A new statistical model for extreme water levels along the Dutch coast

Pieter H.A.J.M. van Gelder
Delft University of Technology, Faculty of Civil Engineering, Delft, Netherlands

ABSTRACT: The Netherlands is a low-lying country which has to protect itself against flooding from the sea and its rivers. Reliable flood defenses are essential for the safety of the country. The sea dikes are designed to withstand floods with a height of once every 10,000 years. This height is used to be calculated using statistics on sea levels measured along the Dutch coast since 1880. In this paper new calculations will be presented taking account of known sea floods in the period before 1880. A Bayesian framework has been developed for this.

1 INTRODUCTION

Approximately 40% of the Netherlands is below mean sea level; much of it has to be protected against the sea and rivers by dikes. In figure 1, a topographical map of the Netherlands with the North Sea, the river Rhine, its tributaries Waal and IJssel, and the Meuse is shown. The sea - and river dikes are clearly visible. The total length of the dikes is about 3000 km. Hardly any systematic statistical research has been done in connection with the height of the dikes before 1953. On February 1 1953, during a severe storm surge, in several parts of the Netherlands the sea dikes broke, approximately 150,000 ha polderland inundated through 90 breaches and 1800 people and thousands of cattle lost their lives. After this flood the Dutch government appointed a committee to recommend on an appropriate height for the dikes. The statistical work for this committee was under direction of professor D. Van Dantzig (Deltacommissie 1960). The recommendation of the committee was adopted by the government and most of the Dutch sea dikes have since been adapted to meet the new standard. Although the statistical analysis was the decisive argument for determining the new standard, a serious effort has been made to study the problem of determining an appropriate height for the sea dikes from an economic and also from a physical point of view. The economic analysis compared the cost of building dikes of a certain height with the expected cost of a possible inundation given that specific height. In the

Figure 1: Topographical map of the Netherlands

mathematical-physical analysis the effects of a wind field on a rectangular water bassin was studied. Now, more than 30 years later, both the number of observations of sea levels and the statistical methodology have grown considerably. This has led to a new investigation (Dillingh et. al. 1993, De Haan 1990). They investigated the problem solely from a statistical point of view of determining a level of the sea dikes such that the probability that there is a flood in a given year, equals p, where the number p is to be determined by the politics ranging between 10^{-3} and 10^{-4} (depending on the importance of the area under subject). For the data they used the complete set of high tide water levels recorded along the Dutch coast since the end of the 19th century. Although this set consists of lots of observations (over 100,000 per location), the problem of determing the dike level is a problem of far extrapolation: estimating a water level that can occur once every 10,000 years out of a data set of a small 100 years.

After the 1953 flood historical data about floods in previous centuries have been studied systematically (Gottschalk 1977, Jonkers 1989). Their reports give a good overview about the severeness of the floods that took place between 1500 and 1850. Up till now, this information hasn't been used in the calculation of the dike levels. In this paper Bayesian statistics will be shown to be useful to take account of historical data in the extrapolation problem. It will appear that in this Bayesian framework it is possible to include the global warming effect on earth and its result on the sea level rise as well.

The paper is organised as follows: the flood historical data and the assumptions on the sea level rise will be discussed in Sec.2. The measurements of sea level data since the end of 1800 will be discussed in Sec.3. We make a distinction there in POT and AM data (peaks over threshold and annual maxima). In Sec.4, we will look at a Bayesian model for the AM data, followed by a Bayesian model for the POT data in Sec.5. In Sec.7 a decision problem for the dikes. Finally in Sec.7, we will draw the conclusions.

