Prepared for:

Rijkswaterstaat Rijksinstituut voor Kust en Zee (RIKZ)

Extreme wave statistics

Methodology and applications to North Sea wave data

Report

February, 2007

WL | delft hydraulics

H4803.30

Prepared for:

Rijkswaterstaat Rijksinstituut voor Kust en Zee (RIKZ)

Extreme wave statistics

Methodology and applications to North Sea wave data

Sofia Caires

Report

February, 2007

wL | delft hydraulics

CLIENT:	Rijkswaterstaat, Rijksinstituut voor Kust en Zee
TITLE:	Extreme wave statistics: methodology and applications to North Sea wave data

ABSTRACT:

Currently, the offshore statistics on extreme values used in the definition of the hydraulic boundary conditions for the Dutch coast are estimated using certain standardized frequency analysis methods that are not completely in accordance with the principles of extreme value theory. Specifically, although the data are sampled in a way consistent with one of the approaches provided by extreme value theory, the probability distributions used to obtain extrapolations to high return periods are not the asymptotic distributions predicted by the theory.

In this work we adopt an approach for estimating offshore extremes that is consistent with extreme value theory, and assess and illustrate the differences between the results of such an approach and the one currently in use. This approach allows us to estimate possible physical bounds on wave heights and compare these with estimates obtained with the wave model SWAN.

REFERENCES: RKZ-1697 SAP bestelnummer “4500041303 pos 20”

VER	AUTHOR		DATE	REMARKS	REVIEW		APPROVED BY	
1.0	S. Caires		31/10/2006		H.F.P. van den Boogaard		W. Tilmans	
2.0	S. Caires		18/12/2006	after external review	H.F.P. van den Boogaard		M. van Gent	
3.0	S.Caires	SC	5/2/2007		H.F.P. van den Boogaard		W. Tilmans	

PROJECT NUMBER:			
KEYWORDS:	Extreme values, significant wave height, SBW, hydraulic boundary conditions		
NUMBER OF PAGES:	21 + figures		
CONFIDENTIAL:	☐ YES	☒ NO	
STATUS:	☐ PRELIMINARY	☐ DRAFT	☒ FINAL

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

In compliance with the Flood Defences Act of The Netherlands ("Wet op de Waterkering, 1996"), the primary coastal structures must be checked every five years (2001, 2006, 2011 etc.) for the required level of protection on the basis of the Hydraulic Boundary Conditions (HBC) and the Safety Assessment Regulation (VTV: *Voorschrift op Toetsen op Veiligheid*). These HBC must be derived anew every five years and established by the Minister of Transport, Public Works and Water Management.

At this moment, there is a degree of uncertainty concerning the quality of the current HBC, in particular those for the Waddenzee. This is because they were obtained from an inconsistent set of measurements and design values (WL, 2002), while for the rest of the Dutch coast (the closed Holland Coast and the Zeeland Delta) the SWAN wave transformation model has been applied (Rijkswaterstaat, 2001).

For 2011 and later the Dutch government plans to define the HBC for the Wadden Sea in the same way as for the rest of the Dutch Coast. In order to produce the best possible hydraulic boundary conditions for that region, and to assess the uncertainty that must be associated with such conditions, the Dutch Directorate for Public Works and Water Management (Rijkswaterstaat) is financing a large study led by WL | Delft Hydraulics. Specifically, Rijkswaterstaat requested WL | Delft Hydraulics to formulate, in the scope the subproject "Boundary Conditions", which is part of the main project "Strength and Loading of Coastal Structures (SBW: *Sterkte en Belasting Waterkeringen*)", a Plan of Action (WL, 2006). This plan establishes a strategy to answer the principle question of "How do we arrive at reliable Hydraulic Boundary Conditions for the Wadden Sea for 2011", and lists a sequence of associated activities; one of these is reported here.

1.2 Objectives of this study

One of the initial steps in defining HBC is the determination of offshore statistics on extreme values. These are used by the HYDRA-K program and by the wave model SWAN (Booij et al., 1999).

The offshore statistics of extremes consist of marginal (or univariate) statistics of water level, wind speed, significant wave height and different wave period parameters—defined as functions of the wind direction, and omni-directionally—as well as of multivariate statistics computed mainly on the basis of deterministic relations between the different variables.

Currently, the offshore wave statistics (WL, 2004 and 2005) are computed using certain standardized methods that are not fully in accordance with the principles of extreme value theory. Specifically, although the data are sampled in a way consistent with one of the approaches provided by extreme vale theory, the probability distributions used to obtain extrapolations are not the asymptotic distributions predicted by extreme value theory (Coles, 2001).

The aim of this work is to propose an approach for estimating offshore extremes that is more in line with extreme value theory, and to assess and illustrate the differences between the results of such an approach and the one currently in use. This approach also allows the estimation of possible physical bounds on wave heights which can be compared with estimates obtained from the wave model SWAN.

Because the goal is mainly to assess differences between methodologies rather than determining final estimates, only the omni-directional marginal distribution of significant wave height has been considered in this study. The extension of the proposed methods to direction-dependent and multivariate cases, however, may not be straightforward, and in fact our results must be seen as an initial step in tackling more complex problems.

Useful references for possible extensions include the works of Ewans and Jonathan (2006), who look at the problem of defining direction-dependent statistics of extremes, and Zachary et al. (1998), who look at the problem of estimating multivariate statistics of extremes. These and related works propose new methods based on extreme value theory which may turn out to be useful for our purposes and therefore should be kept in mind.

2 Extreme value theory

Extreme value theory provides analogues of the central limit theorem for the extreme values in a sample. According to the central limit theorem, the mean of a large number of random variables, irrespective of the distribution of each variable, is distributed approximately according to a Gaussian distribution. For example, the sea surface elevation is often modelled as a sum of several individual random waves and therefore its distribution can often be assumed to be Gaussian. According to extreme value theory, the extreme values in a large sample have an approximate distribution that is independent of the distribution of each variable.

In order to explain the basic ideas, let us define $M_n = \max\{X_1, \ldots, X_n\}$, where $X_1, X_2 \ldots$ is a sequence of independent random variables having a common distribution function *F*. In its simplest form, the *extremal types theorem* states the following: If there exist sequences of constants $\{\sigma_n > 0\}$ and $\{\mu_n\}$ such that $\mathrm{P}\{(M_n - \mu_n)/\sigma_n \le z\} \to G(z)$ as $n \to \infty$, where *G* is a non-degenerate cumulative distribution function, then *G* must be a generalized extreme value (GEV) distribution, which is given by

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \tag{2.1}$$

where *z* take values in three different sets according to the sign of ξ: $z > \mu - \sigma/\xi$ if $\xi > 0$ (the domain of *z* has a lower limit), $z < \mu - \sigma/\xi$ if $\xi < 0$ (the domain of *z* has an upper limit), and $-\infty < z < \infty$ if $\xi = 0$.

In other words, if the distribution function of (a normalization of) the maximum value in a random sample of size *n* converges to a distribution function as *n* tends to infinity, then that distribution function must be a GEV distribution. Moreover, this and other results of extreme value theory *hold true even under general dependence conditions* (Coles, 2001).

In Eq. (2.1), the parameters μ, σ and ξ are called the location, scale, and shape parameters and satisfy $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$. For $\xi = 0$ the GEV is the Gumbel distribution, for $\xi > 0$ it is the Fréchet distribution, and for $\xi < 0$ it is the Weibull distribution. For $\xi > 0$ the tail of the GEV is "heavier" (i.e., decreases more slowly) than the tail of the Gumbel distribution, and for $\xi < 0$ it is "lighter" (decreases more quickly and actually reaches 0) than that of the Gumbel distribution. The GEV is said to have a type II tail for $\xi > 0$ and a type III tail for $\xi < 0$. The tail of the Gumbel distribution is called a type I tail. See the book of Coles (2001) for more information.

The extremal types theorem gives rise to the *annual maxima* (AM) method of modelling extremes, in which the GEV distribution is fitted to a sample of block maxima (in this case annual maxima, but the same could be done for e.g. biannual, monthly or even daily maxima).

The sample sizes of annual maxima data are usually small, so that model estimates, especially return values, have large uncertainties. This has motivated the development of more sophisticated methods that enable the modelling of more data than just block maxima. These methods are based on two well-known characterizations of extreme value distributions: one based on exceedances of a threshold, and the other based on the behaviour of the *r* largest, for small values of *r*, observations within a block.

We will not describe the *r*-largest approach in detail because it is not often used in practice. Briefly, it consists of collecting the *r*-largest values per year (instead of merely the annual maxima) and fitting the *r*-largest distribution (see, for instance, p. 68 of the book of S. Coles mentioned above) to the data. An example of the application of this method to estimate return values of significant wave height is given by Guedes Soares and Scotto (2004).

The approach based on the exceedances of a high threshold, hereafter referred to as the POT (Peaks Over Threshold) method, consists of fitting the generalized Pareto distribution (GPD) to the peaks of clustered excesses over a threshold, the excesses being the observations in a cluster minus the threshold, and calculating return values by taking into account the rate of occurrence of clusters (see Pickands, 1971 and 1975, and Davidson and Smith, 1990). Under very general conditions this procedure ensures that the data can have only three possible, albeit asymptotic, distributions (the three forms of the GPD) and, moreover, that observations belonging to different peak clusters are (approximately) independent. In the POT method, the peak excesses over a high threshold u of a time series are assumed to occur in time according to a Poisson process with rate λ_u and to be independently distributed with a GPD, whose distribution function is given by

$$F_u(y) = 1 - \left(1 + \xi \frac{y}{\tilde{\sigma}}\right)^{-1/\xi},$$

where $0 < y < \infty$, $\tilde{\sigma} > 0$ and $-\infty < \xi < \infty$. The two parameters of the GPD are called scale ($\tilde{\sigma}$) and shape (ξ) parameters. For $\xi = 0$ the GPD is the exponential distribution with mean $\tilde{\sigma}$, for $\xi > 0$ it is the Pareto distribution, and for $\xi < 0$ it is a special case of the beta distribution. As for the GEV, the GPD is said to have a type II tail for $\xi > 0$ and a type III tail for $\xi < 0$. The tail of the exponential distribution is a type I tail.

Just as block maxima have the GEV as their approximate distribution, the threshold excesses have a corresponding approximate distribution within the GPD. Moreover, the parameters of the GPD of threshold excesses are uniquely determined by those of the associated GEV distribution of block maxima. In particular, the shape parameter is the same, and the scale parameters of the two distributions are related by $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

The choice of threshold (analogous to the choice of block size in the block maxima approach) represents a trade off between bias and variance: too low a threshold is likely to violate the asymptotic basis of the model, leading to bias; too high a threshold will generate fewer excesses with which to estimate the model, leading to high variance.

An important property of the POT/GPD approach is the threshold stability property: if a GPD is a reasonable model for excesses of a threshold u_0, then for a higher threshold u a GPD should also apply; the two GPD's have identical shape parameter and their scale parameters are related by $\tilde{\sigma}_u = \tilde{\sigma}_{u_0} + \xi\left(u - u_0\right)$, which can be reparameterized as

$$\sigma^* = \tilde{\sigma}_u + \xi\mu \tag{2.2}$$

Consequently, if μ_0 is a valid threshold for excesses to follow the GPD then estimates of both σ^* and ξ should remain nearly constant above μ_0. This property of the GPD can be used to find the minimum threshold at which a GPD model applies to the data.

There are several methods available for the estimation of the parameters of extreme value distributions. Most of them, for instance the methods of moments and of probability weighted moments, give explicit expressions for the parameter estimates. The maximum likelihood (ML) method tends to be the preferred estimation method since it is quite general and more flexible than other methods, especially when the number of parameters is increased as for instance when extending the extreme value approach to account for non-stationarity. When obtaining ML estimates, the variances of the estimates can be obtained from the expected information matrix or from the observed information matrix. An alternative, and usually more accurate, method is the profile likelihood method (Coles, 2001, p. 57), which is based on the deviance function and yields asymmetric confidence bands intervals. However, in ordinary extreme value analyses like the ones we are concerned with in this report the flexibility provided by the ML method is not necessary, and for the range of tails typically found with wave data (not too heavy-tailed distributions) and for small to moderate sample sizes the method of Probability-Weighted Moments (PWM) performs better (for details, see Hosking and Wallis, 1987, and Hosking et al., 1985). For this reason, it is this method that will be used in this work. Since, the method of PWM does not provide asymmetric confidence bands, we will combine it with adjusted bootstrap estimates (see Coles and Simiu, 2003) in order to compute confidence intervals. A study on the coverage rate of confidence intervals of extreme value estimates based on different methods is to be reported soon in the same framework of this study.

Model checking in extreme value analyses is usually done by means of probability plots, quantile plots and return level plots. In this report we have chosen to illustrate the fits with return level plots.

3 Remarks on the current approach

The way in which the offshore statistics of wave extremes used to define the hydraulic boundary conditions for 2006 have been calculated is documented in WL (2004, 2005); for simplicity, we shall refer to it here as the *current approach*. In the following subsections we briefly describe the current approach and give some remarks on the choice of the distribution.

3.1 Choice of distribution

The sampling of the buoy data used to make predictions about extremes followed, in essence, the POT method (details will be given later in Section 3.2). However, instead of fitting the GPD distribution to the data, the fitted model was the so-called *conditional Weibull distribution*, whose distribution function is given by

$$F(x) = 1 - \exp[(\omega/\sigma)^{\alpha} - (x/\sigma)^{\alpha}], \text{ for } x \geq \omega, \tag{3.1}$$

where α is the shape parameter, σ is the scale parameter, and ω is the threshold.

The choice of this distribution is based on historical usage supported by recommendations of RIKZ (1995b) and has been partly motivated by the work of Battjes (1970); see WL (2004, p. 4-1). However, some comments must be formulated with regard to the motivation of this choice. We will first try to explain the genesis of the conditional Weibull distribution, then analyse the motivation interpreted as coming from Battjes (1970), and finally examine the study and recommendations of RIKZ (1995b).

The *Weibull distribution* has distribution function

$$F_W(x) = 1 - \exp\left[-\left(\frac{x-\nu}{\sigma}\right)^{\alpha}\right], \quad x > \nu, \tag{3.2}$$

where ν is the location parameter, σ is the scale parameter, and α is the shape parameter. This Weibull distribution is not the Weibull distribution *of maxima* referred to in Section 2, but the Weibull distribution of minima (a form of the GEV distribution for minima). In particular, while the latter has a Type III upper tail, the former has a Type I (exponential) upper tail. It was introduced by Weibull in connection with failure data because it is the approximate distribution of the *minimum* of many variables, which could be seen as the weakest link among many links that can be broken in a structure. Popularized by reliability engineers, its use has spread to other areas, in particular to ocean engineering.

Now suppose that X is a random variable with a Weibull distribution with ν=0, and as usual write $P(E)$ for the probability of the event E. Then the probability that $X \leq x$ conditionally on the event that $X > \omega$ is, for $x > \omega$, given by

$$
\begin{aligned}
P(X \le x \mid X > \omega) &= \frac{P(X \le x \cap X > \omega)}{P(X > \omega)} = \frac{P(X \le x) - P(X \le \omega)}{P(X > \omega)} \\
&= \frac{P(X > \omega) - P(X > x)}{P(X > \omega)} = 1 - \frac{\exp(-(x/\sigma)^\alpha)}{\exp(-(\omega/\sigma)^\alpha)} \\
&= 1 - \exp[(\omega/\sigma)^\alpha - (x/\sigma)^\alpha].
\end{aligned}
$$

Since this last expression coincides with Eq. (3.1), we see that the conditional Weibull distribution is the distribution of the exceedances from a Weibull distribution above the threshold ω.

Thus it seems that the adoption of the conditional Weibull distribution to model the exceedances of significant wave height over a high threshold is based on the idea that the whole (non truncated) data follow a single distribution, and moreover a Weibull distribution. The assumption that the data follow a single distribution, however, is well-known to be unrealistic; see Ferreira and Guedes Soares (1998, 2000) and Anderson et al. (2001). On the other hand, the Weibull distribution is not really a universally appropriate model to describe whole sets of wave data—e.g. the lognormal distribution is another commonly used model.

It is in this context that we must examine the reference by WL (2004) to the work of Battjes (1970). This author compared the fits of the Weibull and Lognormal distributions to one-year long data sets from seven different locations and concluded that the Weibull distribution of Eq. (3.2) fitted the data best. However, despite this conclusion, he clearly points out (Battjes, 1970, p. 18) that in order to predict extreme values of significant wave height one needs to consider annual maxima and to use a Gumbel (i.e., Fisher-Tippett double exponential) distribution. This means that Battjes (1970) does not draw conclusions about the prediction of extremes in terms of what seems to be a good model for the whole data, except perhaps for the fact that the Weibull distribution indicates that a Type I tail should be used for that purpose. For this reason, we think that the work of Battjes (1970) must not be understood as a justification for the adoption of the conditional Weibull distribution.

We now examine the work of RIKZ (1995b). These authors simulated a very long time series of significant wave height using the so-called Bruinsma method, and compared the use of the conditional Weibull distribution and the GPD distribution in estimating the 1/10,000 year return value. These distributions were fitted to 25 sets of exceedances of 10, 25 and 100 yearly samples, the exceedances being obtained over three pre-defined thresholds: 4, 5 and 6 metres. RIKZ (1995b) concluded that the conditional Weibull distribution was overall a better choice.

Their conclusion was based on a certain ‘robustness’ criteria. They found that the variance of the Weibull estimates was generally lower; that at lower thresholds and with shorter time series the GPD distribution underestimated the true return value; and that at higher thresholds and with short time series the variance of the GPD estimates was large. They also found that the Weibull distribution was ‘more resistant’ to outliers than the GPD. Specifically, the estimates of the GPD distribution were more seriously affected by an increase in the maximum value in the sample than the Weibull estimates.

The conclusions of RIKZ (1995b) and their scope for generalization raise several issues. First of all, the data RIKZ (1995b) simulated seem to have a Type I (or exponential) tail, which besides limiting the usefulness of the study does not take into account that fitting the exponential distribution (rather than the GPD, of which it is a special case) might lead to better results. Note that the conditional Weibull distribution has a Type I tail, and therefore does not allow for all three possible types of tail behaviour (unlike the GPD). Therefore, if the data being simulated had a type III tail (which would be realistic to assume in shallow waters or even in a finite depth situation) it is not certain that a Type I tail like that of the conditional Weibull distribution would lead to the best results.

Secondly, in the application of the POT method the choice of the threshold is usually 'data driven', i.e., it is to some extent dictated by the sample, rather than fixed in advance (see Section 2), and this may explain a putative under-performance of the GPD distribution.

