Bayesian Estimation of Quantiles for the Purpose of Flood Prevention

J.M. van Noortwijk[1] and P.H.A.J.M. van Gelder[2]

Abstract

In this paper, the problem of Bayesian estimation of flood quantiles is studied. Bayes estimators of the optimal dyke height under symmetric and asymmetric loss are investigated when the annual maximum sea water levels are exponentially distributed with unknown value of the mean. Three types of loss functions are considered: (i) linear loss, (ii) squared-error loss, and (iii) linex loss. In order to properly account for the statistical uncertainty in the mean, a modified linex loss function is to be preferred. This new modified linex loss function is derived from the economic dyke heigthening problem of Van Dantzig. Since the loss function is based on a benefit-cost analysis, its parameters have a clear economic significance.

Introduction

In statistical analysis of civil engineering data such as water levels, wave heights, soil parameters, etc., many attempts have been made to establish what kind of fitting method is preferable for the parameter estimation of a probability distribution in order to estimate the *q*-quantile, i.e. the value with a probability of exceedance equal to *q* (where *q* is usually very small, in the order of 10^{-3} to 10^{-5}). Recent research on this subject can be found in, for example, Yamaguchi (1997), Fortin et al. (1997), Burcharth and Liu (1994), and Van Gelder (1996). The main idea in their work is to generate random samples from a chosen probability distribution using Monte Carlo simulation, and to investigate the advantage of a certain parameter estimation method over the other method which is based on the viewpoint of bias and variance of the *q*-quantile. The estimation method with the

[1] HKV Consultants, P.O. Box 2120, 8203 AC Lelystad, The Netherland, and Delft University of Technology, Faculty of Mathematics and Computer Science, P.O. Box 5031, 2600 GA Delft, The Netherlands, E-mail: J.M.van.Noortwijk@hkv.nl.
[2] Delft University of Technology, Faculty of Civil Engineering, P.O. Box 5048, 2600 GA Delft, The Netherlands, E-mail: P.vanGelder@ct.tudelft.nl.

smallest bias and/or variance of the q-quantile is then considered to be the best parameter estimation method for that particular probability distribution. In this paper, a new concept is presented in which the parameter estimation method is defined as such that the bias of the estimated q-quantile is minimised within a Bayesian framework of asymmetric loss functions. The use of asymmetric loss functions gives us the possibility to differentiate between underestimation and overestimation of the q-quantile. In civil engineering applications, an underestimation error of the q-quantile generally leads to much higher losses than an overestimation error.

Although Fortin, Bobée and Bernier (1997) also use asymmetric loss functions for comparing statistical distributions and estimation methods, they do not apply these loss functions in a full Bayesian framework. The parameters of the probability distributions have been estimated by using three well-known methods from classical statistics: the method of maximum likelihood, the method of moments, and the method of probability-weighted moments. The Bayesian point of view only comes in when Fortin et al. (1997) use a nonparametric Bayesian simulation methodology, called Pólya resampling, instead of the classical bootstrap to draw observations with replacement from a reference sample.

Three loss functions have been studied: (i) the asymmetric linear loss function, (ii) the asymmetric squared-error loss function, and (iii) the linex loss function. For an overview of asymmetric loss functions, see Zellner (1986) and Thompson and Basu (1996). The linex loss function has been used in real estate assessment by Varian (1974). Basu and Ebrahimi (1991) determine an expression for the linex estimator of the survival function of a system having a Type II censored exponential lifetime. Using simulated data, Basu and Thompson (1992) and Thompson and Basu (1993) obtain linex estimates of the reliability of simple stress-strength systems. Pandey et al. (1994) study the problem of estimating the shape parameter of a Pareto distribution using a linex loss function.

The Bayes estimator of the q-quantile under asymmetric loss minimises the expected loss with respect to the probability distribution of an unknown parameter. In order to find the loss function that can best be applied to decision problems in civil engineering, we have studied the economic dyke-height optimisation problem of Van Dantzig (1956). He assumed the annual maximum sea water levels to be exponentially distributed. It appears that Van Dantzig's economic loss function differs slightly from the linex loss function. This modified linex loss function seems to be a promising candidate for solving quantile estimation problems in other civil engineering benefit-cost analyses.