2 FLOOD HISTORICAL DATA

Recently historical research has been undertaken to retrieve information about floods in the period 1500-1850 (Gottschalk 1977, Jonkers 1989). In diaries, signs of flood levels on old churches etc. a collection of old data could be gathered. The old data on flood levels however have to be taken as an indication, not as accurate data points. Different sources on the same sea flood give contradictive reports about the actual height which occured. That's why in this paper the floods will be classified into 4 classes: Class A for very severe floods; B for heavy floods; C for less heavy floods and D for light floods. Every class will be connected with a water level and uncertainty level, based on the historical research data and based on the sea level rise data. Class A floods are modelled to be realisations from a normally distributed stochast of flood levels with a mean of 390cm and a standard deviation of 10cm, denoted as N(390,10). B - floods are realisations from N(360,10); C - floods are in N(330,10) and D - floods in N(300,10). The following floods in the period 1500-1850 are mentioned including its classification:

1570 - A	1775 - D
1672 - D	1776 - C
1686 - D	1806 - B
1715 - D	1808 - B
1717 - D	1825 - B

It is argued that the sea level rise is, in contrary of popular belief, not a new phenomenon but a known fact from geological observations (Vrijling 1994). About 8,000 BC, the northern edge of the North Sea was dry and connected the British Isles with the continent. As the present depth of this area is equal on average to MSL-35m, a sea-level rise of approximately 35m must have occured in the last 10,000 years. This means a rise of 35cm per 100 year. Since 1888 the sea levels at Hook of Holland are recorded intensively. If we perform a linear regression analysis on the year maxima data of Hook of Holland, we obtain as the estimated regression line:

$$y=\alpha+\beta x=226.7+(\text{year}-1888)*0.2017 \text{ cm}.$$

This line is dotted in figure 2. We observe a sea level rise of 20 cm in 100 years. From the 90% confidence limits, given in the figure by the 2 solid lines, we notice that the hypothesis $\beta=0$ would also be accepted. However, in the Delta committee, the sea level rise for the Netherlands during the last centuries is determined at 20cm per 100 year, including the correction for the sinking of the soft soil of the polders (Deltacommissie 1960). We will apply a sea-level rise correction on both the flood historical data as well as the sea level data since 1888.

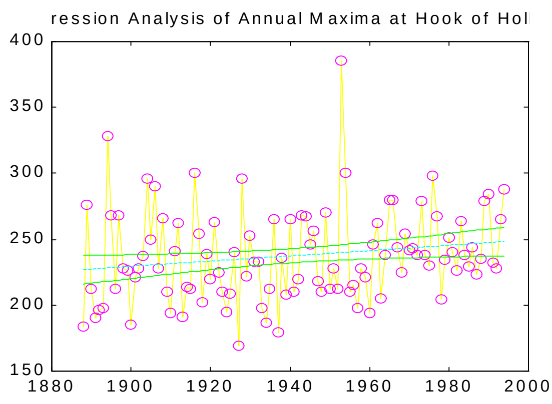

Figure 2: A regression analysis

3 SEA LEVEL DATA SINCE 1888

Since 1888 accurate recordings of the sea levels at Hook of Holland are availabe. For our purpose: determining the level of the sea dikes such that the probability that there is a flood in a given year, equals $p=1/10{,}000\ yr^{-1}$ (denoted as $x_{10,000}$ and called the 10^{-4} quantile), we filter the recordings to 2 data sets. The AM-data set, which consists of all the annual maxima of sea levels and the POT-data set, which consists of all the peaks of water levels over a certain threshold.

The AM-data set has the data of the period 1888 - 1995; in total n=108 points. The number of data points of the POT-data set depends on the level of the threshold (see Sec.5). Both the AM and POT data sets are corrected for the sea-level rise influence (of 20 cm per century).

4 MODEL AM DATA

4.1 *Extreme value model*

Assume that the maximum sea level in a period of 3 days is modelled by a stochastic variable X with (unknown) probability distribution F_X. Divide 1 year in 122 sections of 3 days. Let X_j be the maximum sea level in section j (j=1..122). For Hook of Holland it appears that the X_j's are independent; i.e. the period of 3 days is large enough to make storm 1 independent of storm 2. Let us now look at $X:=\max_{j=1..122} X_j$. From extreme value theory, it follows that X must tend towards an extreme value distribution (Castillo 1988). So X is a Frechet, Weibull or Gumbel distribution, dependent of the parent distribution of X_j. From a graphical analysis (plotting the AM-data on Gumbel paper and examining convexity, concavity or linearity), we suggest for a Gumbel distribution of the AM sea levels at Hook of Holland (see figure 3).