As to the argument based on 'resistance to outliers', it does not seem to be a practically meaningful one. Indeed, the requirement that the parameter estimates remain more or less unaffected by an increase in the maximum value of the sample, although 'agreeable' from the point of view of policy-making, seems to be extraneous to the problem of making reliable predictions. For if the data indicate more uncertainty or variability (in the underlying data generating process) than what one had assumed on the basis of a previous, smaller sample, it seems more sensible to account for the new information in the data—especially if the new information consists of bigger extremes—than to ignore it, which implies appropriately updating the values of the parameters as well as their confidence intervals.

This last point also brings us to what we think is a common misconception: that a model which gives less variable estimates or narrower confidence intervals is a better model. This will be the case only if the estimates themselves are correct. Thus, if we simulate samples from a GPD with Type II tail and compare GPD and exponential fits, we may find that the return value estimates of the exponential are less variable than the GPD ones, but that is not to say that the estimates of the exponential are to be trusted nor that the variability suggested by the GPD is spurious.

The actual fitting of the models in WL (2004) is carried out in several stages. In the first stage, the parameters of the conditional Weibull distribution are estimated by choosing the threshold in the region where the estimates of the shape parameter are stable (i.e., do not change much with changes in the threshold), as it is usually done in the POT/GPD approach. This is done for data at each location, so that after the first stage one has a collection of estimates across a region. Because it is reasonable to assume that the estimates obtained should depend on fetch and water depth, WL (2004) suggest that the shape parameter estimates should vary smoothly in space as a function of fetch and water depth. Using this principle, WL (2004) carry out a *modified regional frequency analysis* in which the shape parameter estimates of both significant wave height and wave period data are used to model a parametric relation between shape parameters, fetch and water depth (see WL, 2004, pp. 4-7,8, section 4.3). In this way, new, and final, shape parameters estimates are obtained for each location.

The final estimates of the shape parameter are then used to estimate the other parameters of the conditional Weibull model by fitting the ‘reduced’ model (i.e., with fixed shape parameter) to the 200 highest peaks over threshold (see WL, 2004, pp. 4-11, 2005, pp. 4-13).

Finally, because the estimated conditional Weibull distributions have unlimited range to the right, and because at the locations considered the significant wave height is supposed to be fetch and/or depth limited, WL (2005) used the wave model SWAN to estimate the physical limit of wave height at each location, and then truncate the corresponding return value functions.

Our main comment on these procedures is about the exclusion, a priori, of the GPD model. As will be seen later on, the data show that the underlying tail could be of Type III—and this at all the locations considered—which does not support the use of a Type I tail model, like the conditional Weibull distribution, alone. The truncation of the return value function that WL (2005) opt for in the final stage of their estimation process reflects just this. If one adopts a GPD model then truncation is not necessary because the return value function of the type III GPD model is already bounded above. Adopting the GPD model and estimating the upper limit of its support also allows for the comparison with the estimates of such limit obtained via SWAN, and hence for further checks. One of the points of the alternative analysis we propose will consist of comparing the two types of estimates.

We finish this section by pointing out that there are certain situations in which it makes sense to fit a Weibull distribution (*not* the conditional Weibull distribution) to the peak excesses in place of the GPD model. Indeed, suppose that the data really follow a Type I tail, or at least that this has been convincingly demonstrated on the basis of some statistical analyses (which is often the case in deep waters; see Caires and Sterl, 2005). Then the asymptotic distribution of the excesses is exponential. Since the exponential is a special case of the Weibull distribution, one might think that there would be no harm in fitting a Weibull rather than an exponential to the data. Now, if the data are truly exponential, this would actually entail more uncertainty in parameter estimates, which would be undesirable (intuitively, to know that the data are exactly exponential amounts to more information than knowing that they are Weibull). However, it may happen that, because the exponential is only valid asymptotically, the Weibull distribution will provide a better approximation to the data (since it has one more parameter and hence more flexibility), and in that case fitting the latter would provide better results than fitting the former. In any case, if one is to step outside the GPD domain one should do so on the basis of some justification.

Still in this connection, we should add that, because wave data most often exhibits a type I tail or a slightly lighter type III tail, studies comparing estimates from the GPD with estimates from the Weibull are often inconclusive as to what is the most appropriate tail/distribution (e.g. Van Vledder et al., 1993) since there are no statistically significant differences between the two models (as was also the case in the RIKZ (1995b) study). However, a good reason for always considering the GPD is that this model has a substantial and solid theoretical basis stemming from asymptotic considerations.

3.2 Data collection in the POT method

Typically, return value estimates of significant wave height are used to calculate the corresponding *sea state maximum wave heights* on the basis of a short-term wave height distribution characterized by a given value of significant wave height which is assumed constant during a 3-hour interval. However, buoy measurements of significant wave height are usually computed from 20-minute long records. Due to sampling variability, significant wave height storm peaks calculated from such (shorter) records are biased upwards, i.e., they are overestimations of storm peaks based on 3-hour long records. According to Forristall el al. (1996), the overestimation depends on the shape of the spectrum, the length of the wave measurements, the number of measurements, and on how quickly the storm passes its peak. For typical wave spectra and sampling strategies the bias is 3-10%. The shape of the spectrum and the length of the wave measurements can be used to estimate the standard deviation of the measurements; the higher the standard deviation, the higher the overestimation; the longer the storm takes to pass its peak, the higher the overestimation. Also according to Forristall et al. (1996), the upward bias can be eliminated if smooth estimates of significant wave height are calculated from continuous wave records in the severe storms of interest. More precisely, the bias can be eliminated if a light filter (less smoothing) is used for short storms and a heavy filter (more smoothing) used for very long storms.

Another problem that may affect the accuracy of storm peaks is undersampling: in short lasting storms, when wave samples are not taken continuously, it is possible that the true peak of the storm is missed and hence that the significant wave height peak is underestimated. However, for typical storm durations and common sampling strategies this problem is not serious, and the effects of sampling variability are dominant.

The measurements that were used in WL (2004, 2005) and are to be used in this study are based on 24 years of 3-hourly wave data. The data were continuously collected in records of 10 minutes length. Every hour, two 10-minute records were combined to estimate the wave parameters. In the current approach, storm peaks above a threshold (the POT sample) are collected from the 3-hourly time series in such a way that the sample extracted from the original time series can be modelled as independent observations. With wave and similar data this is usually done by a process of *declustering* in which only the peak (highest) observations in clusters of successive exceedances of a specified threshold are retained and, of these, only those which in some sense are sufficiently far apart (so that they belong to more or less ‘independent storms’) are considered. Specifically, in the case of WL (2004, 2005) cluster maxima at a distance of less than 48 hours apart were treated as belonging to the same cluster (storm). In order to remove the effects of sampling variability mentioned above, the resulting sequence of storm maxima are then smoothed/filtered. See WL (2004) for details and examples of filters that are applied depending on the type of the storm.

The current approach to data collection is (apart from the filtering) that of the usual POT approach, and seems to be quite sound. Therefore, in our study we will take exactly the same approach and use the MATLAB routines developed for the WL (2004, 2005) studies to extract the samples of peaks. Thus the POT data to be used in this study is the same that was used in the previous WL studies referred to.

4 Extreme values analysis

4.1 Introduction

In this section we analyse the North Sea buoy measurements used in the WL (2004, 2005) studies with the POT/GPD and AM/GEV approaches, in accordance with the principles of extreme value theory summarized in Section 2. As mentioned earlier, we shall use the same data sampling techniques as that of the WL (2004, 2005) studies for collecting the POT and AM data. In particular, the storm maxima used do not correspond to 'raw' measurements but to smoothed estimates thereof.

The data used by WL (2004, 2005) consisted of measurements covering the period of 1979-2002 from 9 buoys located offshore along the Dutch coast; see Figure 4.1.

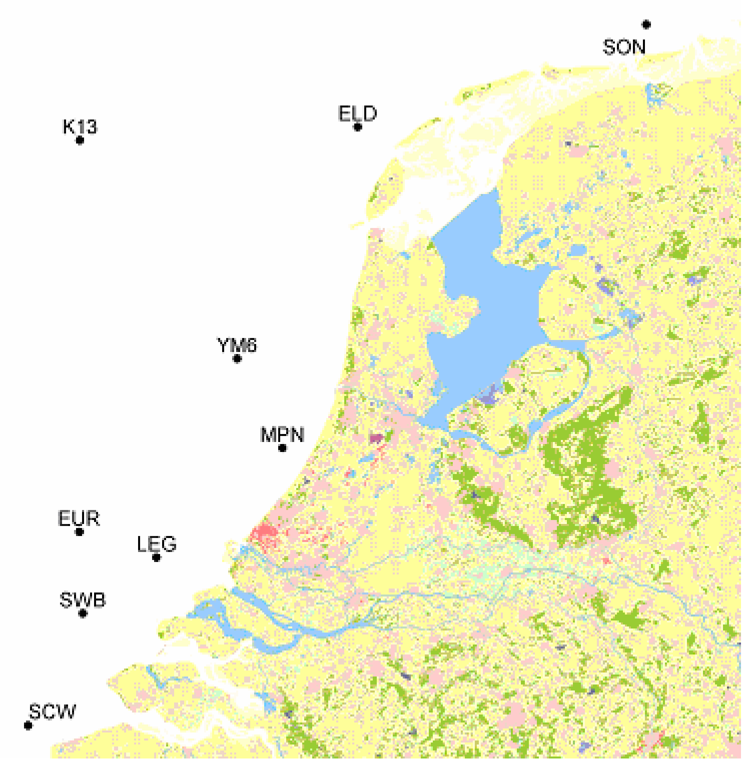


Figure 4.1 Measuring stations along the Dutch coast

The comparison between the current approach and the approach we advocate will be made in terms of the corresponding 1/10000 yr return value estimates.

In a study to be reported soon and in the same framework of this study, SWAN is used in 2D mode to determine physically-based upper bounds for the significant wave height at the SON, ELD, YM6, EUR and SCW measurement locations. We shall also compare these *physical* upper bound estimates with those obtained here from the POT/GPD method.

If the POT/GPD and AM/GEV fits are appropriate—i.e., if the fitted distributions are indeed close to the asymptotic distributions of the data—then the shape parameter and return value estimates of the two models should be compatible (allowing for their sampling errors). If this is not the case, then at least one of the fits is unreliable. The POT/GPD and AM/GEV estimates are compared at the end of this section.

Recall that we use the method of PWM (Probability-Weighted Moments) to estimate the parameters of both the GPD and GEV distributions. Confidence intervals are obtained using percentile bootstrap estimates adjusted according to the recommendations of Coles and Simiu (2003).

All computations were carried out in MATLAB and the GPD and GEV parameter estimates were obtained using WAFO statistical toolbox (see http://www.maths.lth.se/matstat/wafo/ for more information about WAFO).

4.2 POT/GPD estimates

We have used the threshold stability property mentioned in Section 2 to choose the most appropriate threshold for selecting a sample of peak excesses and fitting the GPD to it. More precisely, we have looked for threshold values around which the estimate of the shape parameter and σ^* (see Eq. 2.2) show the least variation. For the purposes of this report we have tried to automatize the choice of the threshold. In WL (2004, 2005) it is assumed that the best threshold lies in a region where the estimates of the conditional Weibull distribution do not vary much. On this basis, they fix the threshold using an automatized procedure. We employ here a similar procedure to choose the most appropriate threshold at which to fit the GPD. The procedure we have used is as follows:

1. POT samples with at least 10 and at most 252 peaks are collected by systematically decreasing the threshold, and for each of these samples GPD fits are obtained. (Note that the fact that there is a POT sample with, say, 20 peaks, does not mean that there is also a POT sample with 19 peaks, since different peaks may have the same value and even a small increase of the threshold can eliminate more than one of the peaks collected at a lower threshold.)
2. For each sample size n, a set of parameter estimates based on sample sizes ranging from n-k to n+k peaks, where k is some fixed value (see below), are obtained, and the standard deviation of such a set of estimates is computed. In the case of the shape parameter, for example, this procedure yields one standard deviation for each value of n, and each standard deviation quantifies the variability of the parameter estimates around a 'window' of 2k+1 sample sizes (2k+1=(n+k)-(n-k)+1).
3. The threshold, or sample size n, at which the GPD is finally fitted corresponds to the one yielding the smallest standard deviation among the standard deviations of the sets of *shape parameter* estimates computed in bullet 2.

Several tests were carried out to determine the best choice of the window size on which the standard deviations of bullet 2. are computed. In our case, k=12 turns out to be the best: using about 25 (2k+1 for k=12) estimates, the threshold automatically chosen coincides with the one that we would have chosen by visual inspection of plots like those in Figure 4.2a, and results in a generally good GPD fit. In most cases the results are resistant against changes in k for k between 10 and 15. With larger values of k the threshold chosen is often too low.

Only in the case of one of the time series analysed, the YM6 data, did the automatically chosen threshold produce a clearly poor GPD fit. In the analysis of that series we have therefore raised the threshold above the one suggested by the automatic procedure, to a region where the variation of shape parameter estimates is also low (see details below).

Figures 4.2 a-i show the estimates of the shape parameter, of σ* and of the 1/10000 return value as functions of the threshold obtained with the data from each of the 9 locations. The threshold at which the corresponding standard deviations defined in the automatized procedure just described are minimal are marked by vertical lines. The return value plot of the final GPD fits (obtained according to point 3. above) are presented in Figures 4.3 a-i. Table 4.1 presents the relevant information on the final fits, namely the sample size (n), the threshold used (u), the estimates of the GPD shape ($\tilde{\sigma}$) and scale (ξ) parameters, the 1/10000 yr return value estimate of significant wave height, all accompanied by their 95% confidence intervals, and the estimated upper bounds for significant wave height (these are estimates of the upper bound ($u - \sigma / \xi$, if $\xi < 0$) of the support of the GPD fitted to the peak excesses of significant wave height).

Buoy	*n*	u (m)	ξ	$\tilde{\sigma}$	1/10000 yr RV (m)	upper bound (m)
SON	157	3.69	-0.13 (-0.36 , 0.05)	1.08 (0.83 , 1.38)	9.99 (7.33 , 16.82)	11.91
ELD	171	4.09	-0.19 (-0.37 , -0.01)	1.06 (0.82 , 1.31)	9.02 (7.35 , 13.28)	9.71
K13	219	4.09	-0.10 (-0.27 , 0.04)	0.88 (0.71 , 1.06)	9.94 (7.67 , 15.33)	12.50
YM6	136	4.18	-0.07 (-0.27 , 0.12)	0.73 (0.55 , 0.92)	9.71 (7.12 , 17.45)	14.37
MPN	186	3.20	-0.16 (-0.32 , -0.01)	0.88 (0.69 , 1.09)	7.85 (6.32 , 10.98)	8.80
EUR	215	3.55	-0.14 (-0.33 , 0.03)	0.79 (0.63 , 0.97)	7.99 (6.33 , 12.28)	9.10
LEG	199	3.37	-0.22 (-0.40 , -0.06)	0.94 (0.74 , 1.16)	7.23 (6.02 , 9.76)	7.57
SWB	233	3.05	-0.20 (-0.36 , -0.05)	0.78 (0.64 , 0.95)	6.57 (5.51 , 8.84)	6.96
SCW	215	2.46	-0.32 (-0.52 , -0.14)	0.77 (0.62 , 0.95)	4.83 (4.25 , 6.08)	4.90

Table 4.1 Estimates of the parameters of the POT/GPD model and of the 1/10000 yr return value of significant wave height with the corresponding 95% confidence intervals

The results presented in Table 4.1 suggest the following remarks:

- At each location, the POT/GPD approach yields a type III tail (suggesting that the data have an upper bound), the point estimates of ξ ranging from -0.32 to -0.07. In some cases the exponential distribution, corresponding to ξ=0, cannot be excluded as a good model for the data, but in most cases the confidence intervals for ξ do not include ξ=0.
- The threshold and sample sizes adopted vary with location.
- The largest return value estimates occur in the most northern North Sea locations (those that are particularly relevant for the Wadden Sea HBC, see Figure 4.1).
- The relative uncertainty in the ξ estimates is rather large.
- As is usually the case, the point estimate of ξ determines the uncertainty of the return value estimates (i.e. the width of their confidence intervals) in a crucial way: the lower the estimate of ξ, the narrower the confidence intervals. For this reason, the effect of sample size in reducing the width of the confidence intervals can be quite different across different values of ξ (compare the MPN and EUR estimates in Table 4.1).

We shall now analyse the choice of the threshold and the fitted GPD at each location:

- The estimates based on the SON measurements indicate that at this location there is an upper bound of about 12 m, at least if we ignore the considerable uncertainty in the estimates of ξ. The threshold at which the minimum variation in estimates is achieved is the same for ξ, σ^* and for the 1/10000 yr return value; see Figure 4.2a. The GPD fit looks reasonable, with some scatter of the peaks around the return value line; see Figure 4.3a.
- For the ELD data, the thresholds at which there is less variation in the ξ, σ^* and 1/10000 yr return value estimates differ, see Figure 4.2b. However, the return value estimates based on each threshold do not significantly differ and the GPD fit based on the stability of the shape parameter looks reasonable.
- The estimates based on the K13 measurements show that the thresholds at which the estimates of ξ, σ^* and of the 1/10000 yr return value have minimal variation coincide; see Figure 4.2c. The return value plot of the final GPD fit, Figure 4.3c, shows that the estimated distribution fits the data well.
- Using our automatized procedure the thresholds for the YM6 data at which there is minimal variation in the ξ, σ^* and 1/10000 yr return value estimates coincide (u=3.68 m, n=212). However, the GPD fit of the POT data sampled above that threshold is rather poor and significantly underestimates some of the highest observed storms, see Figure 4.3d2. We have therefore decided to increase the threshold choosing the new value on the basis of visual inspection of the shape parameter estimate as function of the threshold (the top panel of Figure 4.2d). The new chosen threshold, u=4.18 m and n=136, is in a region where the estimates of the shape parameter are also rather stable, but the value of the shape parameter is higher. In the top panel of Figure 4.2d the threshold at which the final POT sample was collected is marked with a vertical line. The vertical lines in the remaining panels mark the threshold chosen by the automatized procedure. As can be verified in Figure 4.3d, the GPD fit to the data is rather good. Relative to the estimates based on the data from other locations, the shape parameter and upper bound estimates obtained with these data are larger. Because there are no reasons to assume that the behaviour of the extremes at this location should differ much from those at the other locations, we think that there may have been some problems with the measurements at this location.
- The estimates based on the MPN measurements indicate that the thresholds at which there is minimal variation in the ξ, σ^* estimates are the same but differ from the threshold at which there minimal variation in the 1/10000 yr return value estimates. As can be seen in Figure 4.2e, at the latter threshold, the estimates of the shape parameter are rather low. These low shape parameter estimates are responsible for the small variation in the return value estimates. Incidentally, this illustrates the fact that the return value estimates cannot be used as a proxy for the stability property of the POT/GPD approach. In spite of the scatter of the observed peaks around the estimated return value line, the GPD fit looks reasonable; see Figure 4.3e.
- The estimates based on the EUR measurements show that the thresholds at which there is minimal variation in the ξ, σ^* and 1/10000 yr return value estimates are the same; see Figure 4.2f. The return value plot in Figure 4.3f show a good fit, although there is a storm peak rather close to the upper bound of the 95% confidence intervals.
- The estimates based on the LEG measurements indicate that the thresholds at which the estimates of ξ and σ^* have minimal variation coincide (but differ from the threshold at which there is less variation in the 1/10000 yr return value estimates). The return value

plot presented in Figure 4.3g shows a rather good agreement between the observed peaks and the return value estimates.