The outline of the paper is as follows. First, an overview is given of the Bayesian estimation of quantiles by using the three above-mentioned loss functions. The economic optimisation of dyke heights is subsequently addressed. Next, we derive the relation between the economic optimisation and the loss function approach. This relation results in a modified linex loss function. The different methods are compared in a Dutch polder case study. Finally, conclusions are drawn.

Bayesian Estimation of Quantiles Using Loss Functions

Define the random quantity X_i to be the maximal sea water level in year i, $i = 1,\ldots,n$. We assume the random quantities $X_1,\ldots,X_n$ to be mutually independent, identically distributed, random quantities with a cumulative distribution function $\Pr\{X_i \le x\} = F(x \mid \lambda)$ with parameter λ, $i = 1,\ldots,n$. As a function of λ, the *q*-quantile of the probability distribution of *X* is defined to be

$$y_q = g(\lambda) = F^{-1}(1-q \mid \lambda),$$

where $g'(\lambda) > 0$. Suppose the parameter λ is unknown with a prior probability density function $\pi(\lambda)$. After observing the data $\mathbf{x} = (x_1,\ldots,x_n)$, this prior density can be updated to the posterior density using Bayes' theorem:

$$p(\lambda) = \pi(\lambda \mid \mathbf{x}) \propto l(\mathbf{x} \mid \lambda)\pi(\lambda) = \prod_{i=1}^{n} f(x_i \mid \lambda)\pi(\lambda),$$

where $l(\mathbf{x} \mid \lambda)$ is the likelihood function of the observations $\mathbf{x}$ when the value of λ is given.

For the purpose of flood prevention, we are interested in estimating the *q*-quantile of the probability distribution of the maximal sea water level *X* per year, denoted by $g(\lambda^*)$. In a Bayesian framework, this can be achieved by minimising the loss due to the simple estimation error $\Delta = g(\lambda^*) - g(\lambda)$. Since the loss due to flooding increases with overestimation error (i.e. the real *q*-quantile is less than its estimated value: $g(\lambda) < g(\lambda^*)$ or $\Delta > 0$) and, at a much faster rate, with underestimation error (i.e. the real *q*-quantile is greater than its estimated value: $g(\lambda) > g(\lambda^*)$ or $\Delta < 0$), we focus on asymmetric loss functions (see Thompson and Basu, 1996). Beside loss functions of the simple estimation error, loss functions of the relative estimation error can also be considered.

The three most well-known asymmetric loss functions are: (i) the asymmetric linear loss function, (ii) the asymmetric squared-error loss function, and (iii) the linex loss function.

Asymmetric Linear Loss

The asymmetric linear loss function is defined by

$$L(\Delta) = \begin{cases} a\Delta & \text{if} \quad \Delta \ge 0 \ \text{ or } \ \lambda \le \lambda^*, \\ -b\Delta & \text{if} \quad \Delta < 0 \ \text{ or } \ \lambda > \lambda^*, \end{cases} \qquad (1)$$

where $a, b > 0$. This loss function is asymmetric for $a \ne b$. We can best choose the estimate λ^* for which the expected loss is minimal with respect to the probability distribution of λ:

$$E(L(\Delta)) =$$
$$= \int_{-\infty}^{\lambda^*} a[g(\lambda^*) - g(\lambda)]p(\lambda)\,d\lambda + \int_{\lambda^*}^{\infty} b[g(\lambda) - g(\lambda^*)]p(\lambda)\,d\lambda =$$
$$= ag(\lambda^*)P(\lambda^*) - a\int_{-\infty}^{\lambda^*} g(\lambda)p(\lambda)\,d\lambda - bg(\lambda^*)[1 - P(\lambda^*)] + b\int_{\lambda^*}^{\infty} g(\lambda)p(\lambda)\,d\lambda\,,$$

where $P(\lambda)$ is the cumulative distribution function of λ. The Bayes estimator under asymmetric linear loss is the solution of the equation

$$\frac{dE(L(\Delta))}{d\lambda^*} = g'(\lambda^*)([a+b]P(\lambda^*) - b) = 0,$$

which results in $\lambda^* = P^{-1}(b/[a+b])$. Hence, the Bayes estimator λ^* equals the $b/[a+b]$-quantile of the posterior distribution of λ. When *a* = *b*, the linear loss function is symmetric and its Bayes estimator reduces to the posterior median $P^{-1}(0.5)$.