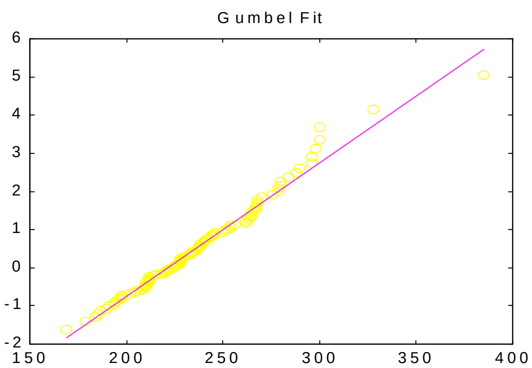

Figure 3: Graphical analysis of AM-data

4.2 *Frequentistic analysis*

With classical methods the calculation of the 10^{-4} quantile and its uncertainty will be presented. From Sec.4.1, we assume the AM data to be fitted by the Gumbel distribution with cumulative distribution function $F(x|\lambda,\delta)=\exp(-(x-\lambda)/\delta)$. The parameters λ and δ are estimated with a maximum likelihood method. This means that the estimators for λ and δ will be approximately (for large n), normally distributed. Their means are asymptotically equal to the true parameters (i.e. are asymptotically unbiased). Their variances are asymptotically equal to:

$$\mathrm{Var}(\lambda)=-1/nE(\partial^2 \log f(x|\lambda,\delta)/\partial\lambda^2)$$
$$\mathrm{Var}(\delta)=-1/nE(\partial^2 \log f(x|\lambda,\delta)/\partial\delta^2)$$

in which $f(x|\lambda,\delta)=\exp(-\exp(-(x-\lambda)/\delta))\exp(-(x-\lambda)/\delta)/\delta$ is the Gumbel probability density function. Under the assumption of a Gumbel likelihood model we have:

$$E(x_{10,000})=E(\lambda)-E(\delta)*\log(-\log(1-p))=E(\lambda)+qE(\delta)$$
$$\mathrm{Var}(x_{10,000})=\mathrm{var}(\lambda+q\delta)=\mathrm{var}(\lambda)+q^2\mathrm{var}(\delta)+$$
$$+2q\rho(\lambda,\delta)\sigma_\lambda\sigma_\delta$$

Here is $p=10^{-4}$ the probability that there is a flood in a given year, $q=\log(-\log(1-p))=9.2103$ and $\rho(\lambda,\delta)$ the correlation coefficient between λ and δ and related with the covariance which is asymptotically given by $\mathrm{Cov}(\lambda,\delta)=-1/nE(\partial^2 \log f(x|\lambda,\delta)/\partial\lambda\partial\delta)$.

With the AM-data we get $E(x_{10,000})=483.6$ cm and $\mathrm{var}(x_{10,000})=393.7$ or $\sigma(x_{10,000})=19.8$ cm.
The calculation of this 10^{-4} quantile and its uncertainty is based solely on the sea level data from the period since 1888. The data that we have from the period of 1500-1850 is impossible to use in this frequentistic analysis. That is why we like to perform a Bayesian analysis.

4.3 *Bayesian analysis*

The principle of a Bayesian analysis is as follows: First a prior distribution has to be determined in which we put our knowledge before any data is available. This prior distribution is therefore decided on subjective grounds. As soon as data becomes available, the prior distribution can be updated to the so called posterior distribution. Bayes Theorem plays the central role in this analysis:

$$p(\lambda|x) = \frac{f(x|\lambda)p(\lambda)}{\int_0^\infty f(x|\lambda)p(\lambda)d\lambda}$$

In this formula, the prior distribution is denoted by $p(\lambda)$, the likelihood model by $f(x|\lambda)$ and the posterior distribution by $p(\lambda|x)$. The integral assures that the posterior is a distribution function. For a good overview of Bayesian statistics, we refer to Bernardo and Smith 1994.

