Coping with uncertainty in the economical optimization of a dike design

P.H.A.J.M. VAN GELDER, J.K. VRIJLING, K.A.H. SLIJKHUIS
Delft University of Technology, Faculty of Civil Engineering
Delft, The Netherlands

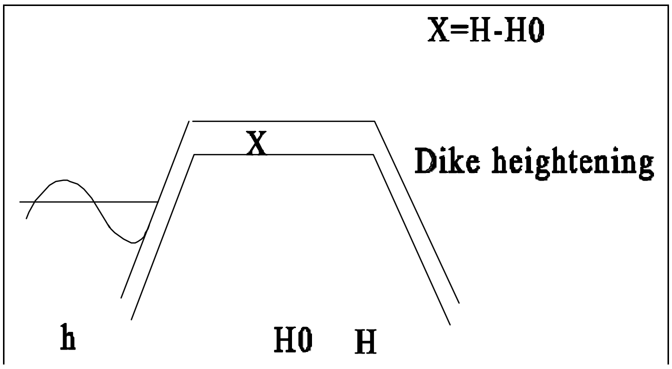
ABSTRACT
A design philosophy for dikes based on an economical cost model is described and examined on its sensitivity with respect to statistical- and model uncertainty. The philosophy is applied on an example of a dike along the Dutch coast.

1 INTRODUCTION

The Netherlands is a unique country by the fact that it can continue to exist thanks to its sea- and river dikes. Without these water retaining structures 2/3 of The Netherlands would be inundated quite regularly and most of its population should have to move elsewhere. In the past dikes were designed by building them as high as the highest known water level at that particular location. More recently other design philosophies have been developed. The approach where a dike has to withstand a certain water level with a fixed probability per year has become quite common. Probabilities of exceedances for water levels once per 10,000 years are adopted for the Dutch sea dikes and once per 1,250 years for the Dutch river dikes [Van Gelder, 1996a]. Other design philosophies have been suggested but didn't become popular so far. However, the approach in which the dike design is determined by an economical optimization [Van Dantzig, 1956] has much advantages in comparison with other design philosophies. This approach will be explained in section 2 of this paper and applied to a location along the Dutch coast. It will be shown how the approach can deal with statistical- and model uncertainties in sections 3.1 and 3.2 respectively. In section 4 the influence of these uncertainties will be examined on the economical optimal dike height. Finally the conclusions are drawn in section 5.

2 ECONOMICAL OPTIMIZATION OF THE DIKE HEIGHT

Taking account of the cost of dike building, of the material losses when a dike-break occurs, and of the frequency distribution of different sea levels, the optimal dike height can be determined by economical optimization [Van Dantzig, 1956]. Assume that H_0 is the current dike level and that we want to determine the amount X by which the dikes must be heightened to the height H (see figure 1).



Let h at any moment denote the sea level along the dikes, then no loss is incurred as long as $h \leq H$; if $h > H$ then we assume a loss with an amount of W including migration costs of the population and cattle, privation of production, damage to houses, buildings, industry etc.

Figure 1: Schematic overview

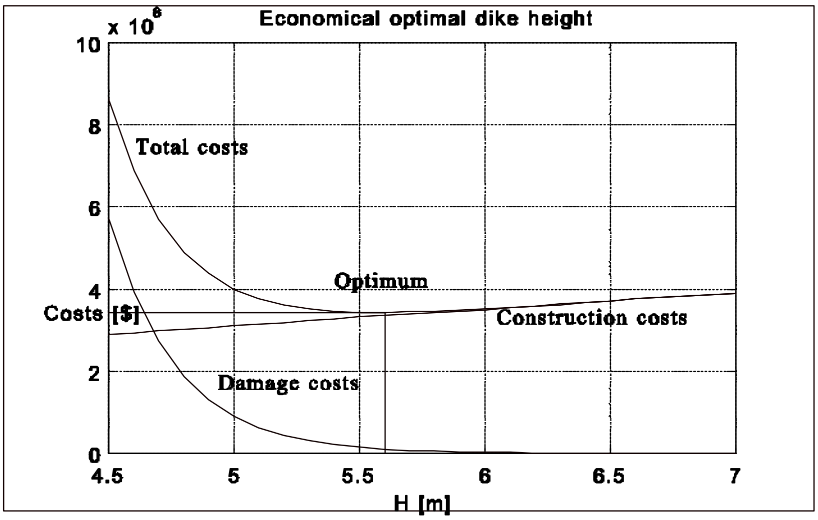
The probability distribution of the sea level h will be denoted by F(h) and will be subject of discussion in sections 3.1 and 3.2. With a dike height of H, each year the expected loss is given by (1-F(H))W. If we assume a constant rate of interest δ the expected value of all future losses are given by:

$$R = F(H)W\Sigma_{t=0}^{\infty}(1+\delta)^{-t} = F(H)W/\delta$$

The total costs of heightening a dike will be assumed linear:

$$I = I_0 + I' X$$

where I_0 is the initial cost and I' the subsequent cost of heightening per meter.
The economical optimal dike height follows by minimizing the expression R+I over the variable H (or X).