<u>Asymmetric Squared-Error Loss</u>

The asymmetric squared-error loss function is defined by

$$L(\Delta) = \begin{cases} a\Delta^2 & \text{if} \quad \Delta \geq 0 \quad \text{or} \quad \lambda \leq \lambda^*, \\ b\Delta^2 & \text{if} \quad \Delta < 0 \quad \text{or} \quad \lambda > \lambda^*, \end{cases} \tag{2}$$

where $a, b > 0$. This loss function is asymmetric for $a \neq b$ with expected value

$$E(L(\Delta)) = \int_{-\infty}^{\lambda^*} a[g(\lambda^*) - g(\lambda)]^2 p(\lambda)\,d\lambda + \int_{\lambda^*}^{\infty} b[g(\lambda) - g(\lambda^*)]^2 p(\lambda)\,d\lambda\,.$$

The Bayes estimator under asymmetric squared-error loss is the solution of the equation

$$\frac{dE(L(\Delta))}{d\lambda^*} = 2g(\lambda^*)g'(\lambda^*)\{[a-b]P(\lambda^*) + b\} -$$
$$- 2g'(\lambda^*)\{[a-b]\int_{-\infty}^{\lambda^*} g(\lambda)p(\lambda)\,d\lambda + b\int_{-\infty}^{\infty} g(\lambda)p(\lambda)\,d\lambda\} = 0,$$

which must be solved for λ^* numerically. When *a* = *b*, the squared-error loss function is symmetric and the Bayes estimator $g(\lambda^*)$ reduces to the posterior mean of $g(\lambda)$.

<u>Asymmetric Linex Loss</u>

The asymmetric linex loss function is defined by

$$L(\Delta) = b[a\Delta + \exp\{-a\Delta\} - 1], \qquad (3)$$

where $a, b > 0$. The expected loss can be written as

$$E(L(\Delta)) =$$
$$= b[\int_{-\infty}^{\infty} a[g(\lambda^*) - g(\lambda)]p(\lambda)\,d\lambda + \int_{-\infty}^{\infty} \exp\{-a[g(\lambda^*) - g(\lambda)]\}p(\lambda)\,d\lambda - 1].$$

The Bayes estimator under asymmetric linex loss, λ^*, is the solution of the equation

$$\frac{dE(L(\Delta))}{d\lambda^*} = ab\,g'(\lambda^*)[1 - \int_{-\infty}^{\infty} \exp\{-a[g(\lambda^*) - g(\lambda)]\}p(\lambda)\,d\lambda] = 0,$$

which results in

$$\lambda^* = g^{-1}\left(\frac{\ln(\int_{-\infty}^{\infty} \exp\{ag(\lambda)\}p(\lambda)\,d\lambda)}{a}\right).$$

Examples of Loss Functions

Examples of the three loss functions are displayed in Figure 1: (i) the asymmetric linear loss function with $a = 5.37\ 10^7$ and $b = 1.94\ 10^7$, (ii) the asymmetric squared-error loss function with $a = 1.07\ 10^8$ and $b = 3.88\ 10^7$, and (iii) the asymmetric linex loss function with $a = 3.03$ and $b = 1.32\ 10^7$. The parameters a and b have been chosen as such that the three loss functions are equal for $\Delta = \pm 0.5$.

Note that both linear loss and squared-error loss are special cases of what Thompson and Basu (1996) called monomial-splined loss, defined, for fixed $m = 1,2,3,...$, by

$$L(\Delta) = \begin{cases} a|\Delta|^m & \text{if} \quad \Delta \geq 0, \\ b|\Delta|^m & \text{if} \quad \Delta < 0, \end{cases}$$

where $a, b > 0$. Fortin, Bobée and Bernier (1997) applied monomial-splined loss for $m = 1,2,3$ within a framework of classical statistics.