We continue the sea level analysis with a Bayesian approach in which the unknown parameters λ and δ are treated as random variabels. As argued in 4.1 we assume the likelihoodmodel to be given by Gumbel:

$$f(x_i|\lambda,\delta)=(1/\delta)^n\exp(-(\Sigma x_i-\lambda)/\delta)\exp(-\Sigma\exp(-(x_i-\lambda)/\delta))$$

in which x_j (j=1..108) are the AM data.

We start the Bayesian analysis with vague (or uninformed) prior distributions:

$$p(\lambda,\delta)=\lambda/(\lambda_{max}-\lambda_{min}) \text{ x } \delta/(\delta_{max}-\delta_{min})$$

i.e. 2 uniform distributed variabels between the wide bounderies $\lambda_{min}=100$, $\lambda_{max}=300$, $\delta_{min}=10$, $\delta_{max}=40$.
The posterior distribution is then given with Bayes by:

$$p(\lambda,\delta|x)=Cp(\lambda,\delta)\ (1/\delta)^n \exp(-(\Sigma x_i-\lambda)/\delta) \text{ x } \exp(-\Sigma\exp(-(x_i-\lambda)/\delta))$$

in which C is the normalisation constant such that $p(\lambda,\delta|x)$ integrates to 1.
The posterior predictive can be calculated from:

$$P(x<x_B)=\int\int F(x_B|\lambda,\delta)p(\lambda,\delta|x)d\lambda d\delta$$

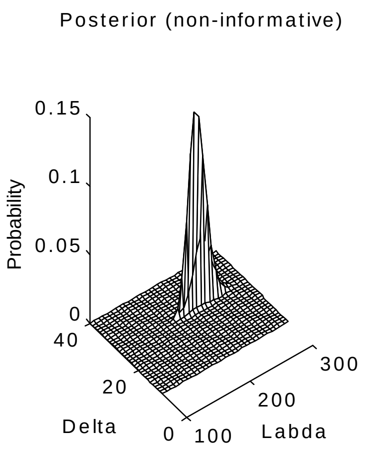

Figure 4: Posterior analysis

From the posterior predictive, $x_{10,000}$ follows from $P(x<x_{10,000})=1-10^{-4}$.

Figure 4 gives us the results of this non-informative analysis. We obtain almost the same estimators for the parameters λ and δ as in the frequentisitc method. $E(x_{10,000})$=490.1cm and $\sigma(x_{10,000})$=21.7cm.

We continue the Bayesian analysis with informed prior distributions. Take in mind that we have historical flood data to our access. We will use this information to get prior information on the δ and λ parameters.

Let $X_1, X_2, ..., X_{350}$ represent the annual maxima in year 1501, 1502, ... 1850. The r-th order statistic of this serie is given by the r-th member in the new sequence $X_{1:350}$, $X_{2:350}$, ..., $X_{350:350}$, which is a rearranging of $X_1, X_2, ..., X_{350}$ in increasing order.

If we assume that the annual maxima of the water level at Hook of Holland is modelled by a Gumbel distribution then the distribution of the r-th order statistic is given by:

$$f_{Xr:350}(x)=350!/(r-1)!/(350-r)!F^{r-1}(x)[1-F(x)]^{350-r}f(x)$$

in which f(x) and F(x) are the Gumbel p.d.f. and c.d.f. respectively. We are interested in the joint probability distribution of $X_{341:350}$, $X_{342:350}$, ..., $X_{350:350}$, given by:

$$f_{X341:350,\ X342:350,...,\ X350:350}(x_1,x_2,...,x_{10})= 350!\Pi_{i=1}^{10}f(x_i)\Pi_{j=1}^{11}[F(x_j)-F(x_{j-1})]^{r(j)-r(j-1)-1}/(r(j)-r(j-1)-1)!$$

in which r(0)=0, r(1)=341, r(2)=342, ... r(10)=350, r(11)=351, $x_0=-\infty$, $x_1 = x_{341:350}$,... and $x_{11}=\infty$.