- The quality of the POT/GPD fits for the SWB and the SCW data is rather similar and in both cases quite good; see Figures 4.3h and 4.3i. Also for both data sets, the thresholds at which less variation in the ξ, σ^* estimates are the same; see Figures 4.2h and 4.2i.

4.2.1 Comparison between the POT/GPD and the physical upper bound estimates

As mentioned above, in a parallel study, SWAN was used in 2D mode to determine physically based upper bounds for the significant wave height at the SON, ELD, YM6, EUR and SCW locations. This was done by computing the significant wave heights generated by extreme wind speeds from different directions applied uniformly over the whole North Sea, and for different spatially uniform water levels. The highest water level considered was 6 m and the highest wind speed 40 m/s. These are also the water level and wind speed at which the significant wave height maxima were attained. We shall refer to these maxima as *the physical upper bounds.*

In Table 4.2 the following quantities are provided:

- bed level;
- wind directions for which the maxima of significant wave height were attained in the SWAN computations;
- the attained maxima significant wave height; and
- comparisons between these and the corresponding POT/GPD upper bound estimates in terms of absolute and relative biases.

Buoy	bed level (m NAP)	SWAN		POT/GPD	bias (m)	relative bias
		wind direction (ºN)	upper bound (m)	upper bound (m)		
SON	20.8	330	10.25	11.91	-1.66	-0.14
ELD	28.0	330	11.81	9.71	2.10	0.22
YM6	23.4	300	10.83	14.37	-3.54	-0.25
EUR	32.4	360	13.08	9.10	3.98	0.44
SCW	12.9	330	7.62	4.90	2.72	0.56

Table 4.2 Estimates of the physical maxima of significant wave height based on SWAN-2D computations and their absolute and relative biases with respect to the estimates obtained using the POT/GPD

The SWAN physical upper bound estimates are closely related to the water depth, varying between 34% and 40% of the water depth (the bed levels plus a 6 m water level), which indicates that the maxima are more likely to be depth- than fetch- limited.

There is no clear relation between the SWAN-based estimated physical upper bounds and the estimated POT/GPD upper bounds. However, the differences between the two are small relative to the large uncertainty in the POT/GPD estimates. Table 4.1 shows that the amplitude of the confidence intervals for the 1/10000 return value varies between 40 and 100% relative to the point estimate, the uncertainty in the POT/GPD upper bound being at least as high. No uncertainty estimates are provided by the SWAN study, but such uncertainty clearly exists and depends not only on the errors in SWAN's results but also on the assumptions used in the definition of the upper bounds, namely the extreme and spatially uniform wind fields and water levels.

At the SON and YM6 locations the POT/GPD upper bound estimates are larger than the physical ones. If we were to trust the POT/GPD estimate, the discrepancies at the SON location could be a consequence of North Atlantic swell entering the North Sea for storms from the Northeast/North, which are not accounted for in the SWAN computations and may affect the local significant wave height. Given the doubts about the quality of the YM6 observations, the bigger POT/GPD upper bound estimate based on the YM6 data is probably an overestimate. The estimated physical upper bounds at the remaining locations are larger than the POT/GPD upper bound estimates. This may indicate that the conditions (extreme and spatially uniform water level, wind speed and direction over the whole North Sea) considered in the SWAN computations cannot be extrapolated on the basis of the measurements analysed using the POT/GPD approach.

4.2.2 Comparison between the POT/GPD and the current approach

Table 4.3 presents the estimates of the 1/10000 yr return values of significant wave height reported in WL (2004) and which were used in the definition of the HBC for 2006. No estimates of the confidence intervals are given. In WL (2005), alternative estimates of return values are given with the associated confidence intervals. However, the method used to estimate these confidence intervals also affects the shape parameter estimates of the conditional Weibull distribution obtained from the modified regional frequency analysis carried out, yielding estimates of the return values up to 2 m above the estimates given in WL (2004). The estimates presented in WL (2005) were, however, not further used.

Buoy	WL (2004)	POT/GPD	bias (m)	relative bias
SON	11.00	9.99	1.01	0.10
ELD	11.10	9.02	2.08	0.23
K13	10.50	9.94	0.56	0.06
YM6	9.30	9.71	-0.41	-0.04
MPN	8.00	7.85	0.15	0.02
EUR	8.20	7.99	0.21	0.03
LEG	8.10	7.23	0.87	0.12
SWB	7.10	6.57	0.53	0.08
SCW	5.70	4.83	0.87	0.18

Table 4.3 Estimates of the 1/1000 yr return value of significant wave height given by WL (2004) and their absolute and relative biases with respect to the estimates obtained using the POT /GPD approach

By comparing the 1/1000 yr return value estimates of significant wave height obtained using the current approach with those obtained using the POT/GPD approach (Tables 4.2 and 4.3) we arrive at the following conclusions:

- The spatial variation of the estimates is quite similar.
- The point estimates of the 1/1000 yr return value of significant wave height based on the conditional Weibull distribution are, except for the YM6 data, larger than the POT/GPD estimates (up to about 23% or 2.1 m). This can be a consequence of the fact that the tail

of the conditional Weibull distribution is heavier than the tails we have estimated with the POT/GPD method.

- The POT/GPD estimates are consistent with those obtained using the current approach in the sense that the latter are well within the 95% confidence intervals of the former.

On the basis of this comparison, we can conclude that in general the current approach yields estimates that are conservative relative to the estimates obtained here, but the differences do not appear to be substantial because the present confidence intervals contain the point estimates obtained in WL (2005).

4.3 AM/GEV estimates

In this section the North Sea buoy measurements used in Section 4.2.1 are reanalysed using the annual significant wave height storm maxima. As described in Section 3.1, the storm maxima used do not correspond to ‘raw’ measurements but represent smoothed estimates thereof. Table 4.4 presents the estimates of the GEV shape (ξ), scale (σ) and location (μ) parameters and of the 1/10000 yr return value of significant wave height together with their 95% confidence intervals, and the estimated upper bounds. Figures 4.4 a-i show the GEV return value plots for each location.

Buoy	ξ	σ	μ (m)	1/10000 yr RV (m)	upper bound (m)
SON	-0.16 (-0.50 , 0.10)	0.89 (0.61 , 1.17)	5.53 (5.13 , 6.00)	9.84 (7.94 , 15.73)	11.15
ELD	-0.18 (-0.57 , 0.16)	0.61 (0.41 , 0.82)	5.80 (5.53 , 6.08)	8.49 (7.02 , 14.74)	9.10
K13	-0.38 (-0.81 , 0.01)	0.73 (0.47 , 0.97)	6.03 (5.72 , 6.33)	7.87 (7.02 , 11.12)	7.93
YM6	0.04 (-0.22 , 0.24)	0.59 (0.39 , 0.86)	5.28 (5.03 , 5.62)	11.91 (8.20 , 19.67)	none
MPN	-0.21 (-0.51 , 0.09)	0.52 (0.37 , 0.67)	4.76 (4.52 , 5.00)	6.90 (5.96 , 10.51)	7.28
EUR	-0.01 (-0.28 , 0.22)	0.51 (0.34 , 0.70)	4.95 (4.73 , 5.21)	9.42 (6.50 , 16.79)	45.87
LEG	0.18 (-0.12 , 0.45)	0.28 (0.15 , 0.42)	4.90 (4.77 , 5.07)	11.31 (6.85 , 26.57)	none
SWB	-0.10 (-0.39 , 0.17)	0.38 (0.24 , 0.53)	4.51 (4.34 , 4.69)	6.80 (5.59 , 10.74)	8.36
SCW	-0.09 (-0.40 , 0.17)	0.30 (0.20 , 0.40)	3.69 (3.56 , 3.85)	5.52 (4.58 , 8.38)	6.83

Table 4.4 Estimates of the parameters of the AM/GEV model and of the 1/10000 yr return value of significant wave height with the corresponding 95% confidence intervals

From Table 4.4 it is readily verified that the confidence intervals of the GEV shape parameter estimates are rather wide and that in all the cases the Gumbel distribution (ξ=0) cannot be excluded as a good model for the data. Because of the large uncertainty in the GEV shape parameter estimates we shall not analyse the corresponding upper bound estimates in detail. The upper bound estimates (or lack thereof, when the estimates of ξ are $\geq$0) based on the YM6, EUR and LEG data are physically impossible. The return value estimates at these locations look rather high and do not follow the general decrease of the 1/10000 yr return values from North to South present in the POT/GPD estimates.

4.3.1 Comparison between the AM/GEV and the POT/GPD estimates

All the GEV 1/10000 yr return value estimates are consistent with the POT/GPD estimates in the sense that the 95% confidence intervals of the POT/GPD estimates contain the point estimates of the AM/GEV estimates. This gives us some trust in the appropriateness of both approaches.

The following remarks can be formulated about the differences in the fits/estimates of both approaches per location:

- The differences in both model estimates based of the SON data are rather small (15 cm in terms of 10000 yr return value) and statistically not significant. Both fits look quite reasonable and the compatibility between both estimates of the asymptotic tail suggest that the extreme values were rather well estimated.
- The same can be said about the fits based of the ELD data, which are also compatible.
- For K13 the GEV shape parameter point estimate is outside the 95% confidence intervals obtained for the POT/GPD estimate. The GEV return value plot shows a worse fit than that of the GPD. This leads us to conclude that the rather small GEV shape parameter and upper bound estimates are unreliable.
- The AM/GEV fit for YM6 is reasonable and the estimates are higher but compatible with the POT/GPD estimates. However, as we had also noted when analysing the POT/GPD estimates, the population of bigger storms at this location does not seem to be compatible with those at the other locations.
- The differences between the POT/GPD and AM/GEV estimates based on the MPN data are not statistically significant on the basis of the 95% confidence intervals. The GEV fit seems to match the data better than the POT/GPD fit.
- The GEV shape parameter estimate based on the EUR data is much larger than the corresponding POT/GPD estimates; nevertheless, both distributions seem to fit the data well apart from almost underestimating one of the high storm peaks; see Figures 4.3f and 4.4f.
- The GEV estimate of the shape parameter for LEG is the largest of all and is not compatible with that obtained with the POT/GPD method. The GEV return value plot in Figure 4.4g shows a reasonable fit given the very high uncertainty. It does not seem very plausible, however, that the asymptotic distribution of the storms at these location have indeed such a heavy tail.
- With the SWB data there is a rather good correspondence between the POT/GPD and AM/GEV return values estimates, and both fits look rather good.
- SCW is the location with the lowest return value estimate. This is so with both the POT/GPD and the AM/GEV methods. The AM/GEV and the POT/GPD estimates are compatible, but the AM/GEV estimates are higher and the fit seems to be better, as can be observed when comparing Figures 4.3i and 4.4i.

5 Final remarks

5.1 Summary

Currently, the offshore extreme wave conditions used in the definition of the Dutch coast HBC are estimated on the basis of the conditional Weibull distribution (WL, 2004, 2005). In the present study the use of this distribution in place of the distributions provided by extreme value theory has been discussed. More generally, it has been argued that the prediction of extreme values should be based on methods furnished by extreme value theory.

These methods involve two steps: the selection of samples of extremes, and the selection of distributions to be fitted to those samples. In the case of the Peaks over Threshold (POT) method, peak observations above a given threshold are collected in a sample and fitted with the generalized Pareto distribution (GPD). In the case of the annual maxima (AM) method, yearly maxima are collected in a sample and fitted with the generalized extreme value (GEV) distribution.

When using the POT and AM methods one must ensure that the samples collected satisfy certain properties of homogeneity and independence.

If other distributions than the GEV and the GPD are fitted to the AM and POT samples, respectively, then it should be explicitly acknowledged that extreme value theory does not justify the use of such distributions, and accordingly some sound justification must be given for using them.

An approach based on the POT/GPD and AM/GEV methods has been proposed and illustrated by means of applications. This approach has been applied to the North Sea buoy measurements used earlier in the WL (2004, 2005) studies and covering the period of 1979-2002.

5.2 Conclusions

The main conclusions of this study are as follows:

1. The estimates from the POT/GPD method are seen to be reliable, and they seem to suggest that in all the considered locations the asymptotic distribution of the data has a type III tail and hence that the wave height has an upper bound.
2. However, this last statement needs to be further investigated due to the large uncertainty in the estimation of the tail.
3. Also, on the basis of the comparison between the POT/GPD estimates and the physical estimates of the upper bounds on the data, it can be concluded that the former estimates do not provide a clear spatial and/or physically meaningful pattern.
4. The POT/GPD estimates obtained in this study are compatible with those given in WL (2004) in the sense that the present confidence intervals contain the point estimates

obtained in WL (2004). However, almost all of our estimates are less conservative (they are lower by at most 2.1 m).

5. The uncertainty in the present estimates of parameters and return values is particularly large in the case of the GEV distribution. Although the AM/GEV method is somewhat more objective than the POT/GPD method, because no threshold is involved, it is rather wasteful of data, and seems to require longer time series than those presently available in order to provide more definite estimates of the shape parameter.

5.3 Suggestions for further work

A way of assessing and further checking the quality of the shape parameter estimates (and consequently upper bounds on the significant wave heights) may be to perform a regional frequency analysis, as the one proposed by Van Gelder et al. (2000), in which spatial information is used to improve the domain of attraction estimates. Such an analysis has been applied by WL (2004) as a means of ‘correcting’ preliminary parameter estimates.

Another approach to determining more accurately the tail underlying significant wave height data is to use the distribution-free method of De Haan and Rootzén (1993), which is applied to water level data and briefly explained in RIKZ (1993, 1995a).

In this report only the omni-directional marginal distribution of significant wave height has been considered. The extension of the proposed methods to direction-dependent and multivariate cases is by no means straightforward, and in fact the approach suggested here must be seen as an initial step in tackling a much more complex problem. In our opinion, further progress requires the establishment of a theoretically sound, and generally accepted framework for multivariate extreme value analyses.

5.4 Recommendations

On the basis of this study it is recommended that estimates of return values of wave data be obtained using methods of extreme value theory, specifically the POT/GPD approach and extensions thereof for dealing with direction-dependent statistics of extremes (Ewans and Jonathan, 2006) and for estimating multivariate statistics (Zachary et al., 1998). Further, such line of approach could be complemented by non-parametric methods of establishing the domains of attraction of the data (e.g. de Haan and Rootzén, 1993) and/or by using spatial information by means of a regional frequency analysis (e.g. Van Gelder et al., 2000).

References

Anderson, C. W., D. J. T. Carter, and P. D. Cotton, 2001: Wave climate variability and impact on offshore design extremes. Shell International and the Organization of Oil and Gas Producers Rep., 99 p.

Battjes, J. A., 1970: Long-term wave height distribution at seven stations around the British Isles. N.I.O. Internal Report No. A.44, National Institute of Oceanography, UK.

Booij, N., R.C. Ris. and L.H. Holthuijsen, 1999: A third-generation wave model for coastal regions, Part 1: Model description and validation, J. Geophys. Res., 104 (C4), 7649– 7666.

Caires, S. and A. Sterl, 2005: 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data. J. Climate, 18(7), 1032-1048.

Coles, S., 2001: An Introduction to the Statistical Modelling of Extreme Values. Springer Texts in Statistics, Springer-Verlag: London.

Coles, S., and E. Simiu, 2003: Estimating uncertainty in the extreme value analysis of data generated by a hurricane simulation model. J. Engrg. Mech., 129 (11), 1288-1294.

Davison, A. C., and R. L. Smith, 1990: Models for exceedances over high thresholds (with discussion). J. Roy. Stat. Soc., 52B, 393–442.

Ewans, K. and P. Jonathan, 2006: Estimating Extreme Wave Design Criteria Incorporating Directionality. Proc. of the 9th Int. Workshop on Wave Hindcasting and Forecasting, September 24-29, Victoria, B.C., Canada. Available at http://www.oceanweather.com/waveworkshop/9thWaves/.

de Haan, L. and H. Rootzén, 1993: On the estimation of high quantiles. J. Stat. Plan. Inference, 35(1), 1-13.

Ferreira, J. A., and C. Guedes Soares, 1998: An application of the peaks over threshold method to predict extremes of significant wave height. J. Offshore Mech. Arct. Eng., 120, 165–176.

Ferreira, J. A., and C. Guedes Soares, 2000: Modelling distributions of significant wave height. Coast. Eng., 40, 361–374.

Forristall, G.Z., J.C. Heideman, I.M. Legget, B. Roskam, and L. Vanderschuren, 1996: Effect of sampling variability on hindcast and measured wave height. J. Wtrwy., Port, Coast., and Oc. Engrg., 122 (5), 216-225.

Guedes Soares, C. and M. G. Scotto, 2004: Application of the r largest-order statistics for long-term predictions of significant wave height, Coastal Eng., 21 (5-6), pp. 387-394.

Hosking, J. R. M., and J. R. Wallis, 1987: Parameter and quantile estimation for the Generalized Pareto Distribution. Technometrics, 29, 339–349.

Hosking, J. R. M.,, J. R. Wallis, and E.F. Wood, 1985: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics, 27, 251-261.

Pickands, J., 1971: The two-dimensional Poisson process and extremal processes. Journal of Applied Probability, 8, 745-756.

Pickands, J., 1975: Statistical inference using extreme order statistics. Annals of Statistics, 3, 119-131.

Rijkswaterstaat, 2001: Hydraulische Randvoorwaarden 2001 voor het toetsen van primaire waterkeringen. (Hydraulic Boundary Conditions Book). Den Haag.

RIKZ, 1993: De basispeilen langs de Nederlandse kust; statistisch onderzoek, Rijkswaterstaat, Dienst Getijdenwateren /RIKZ, Rapport DGW-93.023.