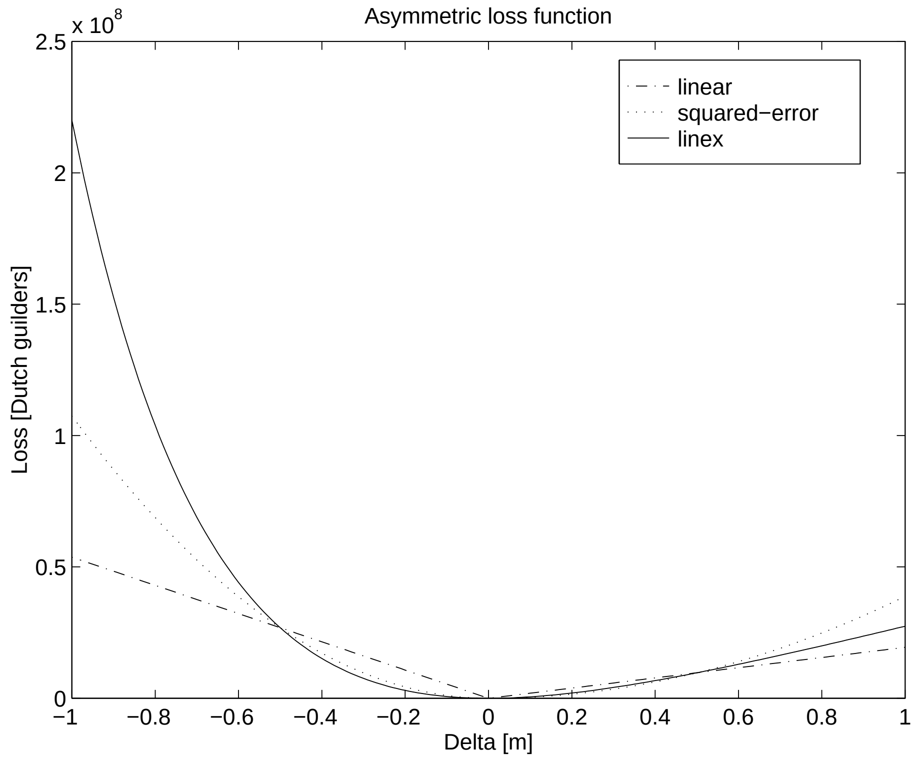

Figure 1. The linear, squared-error and linex loss function according to Eqs. (1-3).

Estimation of Optimal Dyke Height

Let us consider the benefit-cost analysis that is adapted from Van Dantzig (1956). Suppose we have to decide how high the dykes should be to prevent a polder from flooding. Let the height of the dyke *h* be the decision variable, and let $h_0 = 3.25$ metres be the initial height of the dyke at the moment the decision has to be taken. The only failure mechanism that we regard is overtopping, i.e. inundation of the polder will occur as soon as the sea water level exceeds the height of the dyke. To account for the stochastic nature of the sea water level, we assume the maximal sea levels per year X_i, $i = 1,...,n$, to be conditionally independent, exponentially distributed, random quantities with a known location parameter $x_0 = 1.96$ metres and an unknown scale parameter λ with expected value 0.33 metres. Hence, the likelihood function is

$$l(\mathbf{x} \mid \lambda) = \prod_{i=1}^{n} f(x_i \mid \lambda) = \prod_{i=1}^{n} \frac{1}{\lambda} \exp\left\{-\frac{x_i - x_0}{\lambda}\right\}.$$

Accordingly, the *q*-quantile of the probability distribution of *X* is

$y_q = g(\lambda) = x_0 - \lambda \ln(q)$.

The costs of heightening the dykes with $h - h_0$ metres depend on the fixed cost c_f = 1.1 10^8 and the variable cost c_v = 4.0 10^7: i.e. $c_f + c_v[h - h_0]$. If the polder is inundated, an economic value of c = 2.4 10^{10} Dutch guilders is lost. The discount factor is $\alpha = [1 + 0.015]^{-1}$, compounded annually, where $0 < \alpha < 1$. Since the probability of inundation of the polder is $\exp\{-(h - x_0)/\lambda\}$, the expected discounted costs due to inundation of the polder over an unbounded time-horizon can be written as

$$c(\lambda, h) = c_f + c_v[h - h_0] + \frac{\alpha}{1-\alpha} c \exp\left\{-\frac{h - x_0}{\lambda}\right\} \tag{4}$$

if the decision-maker chooses dyke height *h* and when the value of λ is given.