On basis of the historical data we estimate with maximum likelihood the parameters λ and δ. We proceed until both parameters have reached its steady state. After 1000 simulations of $x_1, x_2,..., x_{10}$ this is the case. If we make a plot of λ against δ (figure 5), we observe a high correlation between the two parameters (correlation coefficient = -0.9939).

We have $E(\lambda)$=220.3, $\sigma(\lambda)$=28.4, $E(\delta)$=36.18, $\sigma(\delta)$=6.23 and $E(x_{10,000})$=553.6cm; based on 10 extreme floods in a period of 350 years. We model both parameters by the bivariate normal distribution:

$$f_{\Lambda,\Delta}(\lambda,\delta) = (2\pi\sigma_\Lambda\sigma_\Delta(1-\rho^2)^{-½})^{-1}\ \exp\{-2(1-\rho^2)^{-1} [((\lambda-\mu(\Lambda))/\sigma(\Lambda))^2 - 2\rho\,(\lambda-\mu(\Lambda))(\delta-\mu(\Delta))/\sigma(\Lambda)\sigma(\Delta) + ((\delta-\mu(\Delta))/\sigma(\Delta))^2]\}$$
$$\text{for } -\infty<\lambda<\infty,\ -\infty<\delta<\infty$$

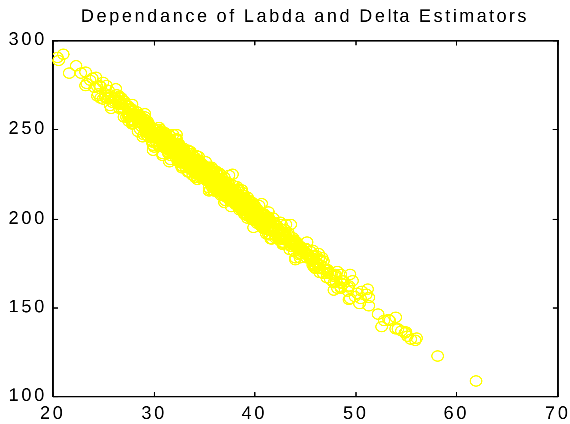

Figure 5: Prior distribution of (δ,λ)

With the informative prior as bivariate normal, we obtain the following results for the posterior: $E(\lambda)=225.5$, $var(\lambda)=8.7$, $E(\delta)=34.70$, $var(\delta)=0.76$, $\rho(\lambda,\delta)=-0.6385$ and $E(x_{10,000})=545.1$cm.

We notice that the role of the amount of sea level data since 1888 on the 10^{-4} quantile is of less effect because the quantile decreased from 553.6 to 545.1cm only.

5 MODEL POT DATA

5.1 *The threshold*

POT data is obtained by selecting peaks of sea levels above a certain level. There are no rules known in literature in order to determine a certain threshold. In this paper some sort of stability criterion is suggested to determine the threshold. First of all the POT data is modelled by an exponential distribution $e^{-\lambda x}$. Let x_j (j=1..n) denote the peak sea-levels above threshold T. Then we can calculate the 10^{-4} quantile as a function of T under the assumption of the exponential model.

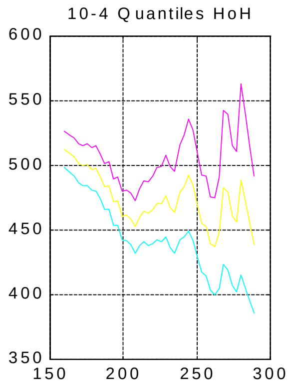

Figure 6: Threshold analysis

Figure 6 give the results (including its standard deviation interval). Where the 10^{-4} quantile remains more or less stable under variation of threshold, we choose the T. For the Hook of Holland data, our choice is T=200cm. The number of peaks over T for this choice is 195.