The design method is applied to the Hook of Holland dike along the Dutch coast near the port of Rotterdam in [Slijkhuis, 1996] with the following values for the above given variables:
$W = \$ 24.2 \times 10^9$
$I_0 = \$ 110 \times 10^6$
$I' = \$ 40 \times 10^6$
$\delta = 1.5\%$
F = Gumbel

Figure 2: The economical optimization lines

The economical optimal dike height follows to be H_{opt} = 5.60m with a corresponding probability of failure of $F(H_{opt}) = 6.43 \times 10^{-6}$ per year.

3 UNCERTAINTIES

In the design of dikes attention should be paid to the statistical - and model uncertainties of the sea levels. The analysis of estimating sea level exceedance probabilities has always been a controversy in the literature. The procedure that is traditionally followed:
(1) observe a historical record of sea levels
(2) pick a probability density function that seems reasonable
(3) estimate the parameters of the pdf from the historical records
(4) make inferences about theoccurrencee of extreme sea levels
In step (1) one usually has to deal with a limited amount of data. Although the sea levels along the Dutch coast are measured quite intensively for more than a century now and also flood historical data is available [Van Gelder, 1996b], the amount of data to predict sea levels with return periods of 1000-10000 years is very limited. For the Hook of Holland dike a data set of year maxima of sea levels in the period 1887 - 1995 is available . This dataset is corrected for the sea level rise (of 20 cm per 100 years along the Dutch coast) and since it consists of year maxima, it can be considered a homogeneous dataset. In steps (2) and (3) there exist a number of uncertainties under which:
- Statistical- or parameter uncertainty which is associated with the estimation of the parameters of the model of the stochastic process due to limited data.
- Model uncertainty which is associated with the uncertainty that a particular probabilistic model of the stochastic process may not be the true or best model.

3.1 PARAMETER UNCERTAINTY

A parameter θ of a distribution function is estimated from the data. It is therefore a function f of the data and because the data is a random variable, the parameter itself is also a random variable. The parameter uncertainty can be described by the distribution function of the parameter. In [Slijkhuis, 1996] an overview is given of the analytical and numerical derivation of parameter uncertainties for certain probability models (Exponential, Gumbel and Log-normal). One of her conclusions was that a bootstrapping method is a fairly easy tool to calculate the parameter uncertainty numerically. Other methods to model parameter uncertainties like Bayesian methods can be very well applied too ([Van Gelder, 1996c]). Bootstrapping methods are described in for example [Efron et.al., 1993]. Given a dataset $x=(x_1,x_2,...,x_n)$, we can generate a bootstrap sample x^* which is a random sample of size n drawn with replacement from the dataset x. The following bootstrap algorithm has been used for estimating the parameter uncertainty:
1. Select B independent bootstrap samples $x^{*1}, x^{*2}, ..., x^{*B}$, each consisting of n data values drawn with replacement from x.
2. Evaluate the bootstrap evaluation corresponding to each bootstrap sample;

$$\theta^*(b)=f(x^{*b}) \text{ for } b=1,2,...,B.$$

3. Determine the parameter uncertainty by the empirical distribution function of θ*.

The algorithm has been applied to the location parameter A and scale parameter B of the Gumbel distribution with an MaxLik-fit to the Hook of Holland data. The following distributions and approximating normal distributions were obtained (figure 3):