The decision with minimal expected costs, i.e. the dyke height for which the expected discounted costs due to inundation are minimal, is

$$h^* = x_0 - \lambda \ln\left(\lambda \cdot \frac{c_v}{c} \cdot \frac{1-\alpha}{\alpha}\right) \tag{5}$$

when the value of λ is given. Accordingly, the inundation probability that balances the cost of investment optimally against the cost of inundation is

$$q = \exp\left\{-\frac{h^* - x_0}{\lambda}\right\} = \lambda \cdot \frac{c_v}{c} \cdot \frac{1-\alpha}{\alpha}. \tag{6}$$

When the value of λ is given to be 0.33 metres, the optimal inundation probability is q = 8.25 10^{-6}.

To account for the statistical uncertainty in the mean of the maximal sea water level per year, the prior density of λ is assumed to be an inverted gamma distribution with scale parameter $\mu > 0$ and shape parameter $\nu > 0$:

$$\text{Ig}(\lambda \mid \nu, \mu) = [\mu^\nu / \Gamma(\nu)] \lambda^{-(\nu+1)} \exp\{-\mu/\lambda\}$$

for $\lambda > 0$. The prior mean and variance are $E(\lambda) = \mu/(\nu - 1)$ and $\text{Var}(\lambda) = E(\lambda)^2/(\nu - 2)$, respectively. Hence, the larger ν, the less uncertain λ. On the basis of this prior density, the expected discounted costs over an unbounded horizon transform into

$$\int_0^{\infty} c(\lambda,h)p(\lambda)\,\lambda = c_f + c_v[h-h_0] + \frac{\alpha}{1-\alpha}c\left[\frac{\mu}{\mu+h-x_0}\right]^{\nu}. \tag{7}$$

The dyke height with minimal expected costs, while taking the uncertainty in λ into account, is

$$h^{\bullet} = x_0 - \mu + \left[\nu\mu^{\nu}\cdot\frac{c}{c_v}\cdot\frac{\alpha}{1-\alpha}\right]^{\frac{1}{\nu+1}}. \tag{8}$$

The inundation probability that balances the cost of investment optimally against the cost of inundation, while taking the uncertainty in λ into account, is

$$q^{\bullet} = \left[\frac{\mu}{\mu+h^{\bullet}-x_0}\right]^{\nu} = \left[\frac{\mu}{\nu}\cdot\frac{c_v}{c}\cdot\frac{1-\alpha}{\alpha}\right]^{\frac{\nu}{\nu+1}}.$$

When the expected value of λ is 0.33 metres, the optimal inundation probability under statistical uncertainty is $q^{\bullet} = 1.02\ 10^{-5}$ for $\nu = 50$, $q^{\bullet} = 9.17\ 10^{-6}$ for $\nu = 100$, and $q^{\bullet} \rightarrow 8.25\ 10^{-6}$ as $\nu \rightarrow \infty$.

An advantage of the inverted gamma distribution as a prior density is that the posterior distribution of λ, when the observations $x_1,\ldots,x_n$ are given, is also an inverted gamma distribution with scale parameter $\mu + \sum_{i=1}^{n}(x_i - x_0)$ and shape parameter $\nu + n$. The inverted gamma distribution is said to be a conjugate family of distributions for observations from an exponential distribution with unknown mean (scale parameter). From now on, when we use the probability density function $p(\lambda)$, we refer to the posterior density. Note that (inverted) gamma priors have als been applied by Basu and Ebrahimi (1991), Basu and Thompson (1992), Thompson and Basu (1993), and Pandey et al. (1994).

<u>Relation Between Economic Loss and Linex Loss</u>

The question arises whether Van Dantzig's economic cost function and the Bayesian loss function are interrelated to each other. In this respect, we reformulate Van Dantzig's cost function in terms of Bayesian loss, i.e. we rewrite the loss function as

$$\begin{aligned} L(\Delta) &= L(g(\lambda^*) - g(\lambda)) = c(\lambda, h^* + \Delta) - c(\lambda, h^*) = \\ &= c_v\Delta + \frac{\alpha}{1-\alpha}c\exp\left\{-\frac{h^*-x_0}{\lambda}\right\}\left[\exp\left\{-\frac{\Delta}{\lambda}\right\} - 1\right]. \end{aligned}$$

There are now two possibilities for rewriting the probability of exceedence

$\exp\{-(h^* - x_0)/\lambda\}$: (i) as a constant and (ii) as a function of the unknown scale parameter λ.