5.2 *Frequentistic analysis*

In contrary with the Gumbel model for the AM data, a lot of analytic solutions can be found for the exponential model for the POT data. The ML-estimator for λ is given by $\lambda=n/(\Sigma_{i=1..n}\, x_i)$.
The ML-estimator for the 10^{-4} quantile $x_{10,000,ML}$ follows from:

$$P(x>x_{10,000,ML})=\exp(-\lambda x_{10,000,ML})=1/10.000 \text{ or}$$
$$x_{10,000,ML}=\log(10.000)*\Sigma_{i=1..n}\, x_i\,/n$$

The distribution of $x_{10,000,ML}$ can also be calculated analytically: From $x_{10,000,ML}=\log(10,000)*\Sigma_{i=1..n}x_i/n$, and the fact that a summation of n exponential distributed stochasts, each with parameter λ is gamma distributed with parameters (n, λ), and that a constant k times a gamma distributed stochast with parameters (α,β) is also gamma distributed with parameters $(\alpha_1, \beta_1)=(\alpha,\beta/k)$, we have that $x_{10,000,ML}$ is gamma distributed with parameters $(n,n\lambda/\log 10000)$. So:

$$E(x_{10,000,ML})=(\log 10000)/\lambda$$
$$Var(x_{10,000,ML})=(\log^2 10000)/(\lambda^2 n).$$

5.3 *Bayesian analysis*

In contrary to the Gumbel model for the AM-data where analytical solutions are very difficult to derive, we can derive the Bayesian formulae of the exponential model analytically very easily. The likelihood model is assumed to be given by an exponential model:

$$f(x|\lambda)=\lambda^n\exp(-\lambda\Sigma x_i)$$

in which x_j (j=1..195) are the POT data. With a non-informative prior $(p(\lambda)=1/\lambda)$, we obtain as posterior:

$$p(\lambda|x)=\lambda^{n-1}\exp(-\lambda\Sigma x_i)/\int_{\lambda=0..\infty} \lambda^{n-1}\exp(-\lambda\Sigma x_i)d\lambda$$

This is nothing else than a gamma distribution with parameters $(n,\Sigma_{i=1..n}x_i)$.

The distribution of the 10^{-4} quantile from a Bayesian point of view $x_{10,000,B}$ can be derived from:
$P(x>x_{10,000,B})=\int_{\lambda=0..\infty}(1-F(x_{10,000,B}|\lambda))p(\lambda|x)d\lambda=$
$=(\Sigma_{i=1..n}x_i \,/\, (x_{10,000,B} + \Sigma_{i=1..n}x_i\,))^n=0.0001$ so that:

$$x_{10,000,B}=(10.000^{1/n}-1)*\Sigma_{i=1..n}x_i$$

Note that for $n\rightarrow\infty$, $x_{10,000,B}\rightarrow x_{10,000,ML}$; because $n(10.000^{1/n}-1)\rightarrow\log 10.000$ $(n\rightarrow\infty)$.
We have $\lambda\sim Ga(n,\Sigma_{i=1..n}x_i\,)$, so $1/\lambda\sim Ig(n,\Sigma_{i=1..n}x_i\,)$ in which Ig is the inverted gamma distribution.
If $X\sim Ig(\alpha,\beta)$ then $kX\sim Ig(\alpha,k\beta)$. With $x_{10,000,B}=n(10.000^{1/n}-1)(\Sigma_{i=1..n}x_i/n)=(1/\lambda)n(10.000^{1/n}-1)$ it follows that:

$$x_{10,000,B}\sim Ig(n,n(10.000^{1/n}-1)\Sigma_{i=1..n}x_i\,)$$

We conclude:

$$E(x_{10,000,B})=n(10.000^{1/n}-1)(\Sigma_{i=1..n}x_i\,)/(n-1)$$
$$var(x_{10,000,B})=(n/(n-1))^2\,(10.000^{1/n}-1)(\Sigma_{i=1..n}x_i\,)^2\,/(n-2)$$

Figure 7 gives us the results of the ML-analysis and the non-informative Bayesian analysis. From the 195 data points, it is estimated that $\lambda=0.0375$. The 10^{-4} quantile is calculated in the frequentistic way to be $E(x_{10,000,ML})=461.6$ cm and $\sigma(x_{10,000,ML})=18.7$ cm. The non-informative Bayesian analysis almost gave the same results $E(x_{10,000,B})=462.9$ and $\sigma(x_{10,000,B})=19.1$cm.

