and thus which flood defence sections require more intensive risk-based maintenance and inspection.

6 References