First, we investigate the probability of exceedence $\exp\{-(h^* - x_0)/\lambda\}$ to be a constant, i.e. to be q = 8.25 10^{-6}:

$$L(\Delta) = c_v\Delta + \frac{\alpha}{1-\alpha}cq\left[\exp\left\{-\frac{\Delta}{\lambda}\right\}-1\right], \tag{9}$$

where $\Delta = g(\lambda^*) - g(\lambda)$ and $g(\lambda) = x_0 - \lambda\ln(q)$. The Bayes estimator under asymmetric loss in terms of Eq. (9), λ^*, is the solution of the equation

$$\frac{dE(L(\Delta))}{d\lambda^*} = g'(\lambda^*)\left[c_v - \frac{\alpha}{1-\alpha}cq\int_0^\infty \frac{1}{\lambda}\exp\left\{-\frac{g(\lambda^*)-g(\lambda)}{\lambda}\right\}p(\lambda)\,d\lambda\right] = 0,$$

which results in

$$g(\lambda^*) = x_0 - \lambda^*\ln(q) = x_0 - \mu + \left[\nu\mu^\nu \cdot \frac{c}{c_v}\cdot\frac{\alpha}{1-\alpha}\right]^{\frac{1}{\nu+1}} = h^\bullet.$$

Second, we consider the probability of exceedence $\exp\{-(h^* - x_0)/\lambda\}$ to be a function of the unknown scale parameter λ, by substituting the optimal dyke height h^* according to Eq. (5):

$$L(\Delta) = c_v\left(\Delta + \lambda\left[\exp\left\{-\frac{\Delta}{\lambda}\right\}-1\right]\right), \tag{10}$$

where $\Delta = g(\lambda^*) - g(\lambda)$ and

$$g(\lambda) = x_0 - \lambda\ln\left(\lambda\cdot\frac{c_v}{c}\cdot\frac{1-\alpha}{\alpha}\right).$$

The Bayes estimator under asymmetric loss in terms of Eq. (10), λ^*, is the solution of the equation

$$\frac{dE(L(\Delta))}{d\lambda^*} = c_v g'(\lambda^*)\left[1 - \int_0^\infty \exp\left\{-\frac{g(\lambda^*)-g(\lambda)}{\lambda}\right\}p(\lambda)\,d\lambda\right] = 0,$$

which results in

$$g(\lambda^*) = x_0 - \lambda^* \ln\left(\lambda^* \cdot \frac{c_v}{c} \cdot \frac{1-\alpha}{\alpha}\right) = x_0 - \mu + \left[\nu\mu^\nu \cdot \frac{c}{c_v} \cdot \frac{\alpha}{1-\alpha}\right]^{\frac{1}{\nu+1}} = h^\bullet .$$

Modified Linex Loss

We ought to notice that the two economic loss functions (9-10) differ slightly from the linex loss function (3). A difference is that both economic loss functions are not only a function of the simple estimation error Δ, but also of the relative estimation error Δ/λ. In terms of λ^* and λ, the loss function (9) can be written as

$$L(\Delta) = -c_v \ln(q) \cdot (\lambda^* - \lambda) + \frac{\alpha}{1-\alpha} cq\left[\exp\left\{\ln(q) \cdot \frac{\lambda^* - \lambda}{\lambda}\right\} - 1\right] = L(\Delta_1, \Delta_2), \qquad (11)$$

where $\Delta_1 = \lambda^* - \lambda$ is the simple estimation error of λ^* and $\Delta_2 = (\lambda^* - \lambda)/\lambda$ is the relative estimation error of λ^*. The general formulation of the modified linex loss function (11) is:

$$L(\Delta_1, \Delta_2) = b(a\Delta_1 + d[\exp\{-a\Delta_2\} - 1]),$$

where

$$a = -\ln(q), \quad b = c_v, \quad d = \frac{\alpha}{1-\alpha} \cdot \frac{c}{c_v} \cdot q .$$

Since the main aim of this paper is estimating the *q*-quantile of a probability distribution, the most appropriate loss functions seem to be the economic loss functions (9) and (11) (in terms of $g(\lambda)$ and λ, respectively). These economic loss functions are modified linex loss functions, for which the parameters have a clear economic significance. The parameters represent the cost of investment (dyke heightening) on the one hand, and the cost of flooding on the other hand. Since the modified linex loss functions are derived from estimating the mean of an exponential distribution, more research has to be undertaken to find out whether they can also be applied to estimate the statistical parameters of other probability distributions.