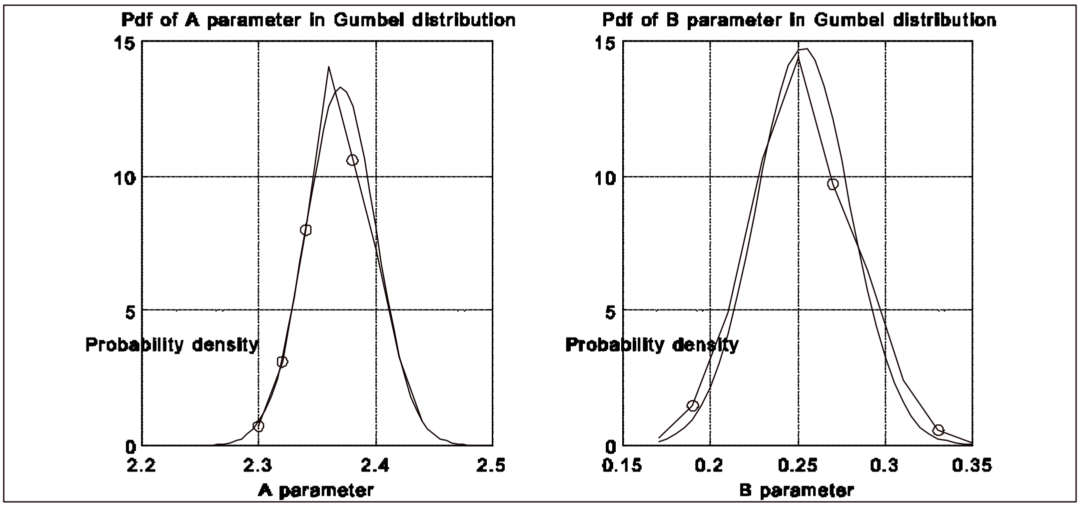


Figure 3: Parameter uncertainties in the Gumbel distribution

The influence of the parameter uncertainty on the frequency line for the sea levels year maxima is examined in the next figure. The 5 lines correspond with values for the location and scale parameters as given in table 1.

Line	A	B
1	μ	μ-2σ
2	μ-2σ	μ
3	μ	μ
4	μ+2σ	μ
5	μ	μ+2σ

Table 1

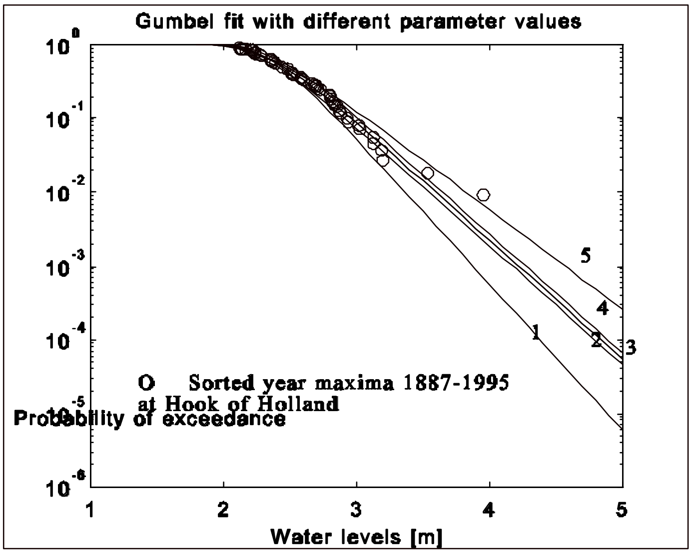


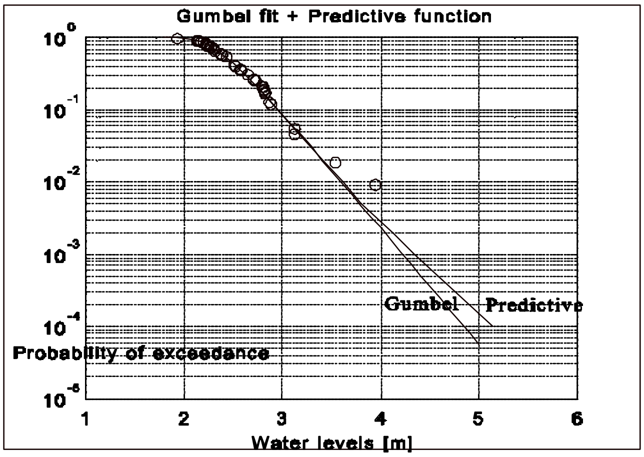
Figure 4: Influence of parameter uncertainty on Gumbel fit

Note the high sensitivity of the frequency lines in Figure 4.

If we are involved in calculating the expected frequency line for the sea levels h, then the inferences we make on h should reflect the uncertainty in the parameters θ. In the Bayesian terminology we are interested in the so-called predictive function:

$$F(h)=\int_{\theta} F(h|\theta)\, f(\theta)d\theta$$

where $F(h|\theta)$ is the probabilistic model of sea levels, conditional upon the parameters θ and F(h) is the predictive distribution of the sea levels, now parameter free. In popular words: Athe uncertainty in the θ parameters have to be integrated out@.