RIKZ, 1995a: De basispeilen langs de Nederlandse kust, De ruimtelijke verdeling, Rijkswaterstaat, Rijksinstituut voor Kust en Zee/*RIKZ,* Rapport RIKZ-95.008

RIKZ, 1995b. Golfrandvoorwaarden langs de Nederlandse kust op relatief diep water, Repport RIKZ -95.024.

Van Gelder, P.H.A.J.M., J.G. de Ronde, N.M. Neykov, Pl. Neytchev, 2000: Regional frequency analysis of extreme wave heights: trading space for time. Proc. 27th ICCE, 2, Coastal Engineering 2000, Sydney, Australia, 1099-1112.

Van Vledder, G., Y. Goda, P. Hawkes, E. Mansard, M.J. Martin, M. Mathiesen, E. Peltier, and E. Thompson, 1993: Case studies of extreme wave analysis: a comparitive analysis. Proc. WAVES'93 Conf., New Orleans, USA, 978-992.

WL, 2002: Kwaliteit randvoorwaardenboek en kwaliteit SWAN. WL|Delft Hydraulics Report H4061, Januari 2002.

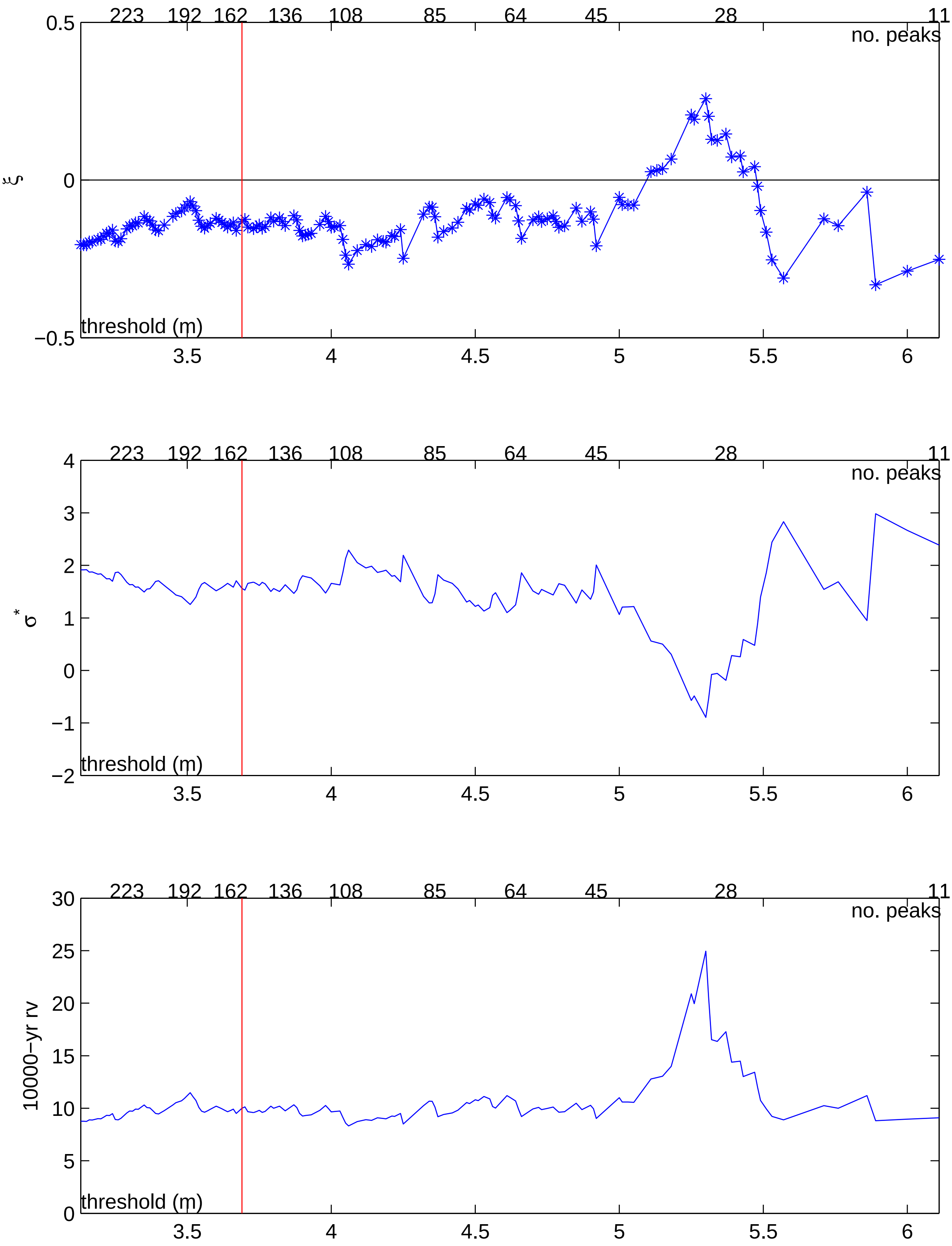
WL, 2004: Golfstatistiek op relatief diep water 1979-2002. WL | Delft Hydraulics Report Q3770, December 2004.

WL, 2005: Analyse golfstatistiek op relatief diep water. WL | Delft Hydraulics Report Q3966.10, September 2005.

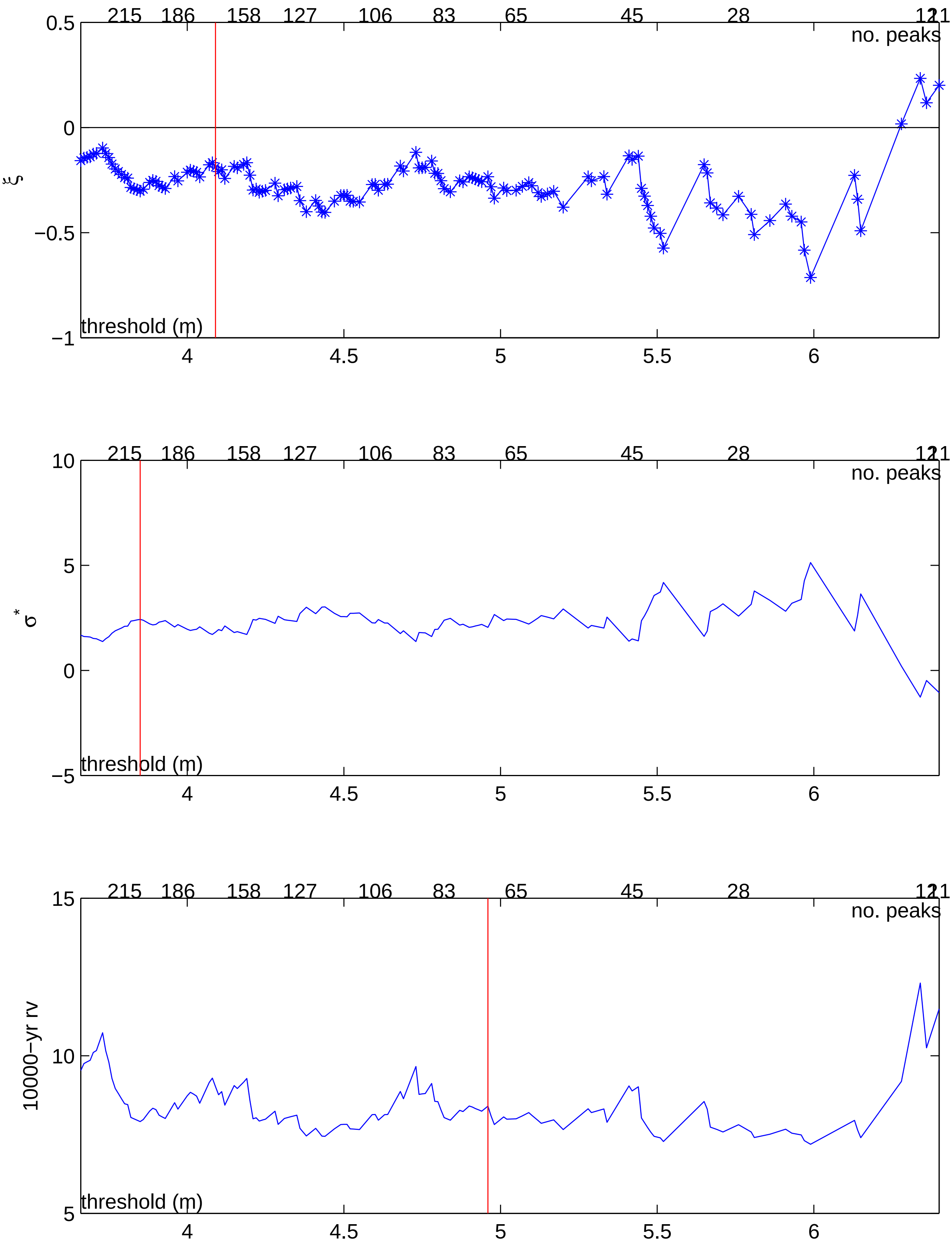
WL, 2006: SBW Plan of Action on the boundary conditions for the Waddenzee. WL | Delft Hydraulics Report H4750, April 2006.

Zachary, S., G. Feld, G. Ward and J. Wolfram, 1998: Multivariate extrapolation in the offshore environment, Applied Ocean Research, 20, 273–295.

Figures

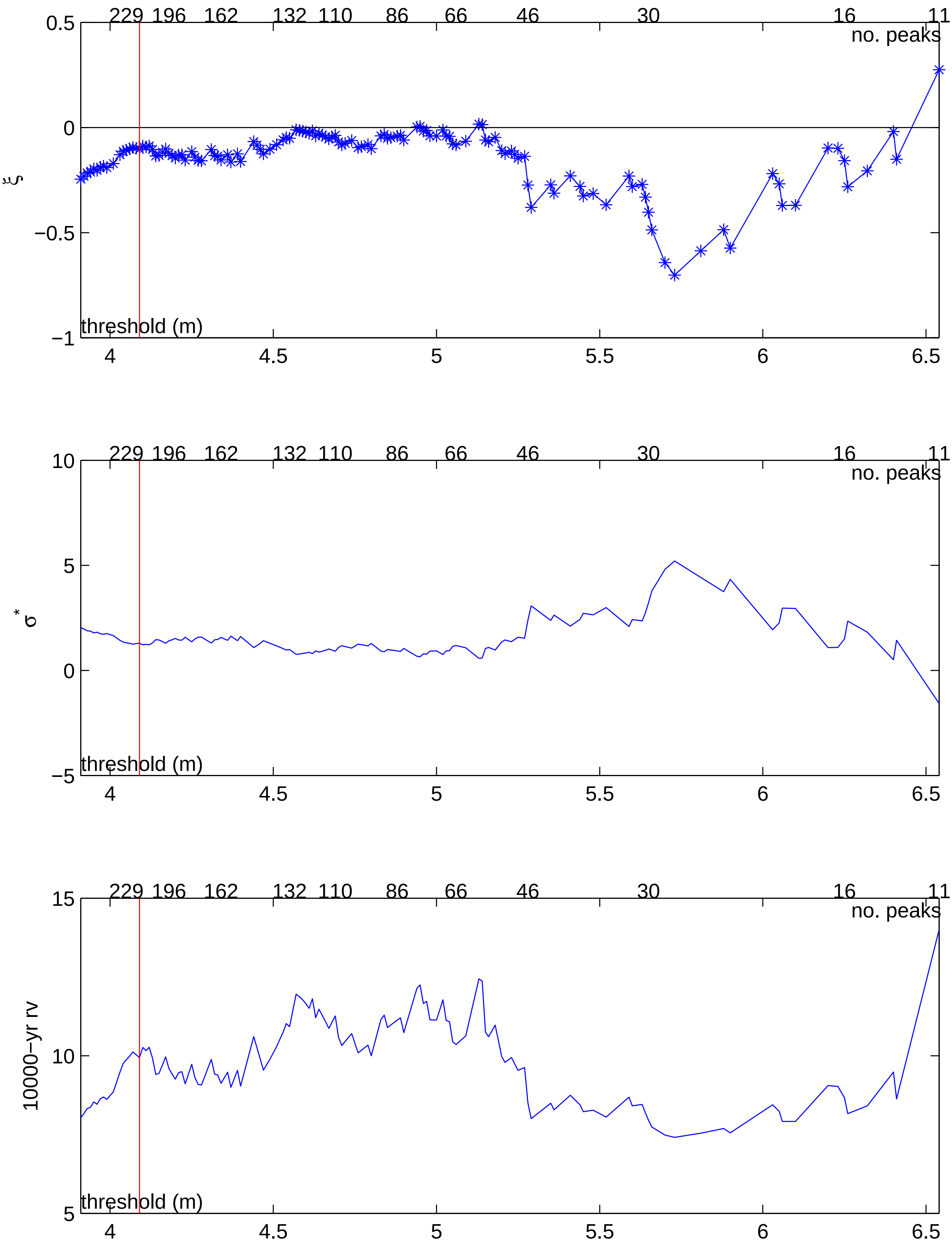


POT/GPD threshold choice buoy: SON	H4803.30	gw1sonwa
WL | DELFT HYDRAULICS		Fig. 4.2a

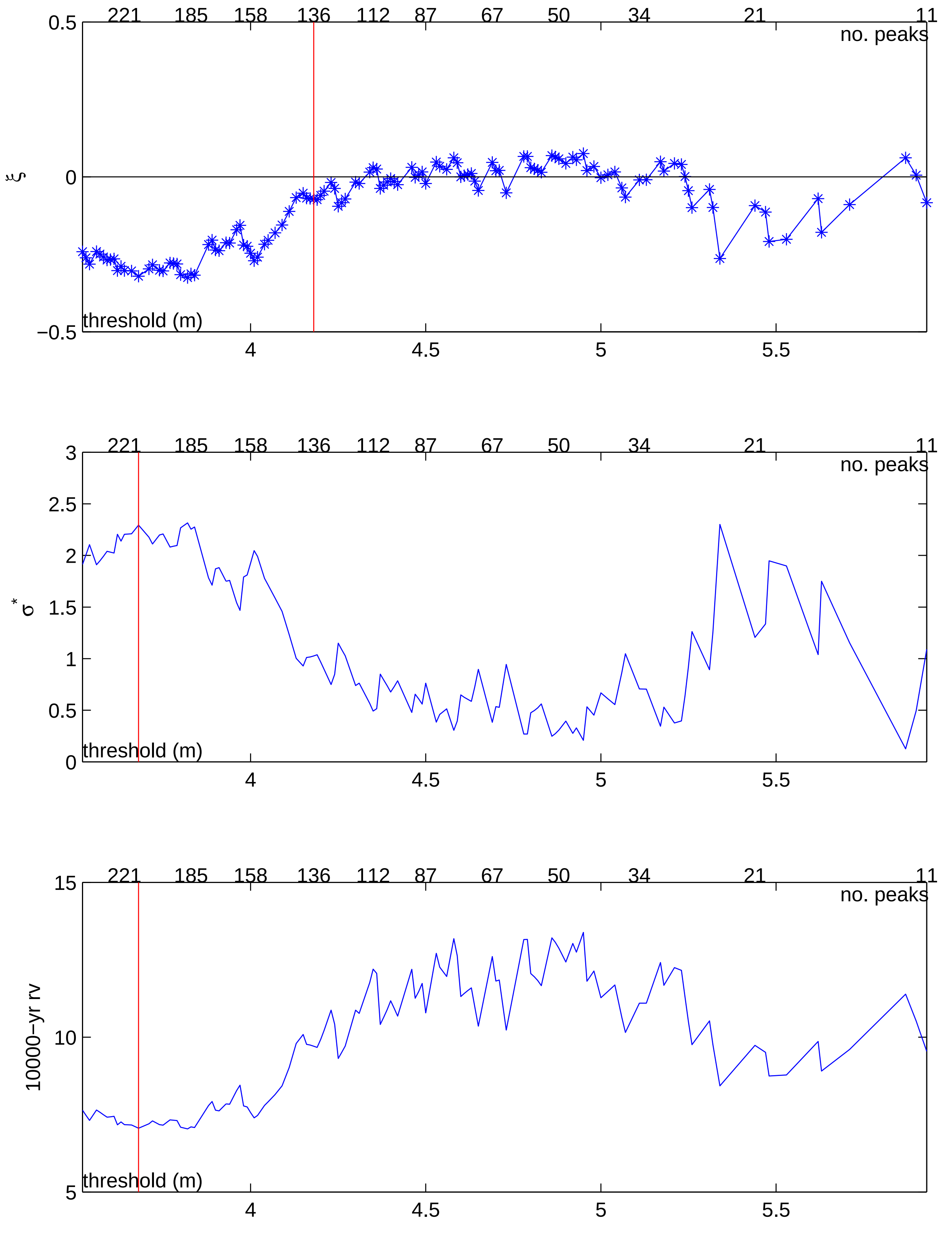


POT/GPD threshold choice
buoy: ELD

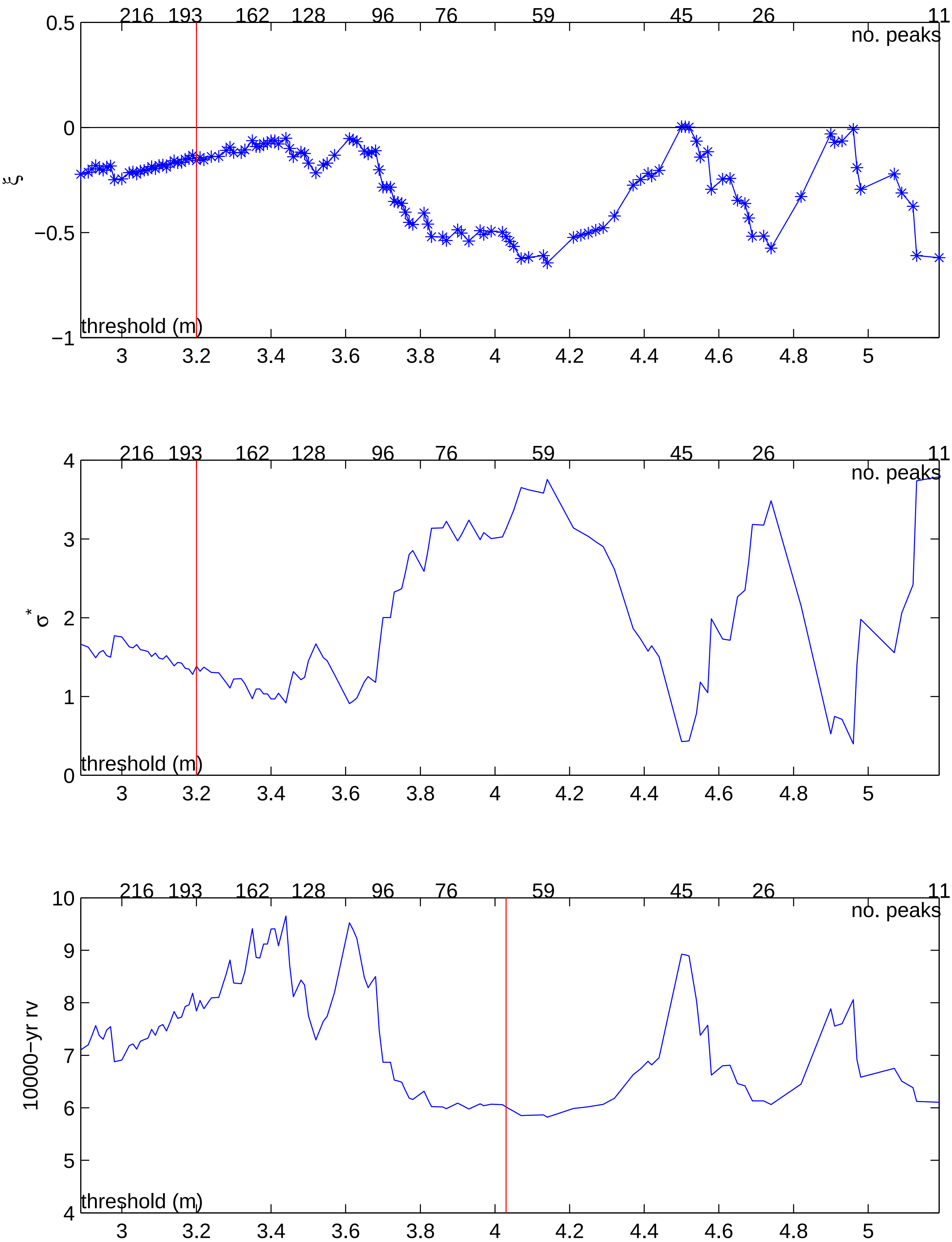
POT/GPD threshold choice buoy: ELD	H4803.30	gw1eldwa
WL | DELFT HYDRAULICS		Fig. 4.2b