Comparative Results

On the basis of the dyke heightening problem, we have compared the linear, squared-error and linex loss function with the economic loss functions. The results are summarised in Tables 1-2. The coefficients *a* and *b* of the linear, squared-error and linex loss function have been assessed in the following way. As suggested by the economic loss functions (9-10), the coefficients of the linex loss function are assumed to be $a = [E(\lambda)]^{-1}$ and $b = c_v E(\lambda)$. Furthermore, asymmetric linear and squared-error loss functions have been fitted to this linex loss function by, somewhat arbitrary, assuming the linear, squared-error and linex loss to be equal to each other for $\Delta = \pm 0.5$ (see Figure 1). Results are also presented for the symmetric linear and squared-error loss function.

Table 1: Bayes estimates of the scale parameter λ and the dyke height h for ν = 50.

Estimation method for ν = 50 observations	Eq.	a	b	λ^* [m]	h^* [m]
Van Dantzig without uncertainty	(5)	-	-	0.330	5.82
Van Dantzig with uncertainty	(8)	-	-	-	6.14
Bayes estimate symmetric linear loss	(1)	1	1	0.326	5.77
Bayes estimate symmetric squared-error loss	(2)	1	1	0.330	5.82
Bayes estimate asymmetric linear loss	(1)	$5.37\ 10^7$	$1.94\ 10^7$	0.356	6.13
Bayes estimate asymmetric squared-error loss	(2)	$1.07\ 10^8$	$3.88\ 10^7$	0.350	6.05
Bayes estimate linex loss	(3)	3.03	$1.32\ 10^7$	0.393	6.56
Bayes estimate modified linex loss	(9)	-	-	0.360	6.14
Bayes estimate modified linex loss	(10)	-	-	0.357	6.14

Table 2: Bayes estimates of the scale parameter λ and the dyke height h for ν = 100.

Estimation method for ν = 100 observations	Eq.	a	b	λ^* [m]	h^* [m]
Van Dantzig without uncertainty	(5)	-	-	0.330	5.82
Van Dantzig with uncertainty	(8)	-	-	-	5.98
Bayes estimate symmetric linear loss	(1)	1	1	0.328	5.80
Bayes estimate symmetric squared-error loss	(2)	1	1	0.330	5.82
Bayes estimate asymmetric linear loss	(1)	$5.37\ 10^7$	$1.94\ 10^7$	0.349	6.05
Bayes estimate asymmetric squared-error loss	(2)	$1.07\ 10^8$	$3.88\ 10^7$	0.344	5.98
Bayes estimate linex loss	(3)	3.03	$1.32\ 10^7$	0.354	6.10
Bayes estimate modified linex loss	(9)	-	-	0.345	5.98
Bayes estimate modified linex loss	(10)	-	-	0.343	5.98

From Tables 1-2, we can conclude the following. The cost-optimal dyke height without taking the statistical uncertainties involved into account is, due to Eq. (5), equal to 5.82 m. When asymmetric loss functions are applied, the optimal dyke height is higher while taking the statistical uncertainty in λ into account. The larger the uncertainty in the scale parameter λ, i.e. the smaller the number of observations ν, the higher the cost-optimal dyke height. On the other hand, a symmetric squared-error loss function results in the same height without uncertainty (5.82 m) and a symmetric linear loss function can result in even lower heights (5.77 m and 5.80 m, respectively). As expected, the optimal dyke height under uncertainty according to Eq. (8) equals the dyke height that follow from both economic loss functions (9-10). Recall that the main difference between Eq. (9) and Eq. (10) is that the former is regarded as a function of the optimal q-quantile, whereas the latter contains the substitution for q in terms of Eq. (6). Since the linex loss function results in a dyke height much greater than the height in case of the economic loss functions, we recommend using the economic loss functions (modified linex loss functions) instead.