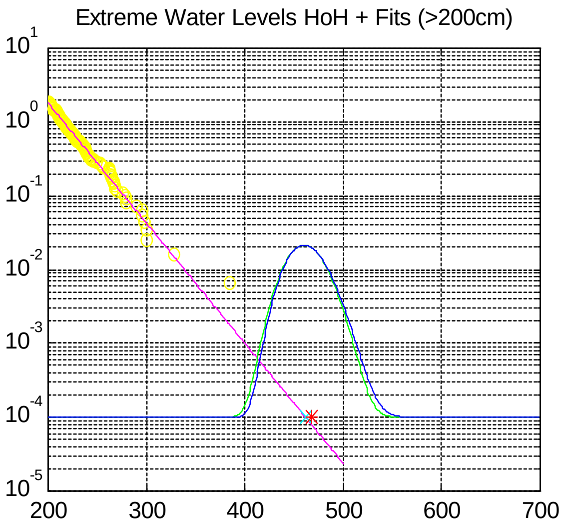

Figure 7: The exponential model

We continue the Bayesian analysis with informed prior distributions. In approximately 100 years, we have about 200 POT data (above 200cm). In the period 1500-1850, we therefore assume 700 POT data (in a sensitivity analysis performed later, it appeared that this amount of 700 was not very sensitive for the results). We have 10 POT data available during this period. In fact they are the 10 largest POT data, so we can use order statistics to estimate the most likely λ value if we assume an exponential parent distribution. We repeat this procedure 100 times and use the λ-estimates to fit a gamma distribution (which is conjugate with an exponential likelihood model). In figure 8 (left) the fit is shown. The prior distribution becomes $p(\lambda)=Ga(\lambda|\alpha,\beta)$, with $\alpha=183$ and $\beta=6392$. The conjugate posterior is also gamma distributed: $p(\lambda|x)=Ga(\lambda|\alpha+n,\beta+\Sigma_{i=1..n}x_i\,)$, where $n=195$ and $\Sigma_{i=1..n}x_i=5204$.

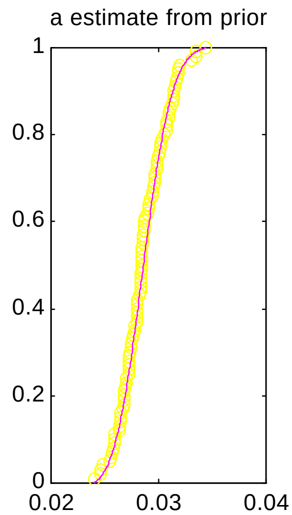

Figure 8: Posterior analysis

In figure 8 (right), the prior distribution (right function), the non-informative posterior distribution $Ga(\lambda|n,\Sigma_{i=1..n}x_i)$ (left function), and the informative posterior distribution (middle function), which is the normalized product of both other functions, are shown. The conjugate posterior predictive is given by:

$$P(x<x_B)=\int_{\lambda=0..\infty}F(x_B|\lambda)p(\lambda|x)d\lambda=$$
$$=Gg(x_B|\alpha+n,\beta+\Sigma_{i=1..n}x_i,1)$$

in which Gg stands for the Gamma-gamma distribution. We determine $x_{10,000}$ for which $P(x>x_{10,000})=1-10^{-4}$. We obtain 482.4cm. So notice that the role of the amount of sea level data since 1888 on the 10^{-4} quantile is of larger effect than in the Gumbel model for the AM data, because this time the quantile decreases from 521.4 to 482.4cm. See figure 9 for the progress of updating the prior beliefs with new data points.