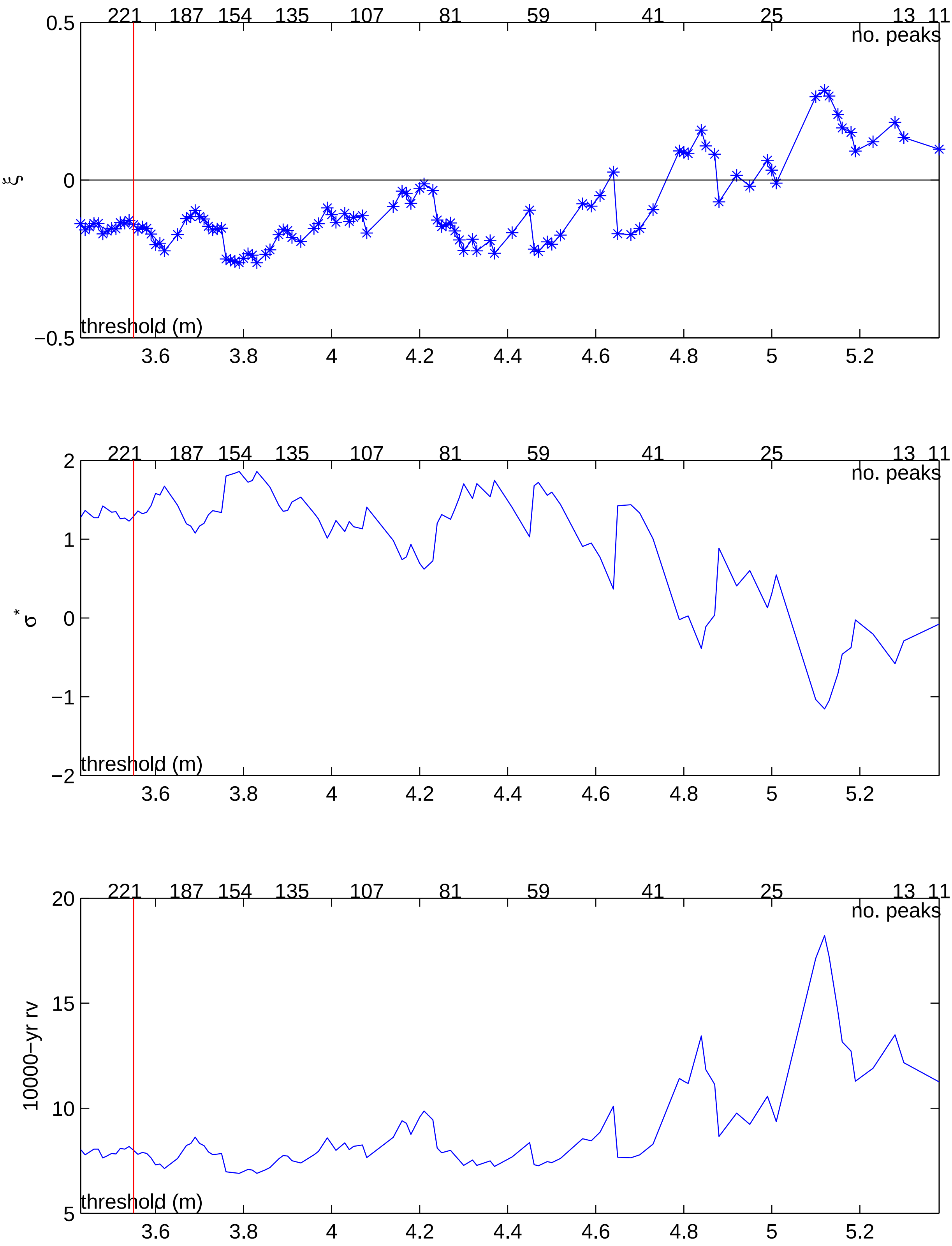
POT/GPD threshold choice buoy: K13	H4803.30	gw1k13wa
WL | DELFT HYDRAULICS		Fig. 4.2c



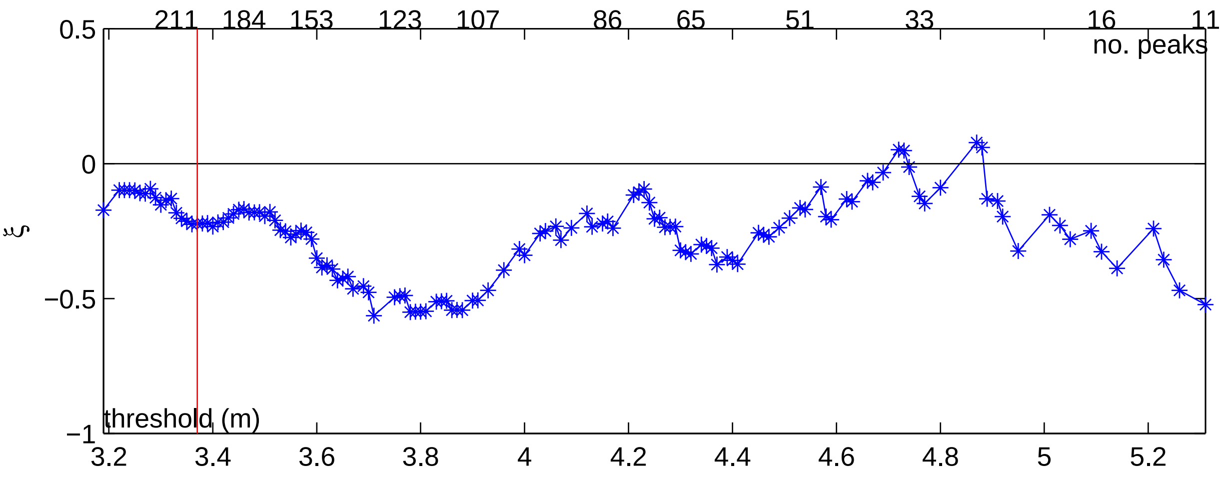
POT/GPD threshold choice buoy: YM6	H4803.30	gw1ym6wa
WL | DELFT HYDRAULICS		Fig. 4.2d

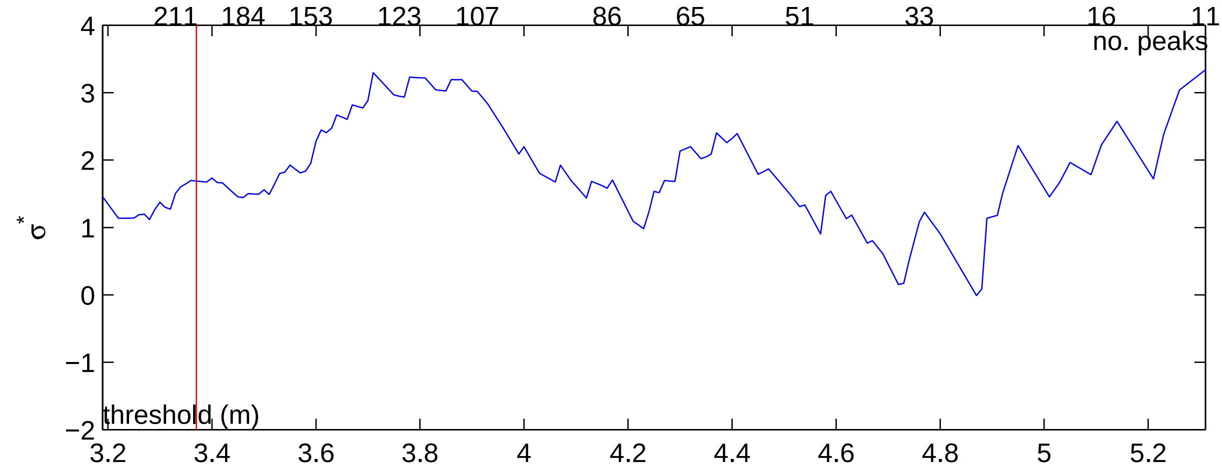


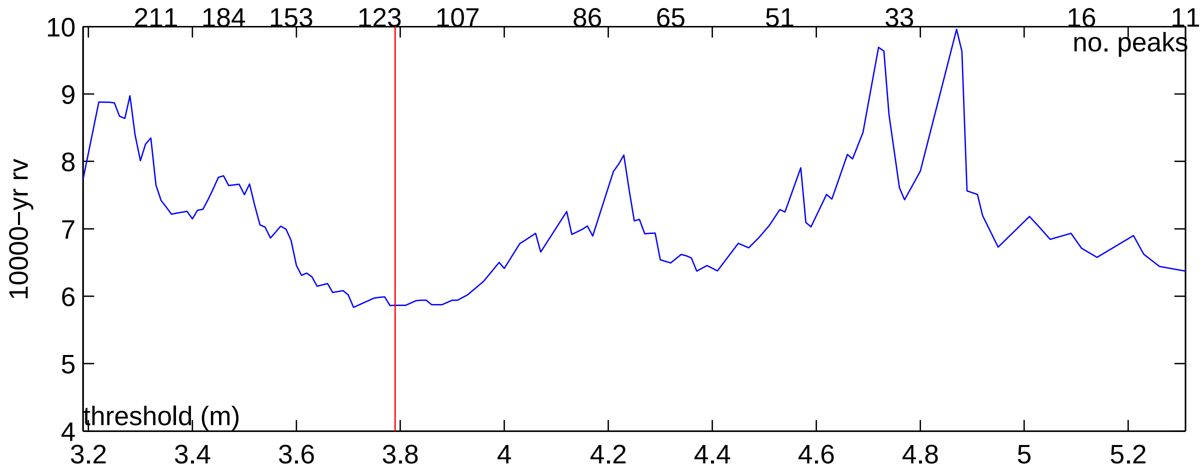
POT/GPD threshold choice buoy: MPN	H4803.30	gw1mpnwa
WL | DELFT HYDRAULICS		Fig. 4.2e



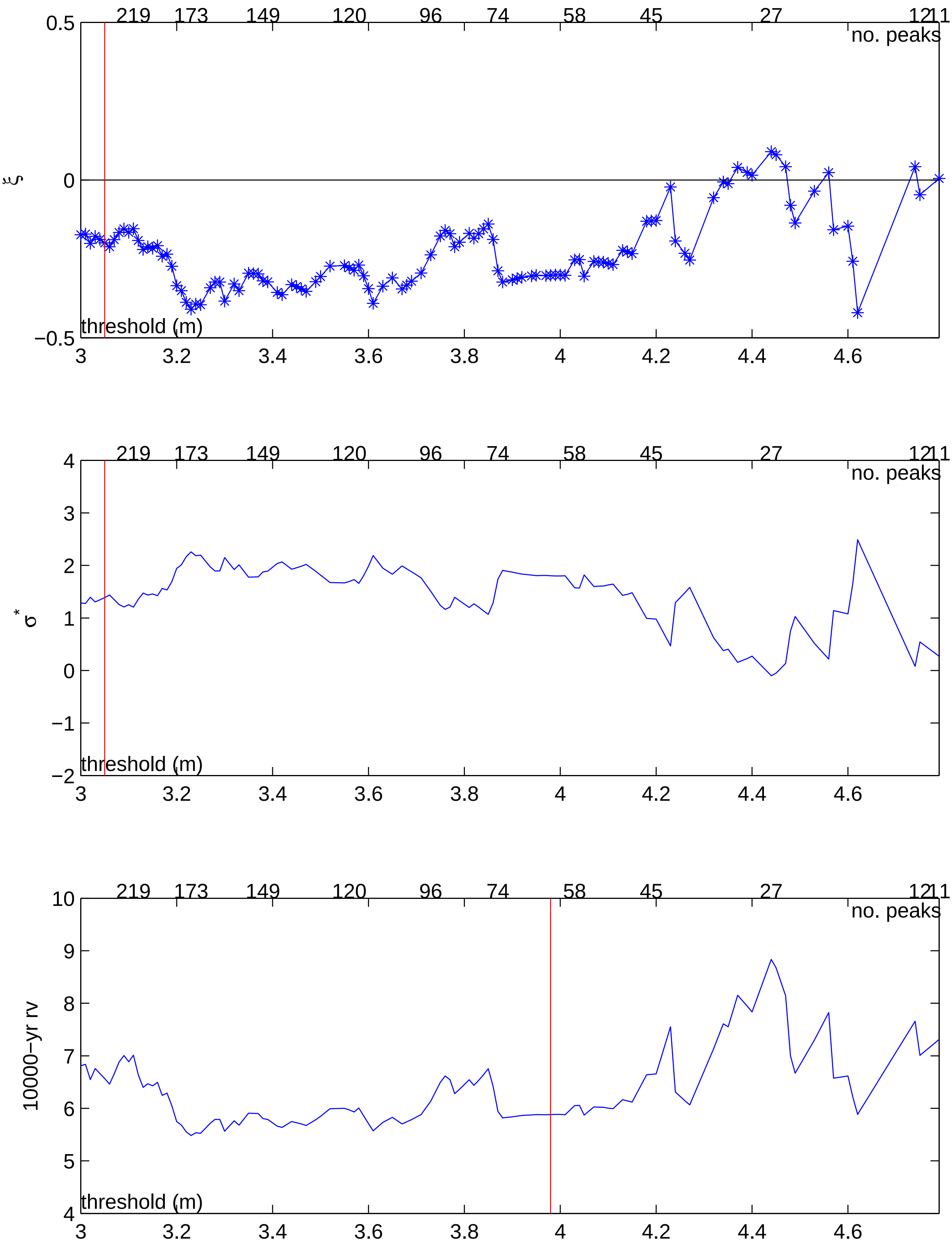
POT/GPD threshold choice buoy: EUR	H4803.30	gw1eurwa
WL | DELFT HYDRAULICS		Fig. 4.2f