The predictive distribution can be interpreted as being the distribution $F(h|\theta)$ weighted by $f(\theta)$. *In making inferences on sea levels it is important to use the predictive function for h,*(fig. 5) as opposed to the probabilistic model for h with some estimator for the parameter set θ, i.e. $f(q|\theta^{*})$. This is because using point estimators for uncertain parameters underestimates the variance in sea levels.

Figure 5: Integrating out the uncertainty

3.2 MODEL UNCERTAINTY

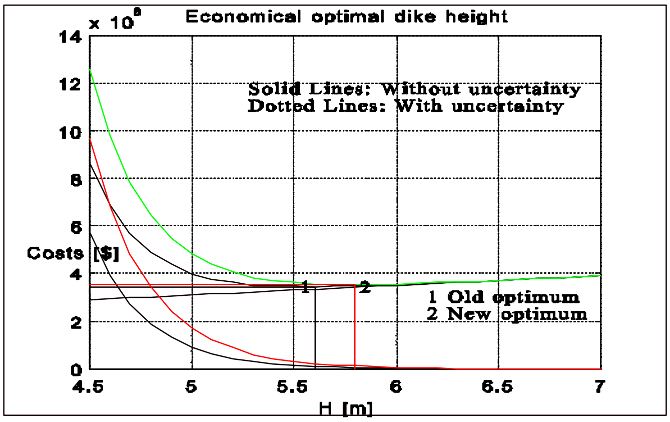
In most cases, the probability model of a physical magnitude cannot be clearly estimated from statistical techniques. It is also seldom possible to derive the exact model theoretically. The correct probability model of a physical magnitude may never be known. Instead of assuming a specific probability model for prediction, it is much better to assume that each model could be potentially correct. Predictions may be made from any probability model. Yet, the model that reveals a wide scatter in the probability plot will have a large standard deviation around its predicted value. Consequently, the prediction from this model should be given less weight relative to those models that exhibit less scatter. Weighting factors can be calculated by techniques described in for example [Pericchi et.al., 1983].For the Hook of Holland data, the following models were fitted and the corresponding weighting factors were calculated:

LogNormal	0.05
LogPearson III	0.46
Gumbel	0.49

Table 2: Weighting factors

4 THE INFLUENCE OF UNCERTAINTY ON THE ECONOMICAL OPTIMIZATION

Rather than performing the economical optimization with the probability distribution F(h|θ), we will perform it with the predictive distribution, where the parameter uncertainty is Aintegrated out@.



An increase in the optimal dike height from 5.60m to 5.80 m is the result when parameter uncertainty in the probabilistic model is taken into account. The corresponding costs almost remain the same however. The influence of model uncertainty on the optimal dike height is rather small, since the Gumbel model already gave a high weighting factor.

Figure 6: Economics with and without parameter uncertainty

5 CONCLUSIONS

In this paper we have studied the influence of statistical- and model uncertainty on an economical design philosophy for dikes. Techniques like Bootstrapping, Bayesian prediction and Model weighting are very suitable for this analysis. An application of the dike height at Hook of Holland is presented where the influence of statistical- and model uncertainties on the economical optimal dike height is in the order of 20cm. The uncertainty in other variables like W, I_0, I: and δ is examined in [Slijkhuis, 1996].

6 REFERENCES

[1993] Efron, B., Tibshirani, R.J., *An Introduction to the Bootstrap*. Chapman & Hall.
[1983] Pericchi, L.R., and Rodriguez-Iturbe, I., *On some problems in Bayesian model choice in hydrology*, The Statistician, Vol.32.
[1996] Slijkhuis, K.A.H., *The influence of statistical uncertainty on the design height*

of dikes [IN DUTCH]. Msc-thesis, Delft University of Technology, October 1996.

[1956] Van Dantzig, D., *Economic decision problems for flood prevention*. Econometrica, 24, pp.276-287, New Haven.

[1996a] Van Gelder, P.H.A.J.M., Roos, A., and Tonneijck, M.R., *On the probabilistic design of dikes in the Netherlands*. Applications of Statistics and Probability, Volume 3, pp.1505-1508, A.A. Balkema.

[1996b] Van Gelder, P.H.A.J.M., *How to deal with wave statistical and model uncertainties in the design of vertical breakwaters*. Probabilistic tools for vertical breakwater design, MAS3-CT95-0041, Workshop Grenoble, 24-25 October 1996.

[1996c] Van Gelder, P.H.A.J.M., *A new statistical model for extreme water levels along the Dutch coast.* Stochastic Hydraulics, pp. 243-250, Mackay.