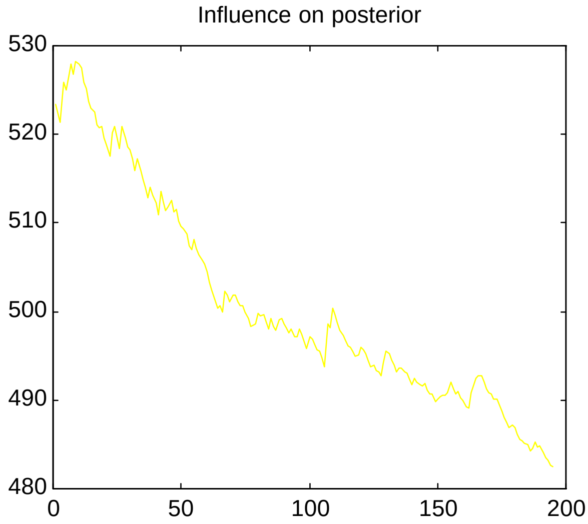

Figure 9: The updating process

6 A DECISION PROBLEM FOR DIKES

Van Dantzig (1956) posed the following economical decision problem for determining the optimal 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 sea levels, determine the optimal height of the dikes. If we take in correspondence with the result in section 5, the exponential model as the distribution for the sea levels: $p(h)=ce^{-\alpha h}$. We denote with H_0 the current dike height, X the optimal dike heightening, V the material loss after an inundation, I_0 the initial costs for dike heightening, k the subsequent costs of heightening per meter and δ the rate of interest, then the solution of the decision problem is given by:

$$X=1/\alpha \ln (100ce^{-\alpha H0}V\alpha/\delta k).$$

In a follow-up research (Van Gelder 1995), a Bayesian analysis on this decision problem, in which the parameters are treated as random variables and the uncertainty in the frequency distribution of sea levels is taken into account, is being performed.

7 CONCLUSIONS

In this paper, we have examined the problem of determining the sea level with a return period of 10,000 years. The known existing statistical models for this problem use the accurate sea level measurements of the period from 1888 up till now. However, flood historical data has been recently published about the period 1500 - 1850. In order to take account of this information into the current statistical model a Bayesian approach has been used in this paper. A classical frequentistic approach wouldn't have been possible in this case. An analysis has been made with a Gumbel model for the Annual Maxima data. The 10^{-4} quantile of the sea level increases significantly if the historical flood data is taken into account; from 490 to 545cm.
A second analysis has been made with an exponential model for the Peaks Over Threshold data. In this case the 10^{-4} quantile increases from 462 to 482cm.
The particular choice for an AM model or a POT model and the choice for the distribution type (Gumbel, exponential, etc.) remains unanswered in this paper. The choice for a Bayesian framework, however, shows to be very useful when taking account of historical data in the analysis of extreme water levels.

REFERENCES

Deltacommissie 1960. *Delta rapport.* Den Haag.
Dillingh, D., De Haan, L., Helmers, R., Können, G.P., Van Malde, J. 1993. *De basispeilen langs de Nederlandse kust.* Den Haag.
De Haan, L. 1990. *Fighting the arch-enemy with mathematics.* Reprint Series No. 602. Erasmus University Rotterdam.
Gottschalk, M.K.E. 1977. *Storm surges and river floods in the Netherlands.* Van Gorcum. Amsterdam.
Jonkers, A.R.T. 1989. *Over den schrikkelijken watervloed.* Den Haag.
Vrijling, J.K. 1994. *Sea-level rise: a potential threat?* Statistics for the environment 2: Water related issues. John Wiley & Sons.
Castillo, E. 1988. *Extreme value theory in engineering.* Academic Press, Inc.
Bernardo, J.M., Smith, A.F.M. 1994. *Bayesian theory.* Wiley Series in Probability.
Van Dantzig, D. 1956. *Economic decision problems for flood prevention*. Econometrica 24, p.276-287.
Van Gelder, P.H.A.J.M. 1995 *Determination of statistical distribution functions for reliability analysis of civil engineering structures.* Technical report. Delft University of Technology.