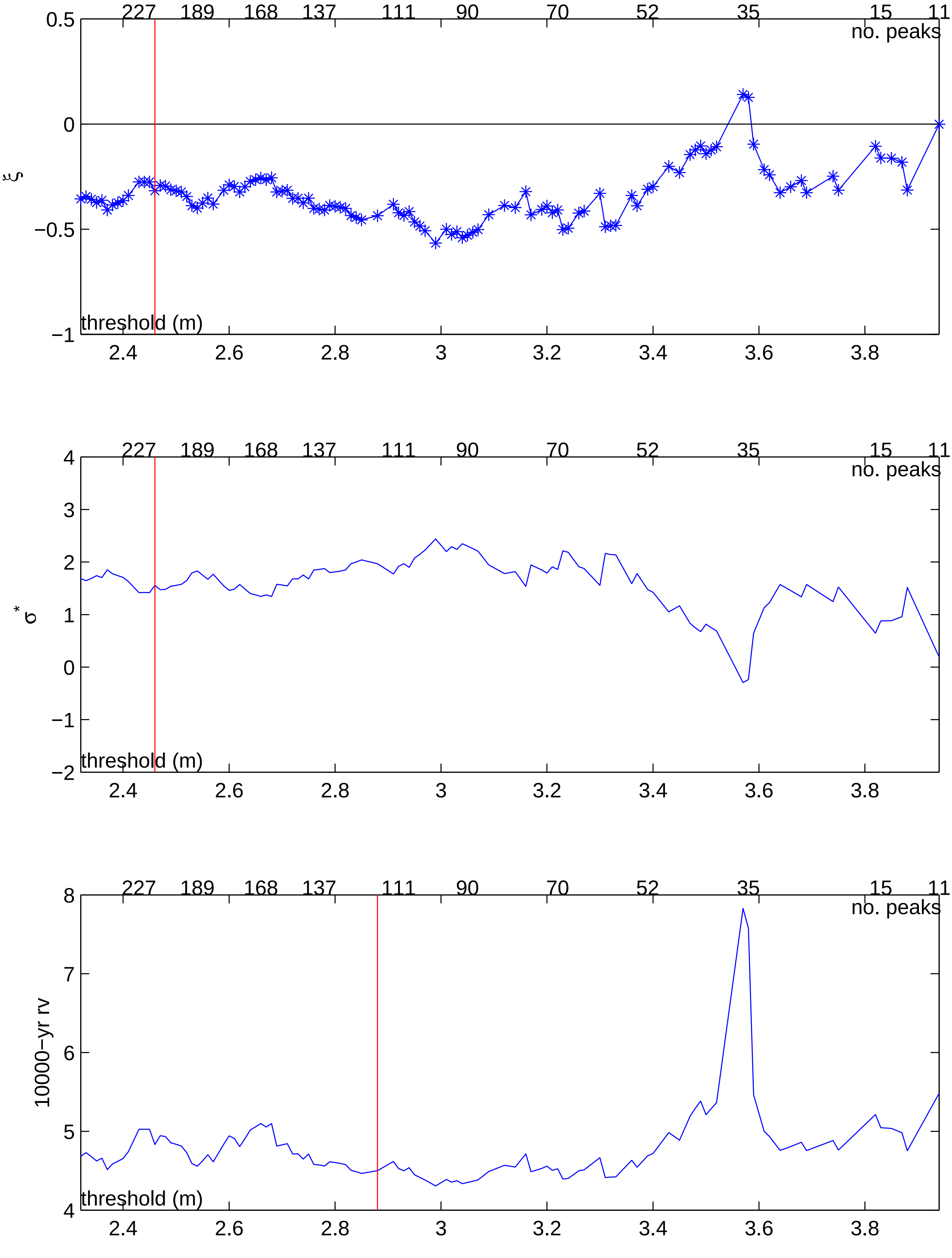




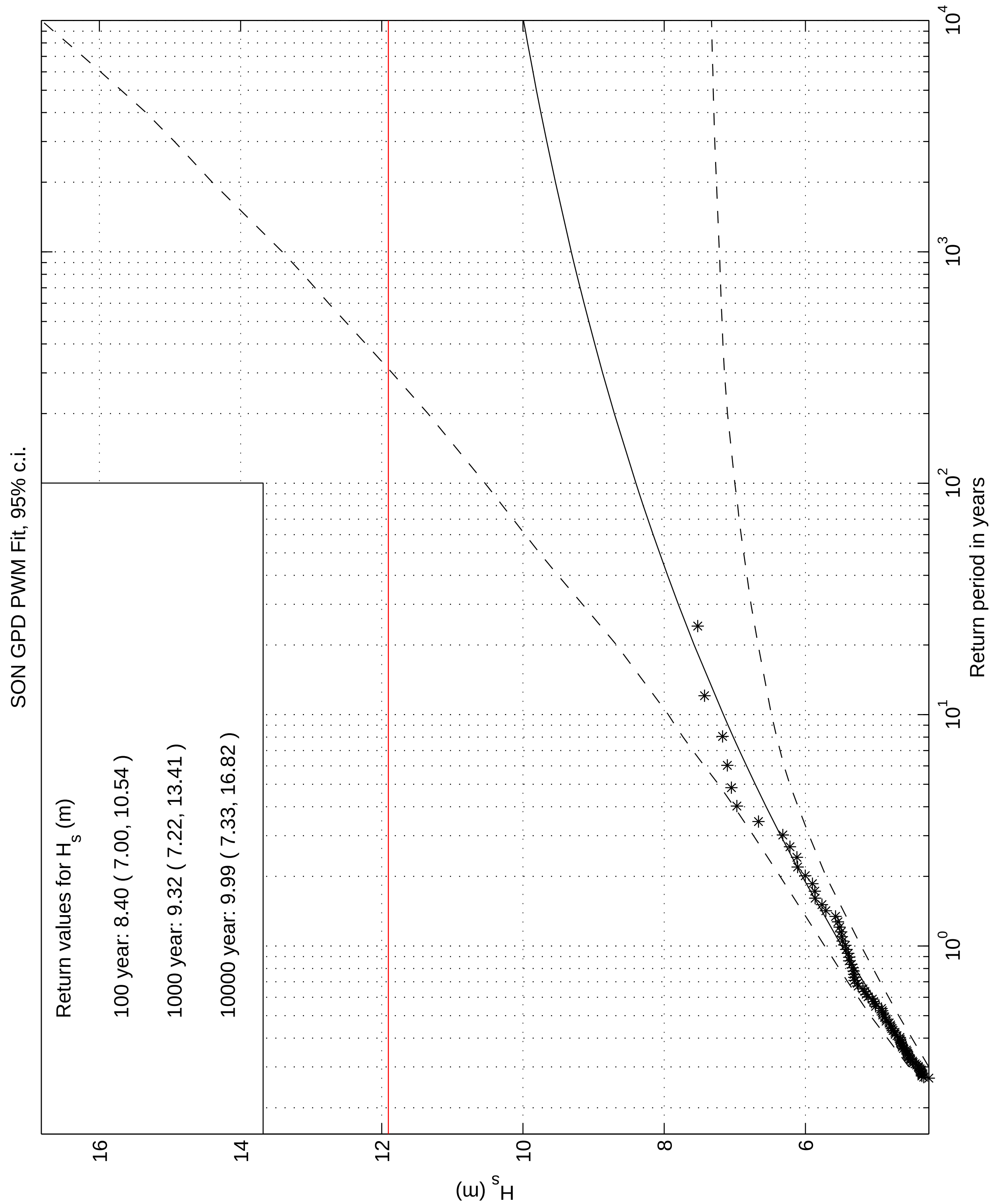
POT/GPD threshold choice buoy: LEG	H4803.30	gw1legws
WL | DELFT HYDRAULICS		Fig. 4.2g



POT/GPD threshold choice buoy: SWB	H4803.30	gw1swbwr
WL | DELFT HYDRAULICS		Fig. 4.2h

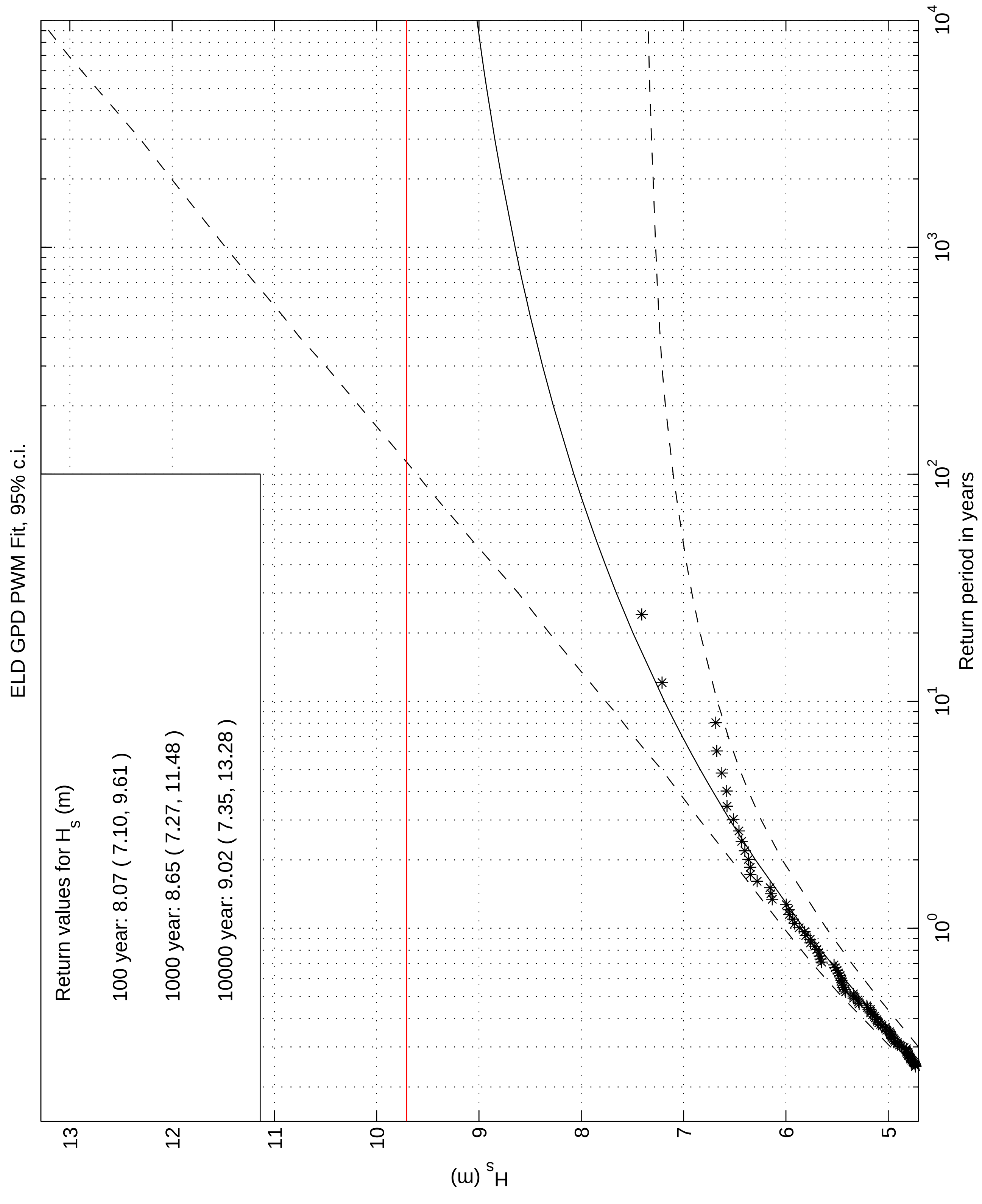


POT/GPD threshold choice buoy: SCW	H4803.30	gwlscwwr
WL | DELFT HYDRAULICS		Fig. 4.2i

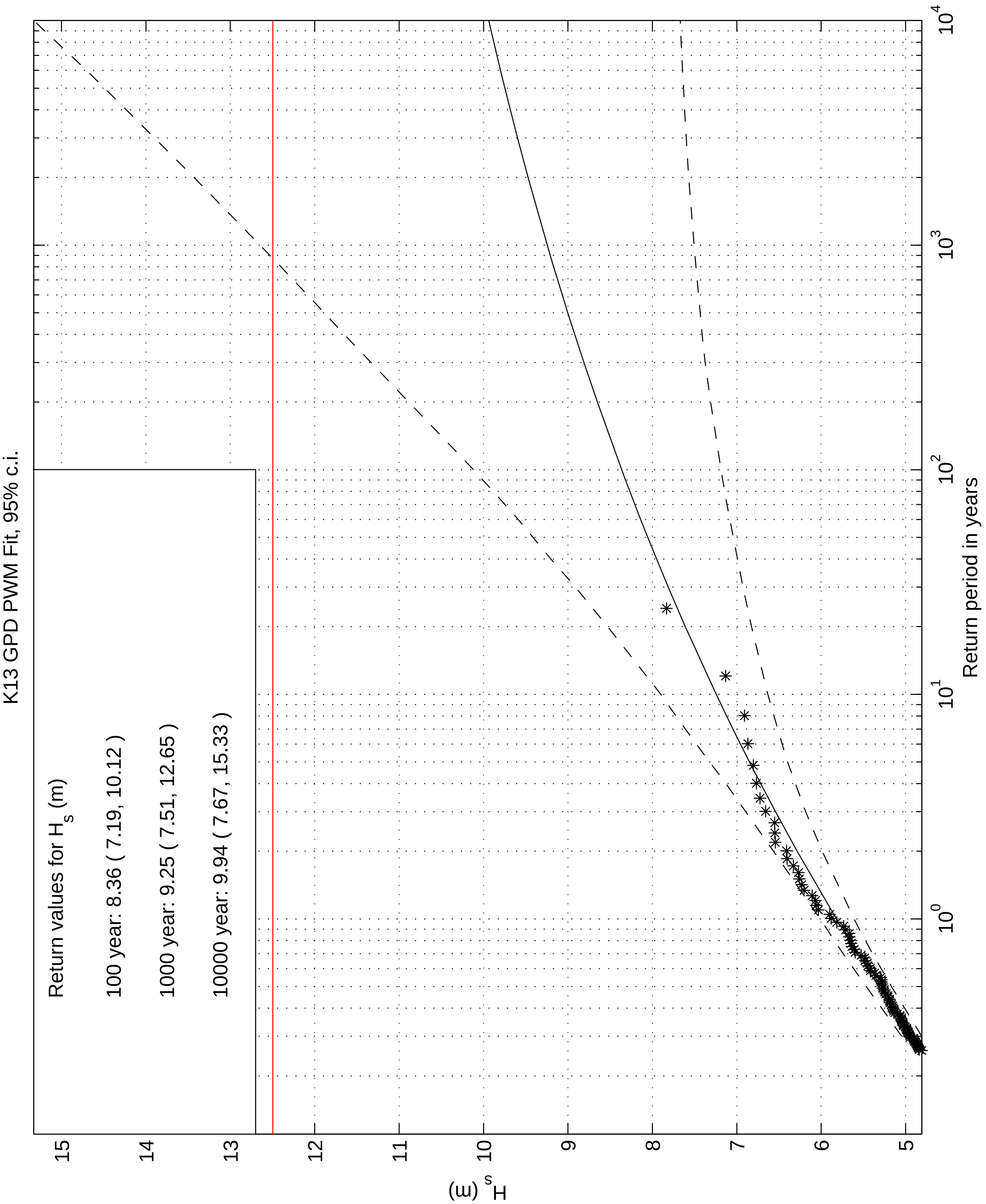


GPD fit
buoy: SON

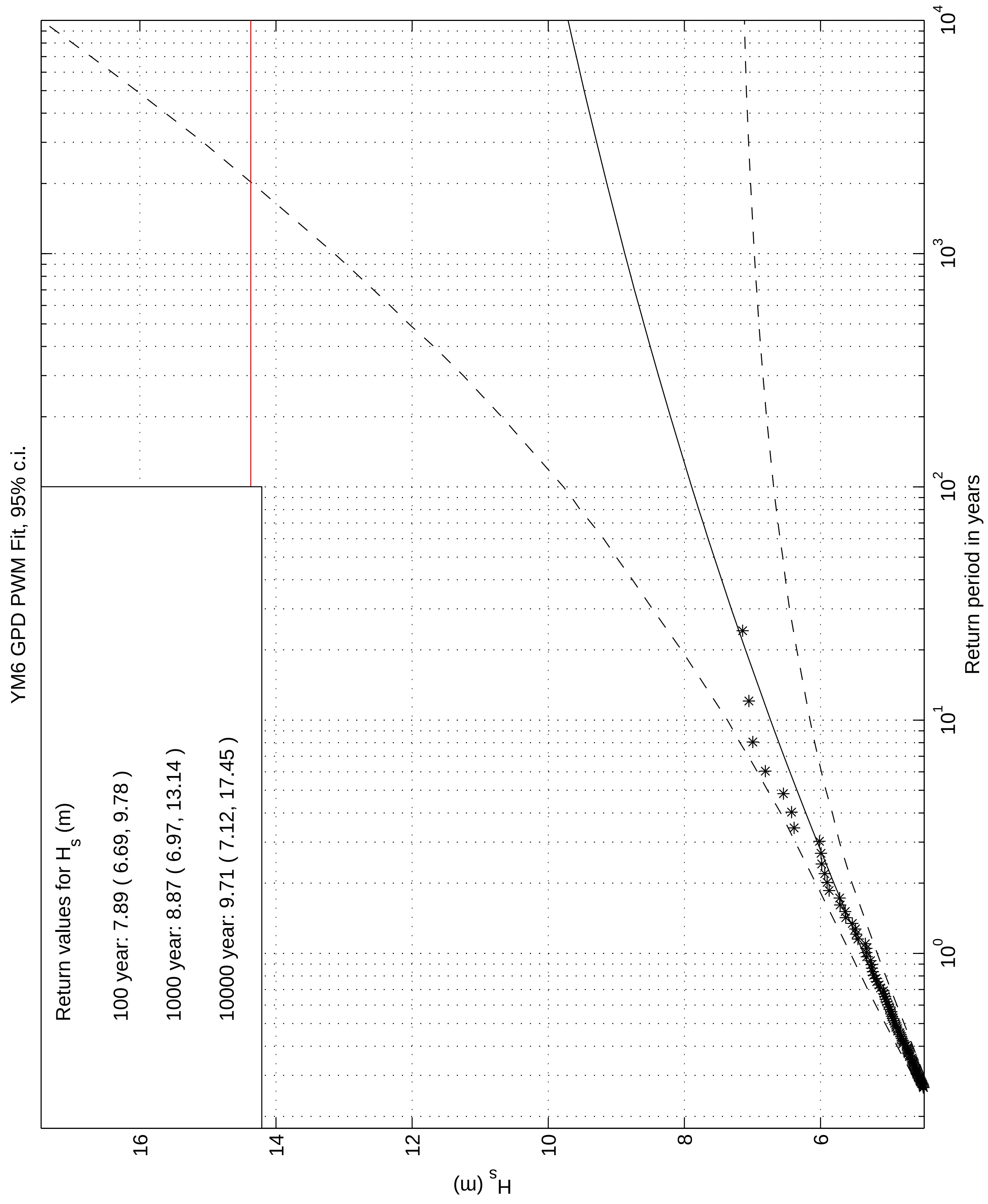
	H4803.30	gw1sonwa
WL | DELFT HYDRAULICS		Fig. 4.3a