Conclusions

A Bayesian approach towards the estimation of flood quantiles has been suggested. Bayes estimators of the optimal dyke height under symmetric and asymmetric loss have been investigated when the annual maximum sea water levels are exponentially distributed with unknown mean. Three types of loss functions have been considered: (i) linear loss, (ii) squared-error loss, and (iii) linex loss. In order to properly account for the statistical uncertainty in the mean, a modified linex loss function can best be applied. This new modified linex loss function is derived from the economic dyke heigthening problem of Van Dantzig. The Bayes estimate of the dyke height under modified linex loss is equivalent to the optimal dyke height for which the economic loss is minimal. The modified linex loss function seems to be a promising candidate to solve quantile estimation problems in other civil engineering benefit-cost analyses. Moreover, unlike in most Bayesian literature, the parameters of the modified linex loss function have a clear economic significance. They represent the cost of investment (dyke heightening) on the one hand, and the cost of flooding on the other hand. The advantage of using a Bayesian loss function approach over a Van Dantzig approach is that the former approach is more closely related to the current design practice of hydraulic structures with fixed quantiles. The difference between under- and overdesign is more visible in the Bayesian loss function approach than in the Van Dantzig approach. The next step would be to repeat the type of work done in this paper on a larger scale in order to estimate the statistical parameters of other probability distributions.

References

Burcharth, H.F., and Liu, Z. (1994). On the extreme wave height analysis, In Port and Harbour Research Institute, Ministry of Transport, editor, *International Conference on Hydro-Technical Engineering for Port and Harbor Construction (Hydro-Port), Yokosuka, Japan, 1994*, p ages 123-142, Yokosuka: Coastal Development Institute of Technology.

Basu, A.P., and Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of Statistical Planning and Inference*, 29, 21-31.

Basu, A.P., and Thompson, R.D. (1992). Life testing and reliability under asymmetric loss. In J.P. Klein, editor, *Survival Analysis - State of the Art*, pages 1-7, Dordrecht: Kluwer Academic Publishers.

Fortin, V., Bernier, J., and Bobée, B. (1997). Simulation, Bayes, and bootstrap in statistical hydrology. *Water Resources Research*, 33(3), 439-448.

Fortin, V., Bobée, B., and Bernier, J. (1997). Rational approach to comparison of flood distributions by simulation. *Journal of Hydrologic Engineering*, 2(3), 95-103.

Pandey, M., Singh, V.P., and Srivastava, C.P.L. (1994). A Bayesian estimation of reliability model using the linex loss function. *Microelectronics Reliability*, 34(9), 1519-1523.

Thompson, R.D., and Basu, A.P. (1993). Bayesian reliability of stress-strength systems. In A.P. Basu, editor, *Advances in Reliability*, pages 411-421, Amsterdam: Elsevier.

Thompson, R.D., and Basu, A.P. (1996). Asymmetric loss functions for estimating system reliability. In D.A. Berry, K.M. Chaloner, and J.K. Geweke, editors, *Bayesian Analysis in Statistics and Econometrics*, pages 471-482, New York: John Wiley & Sons.

Van Dantzig, D. (1956). Economic decision problems for flood prevention. *Econometrica*, 24, 276-287.

Van Gelder, P.H.A.J.M. (1996). How to deal with wave statistical and model uncertainties in the design of vertical breakwaters. In H.G. Voortman, editor, *Probabilistic Design Tools for Vertical Breakwaters; Proceedings Task 4 Meeting, Hannover, Germany, 1996 (MAST III/PROVERBS: MAS3-CT95-0041)*, pages 1-13.

Varian, H.R. (1974). A Bayesian approach to real estate assessment. In S.E. Fienberg and A. Zellner, editors, *Studies in Bayesian Econometrics and Statistics: In Honor of Leonard J. Savage*, pages 195-208, New York: North-Holland.

Yamaguchi, M. (1997). Intercomparison of parameter estimation methods in extremal wave analysis. In B.L. Edge, editor, *25th International Conference on Coastal Engineering, Orlando, Florida, U.S.A., 1996*, p ages 900-913, New York: American Society of Civil Engineers (ASCE).

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, 81(394), 446-451.