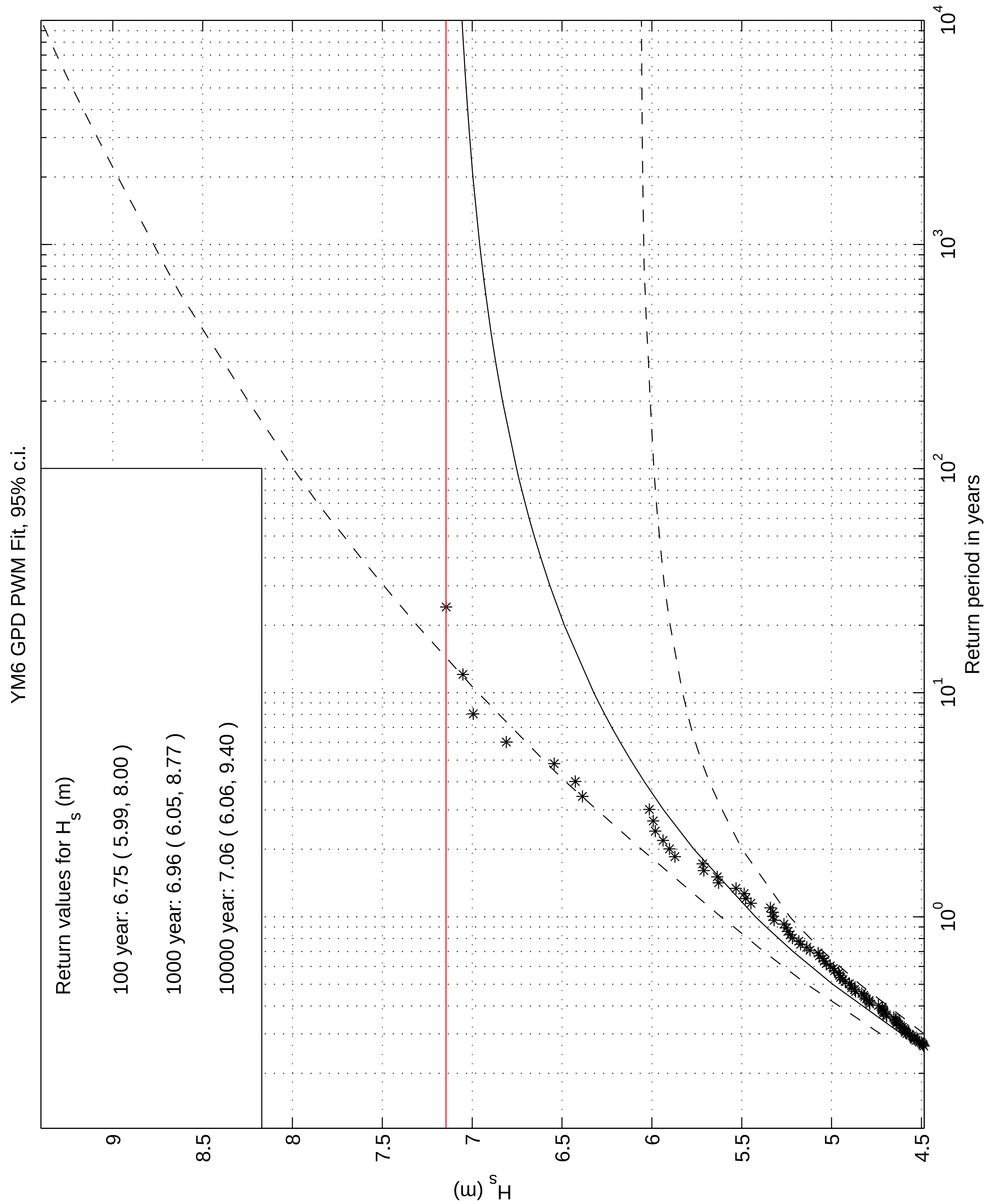
GPD fit buoy: ELD	H4803.30	gw1eldwa
WL | DELFT HYDRAULICS		Fig. 4.3b



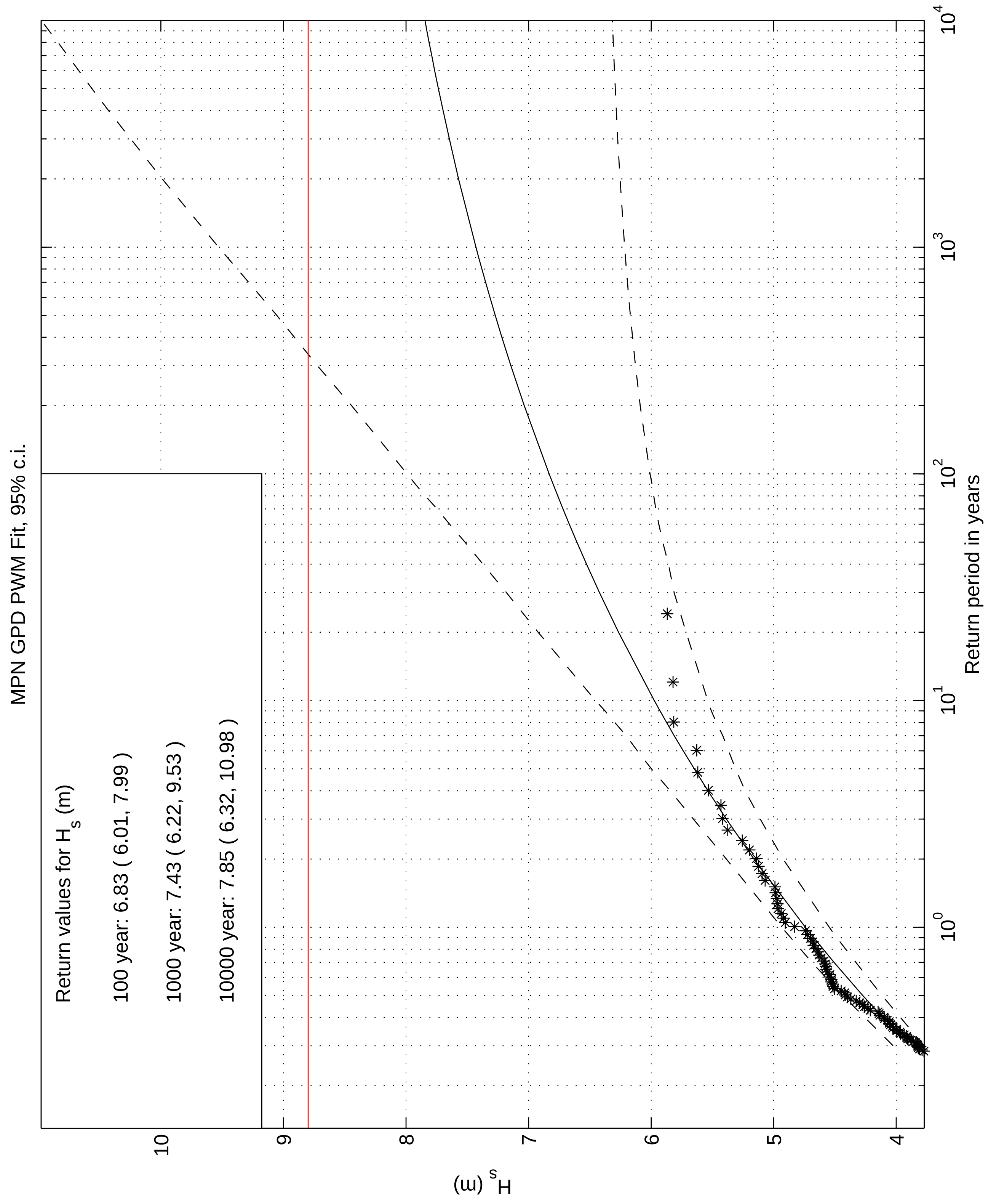
GPD fit buoy: K13	H4803.30	gw1k13wa
WL | DELFT HYDRAULICS		Fig. 4.3c



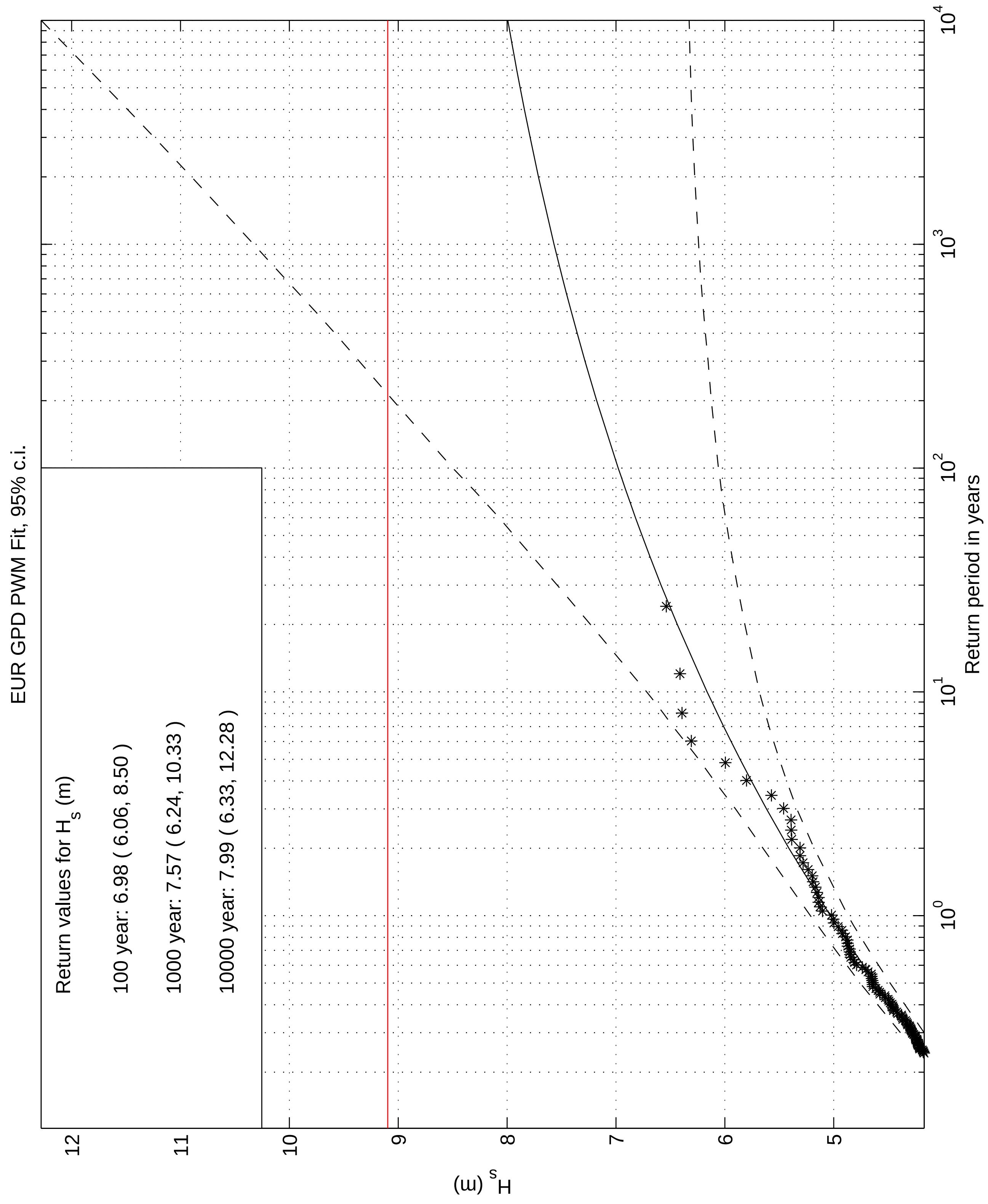
GPD fit buoy: YM6	H4803.30	gw1ym6wa
WL | DELFT HYDRAULICS		Fig. 4.3d



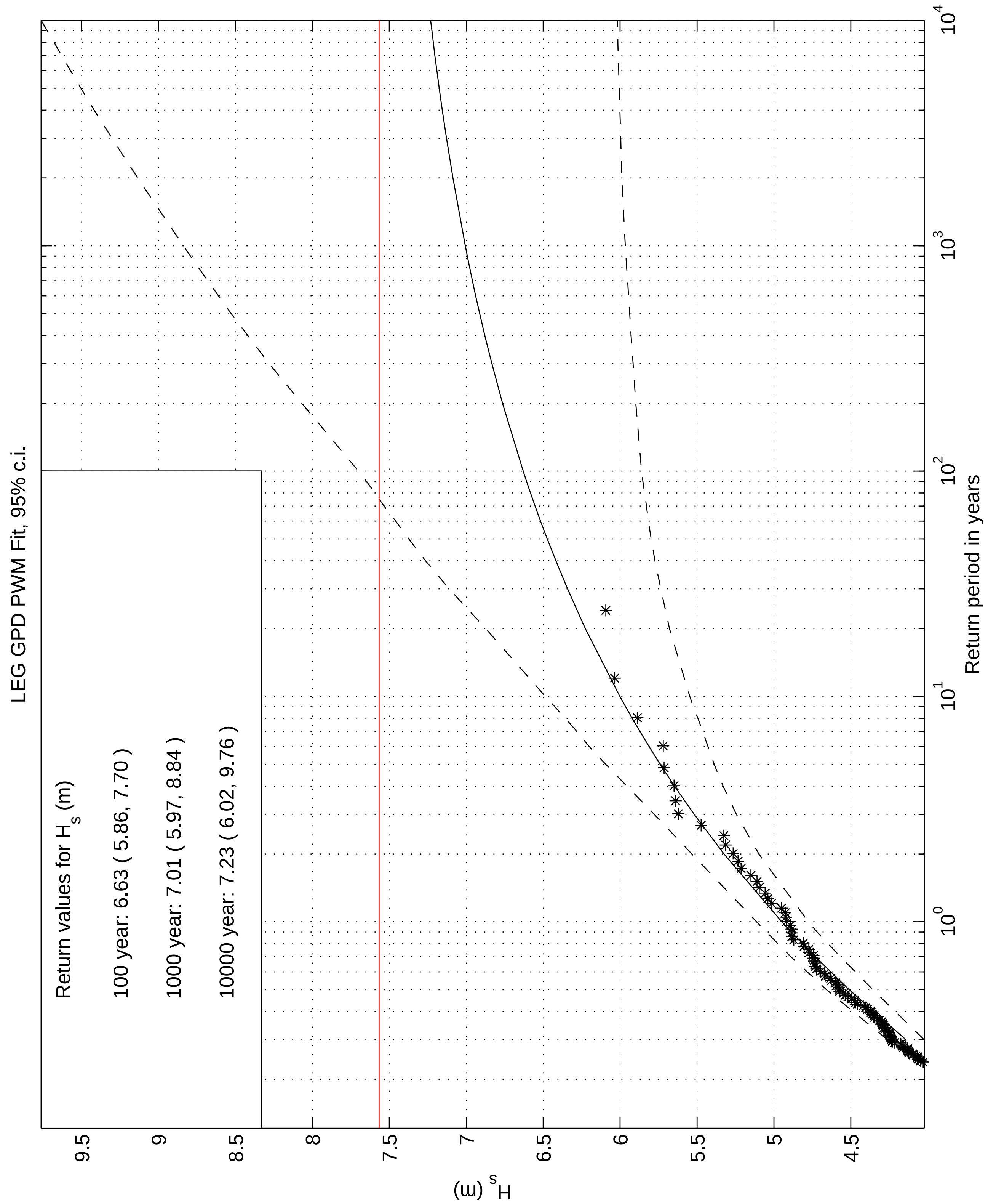
GPD fit buoy: YM6	H4803.30	gw1ym6wa
WL | DELFT HYDRAULICS		Fig. 4.3d2



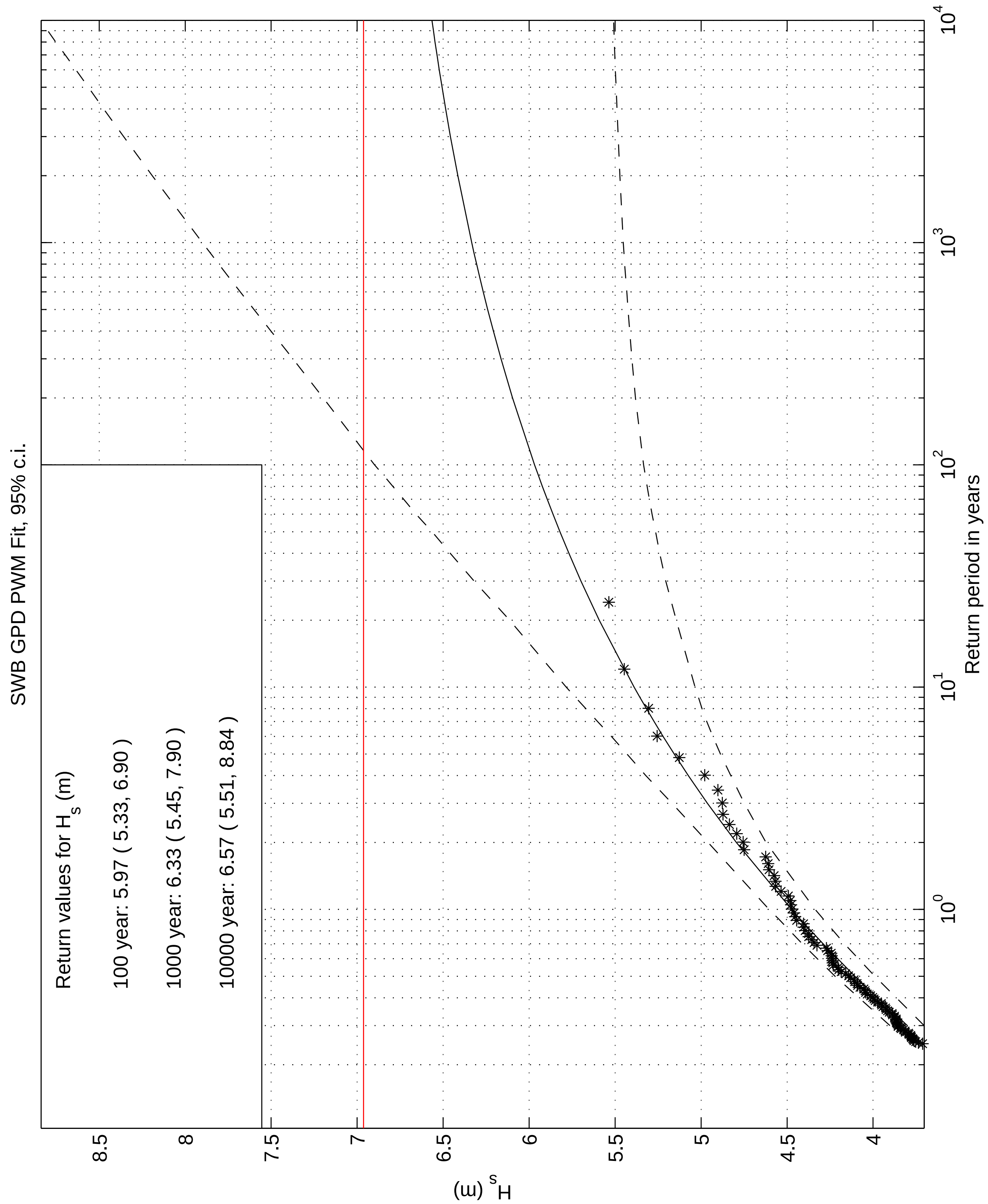
GPD fit buoy: MPN	H4803.30	gw1mpnwa
WL | DELFT HYDRAULICS		Fig. 4.3e



GPD fit buoy: EUR	H4803.30	gw1eurwa
WL | DELFT HYDRAULICS		Fig. 4.3f



GPD fit buoy: LEG	H4803.30	gw1legws
WL | DELFT HYDRAULICS		Fig. 4.3g

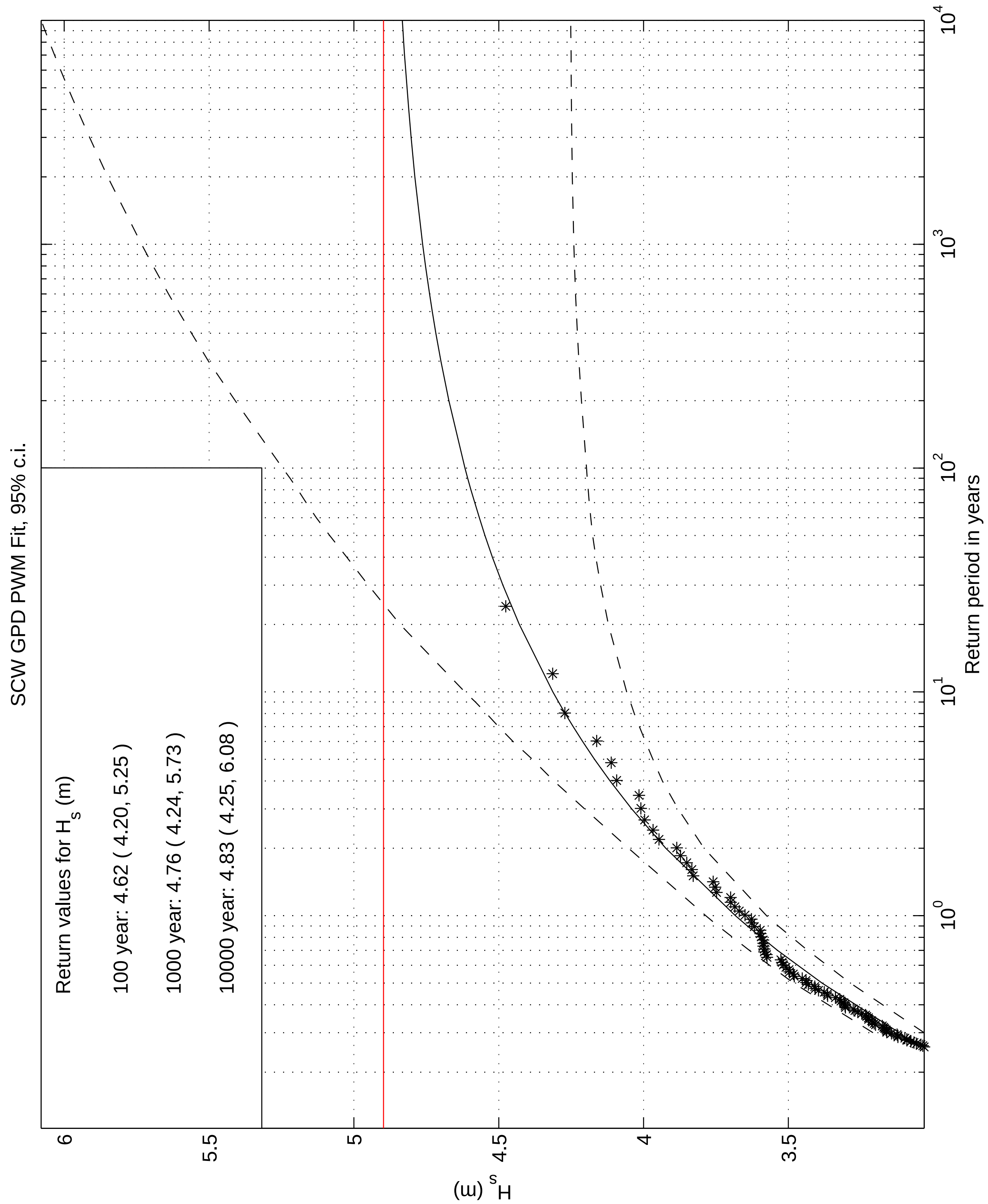


GPD fit
buoy: SWB

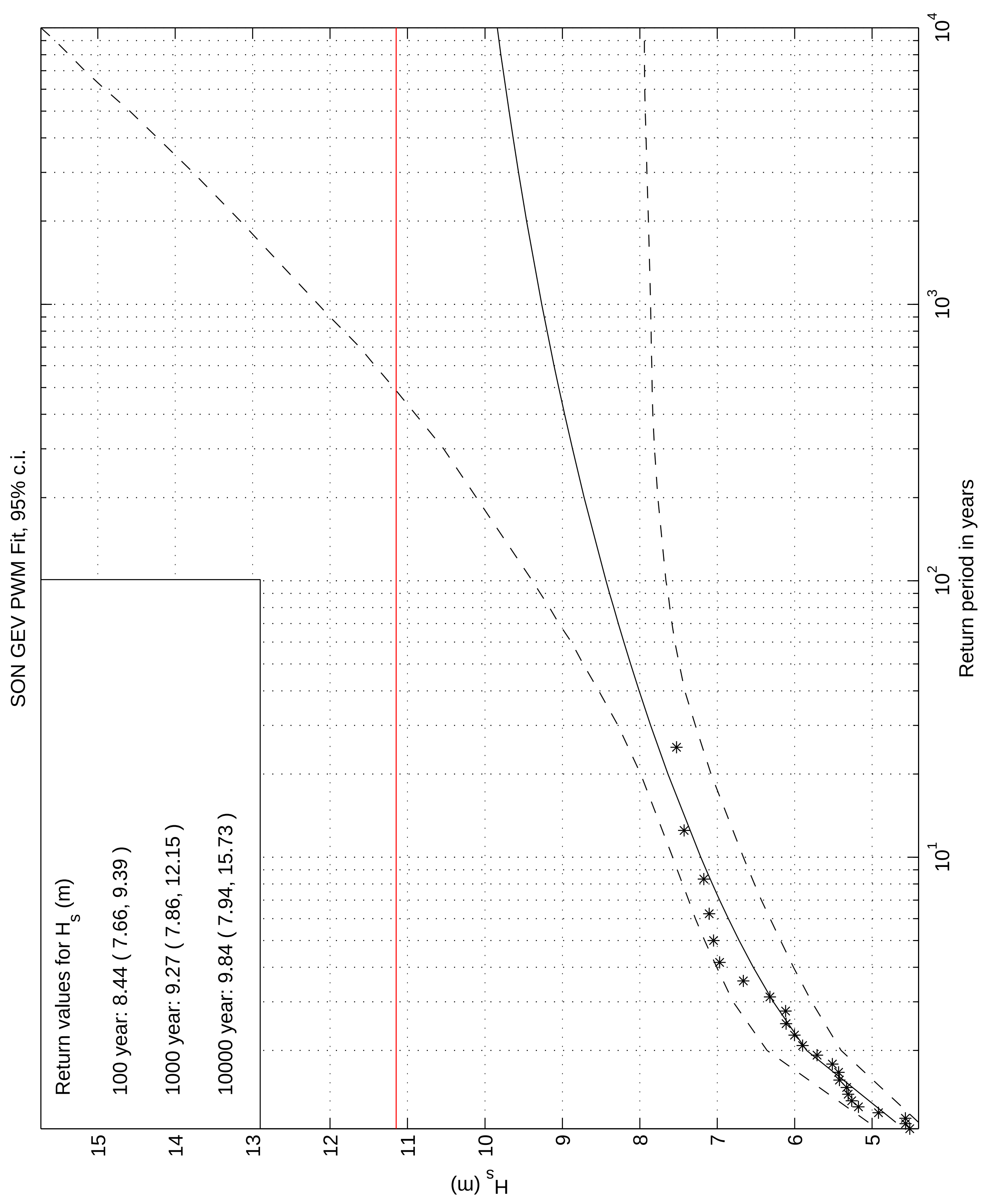
H4803.30	gw1swbwr

WL | DELFT HYDRAULICS

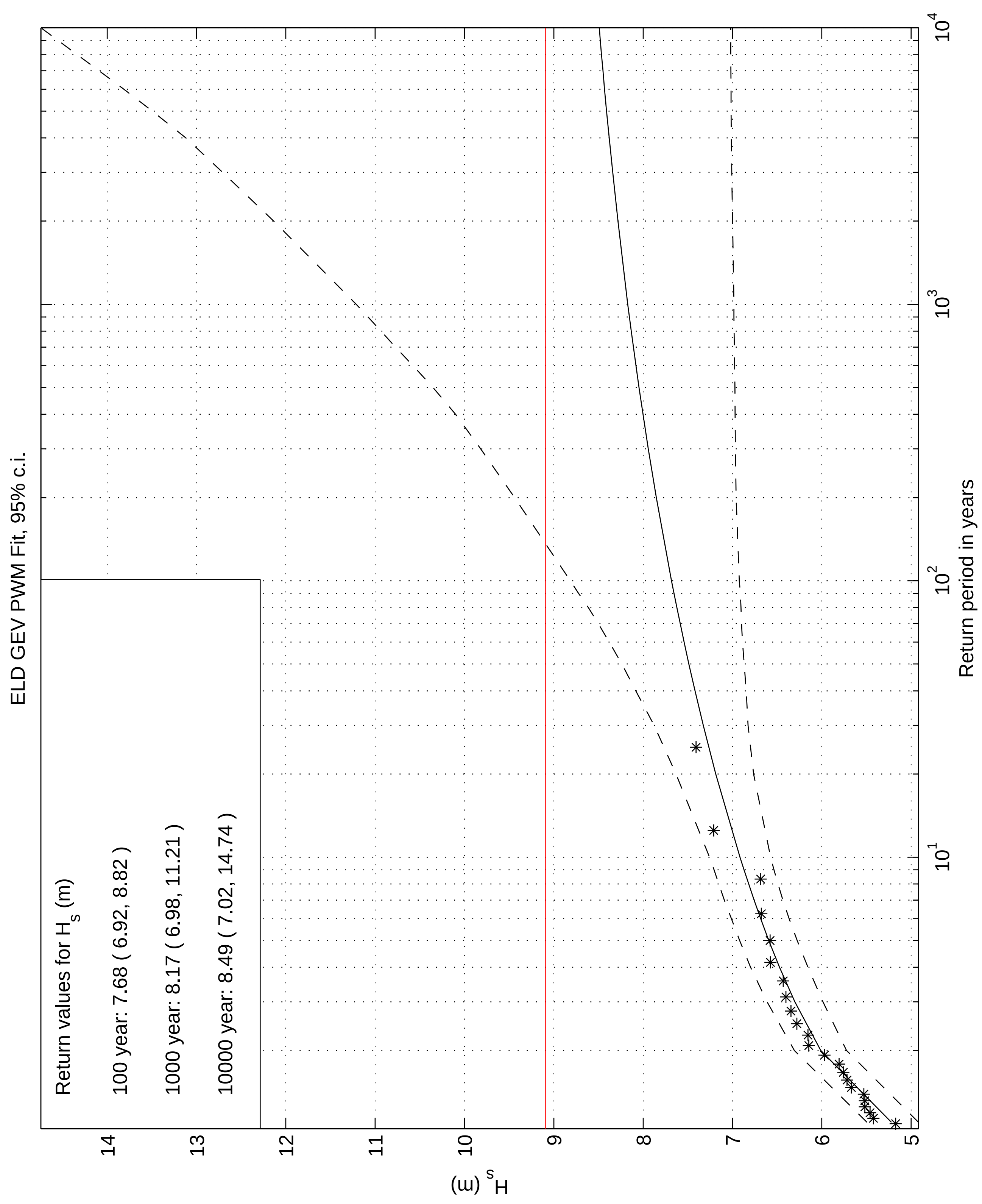
Fig. 4.3h



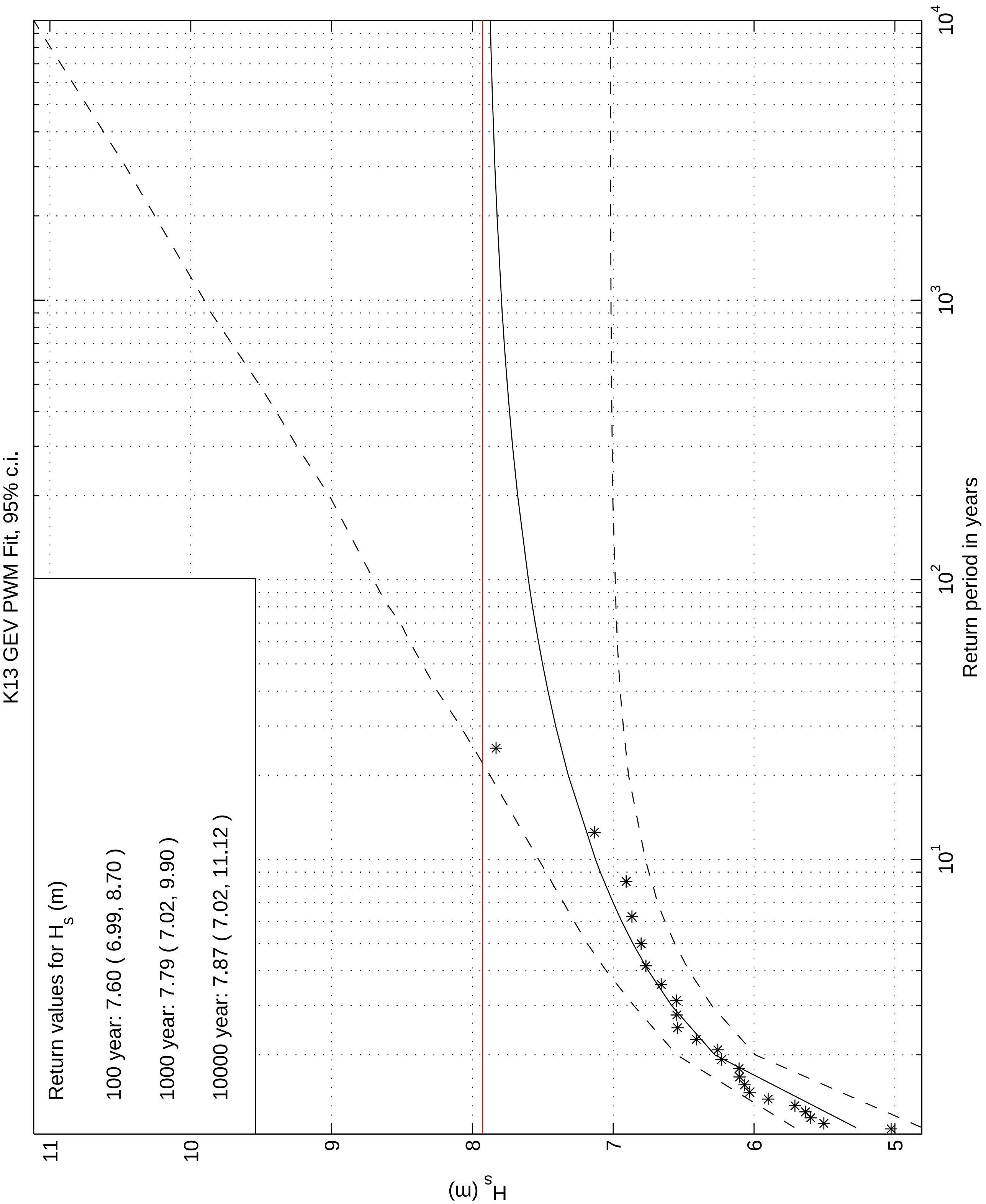
GPD fit buoy: SCW	H4803.30	gw1scwwr
WL | DELFT HYDRAULICS		Fig. 4.3i



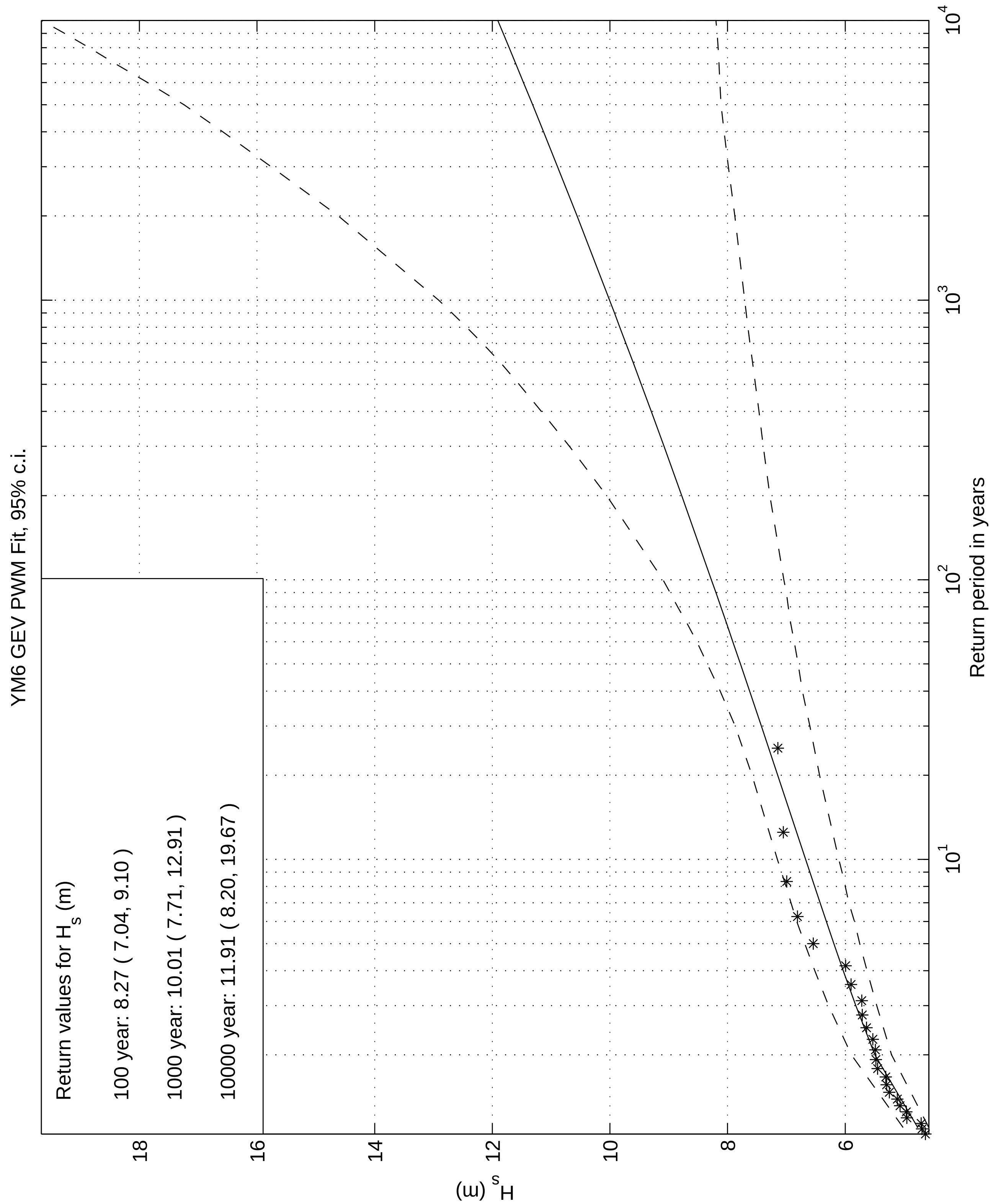
GEV fit buoy: SON	H4803.30	gw1sonwa
WL | DELFT HYDRAULICS		Fig. 4.4a



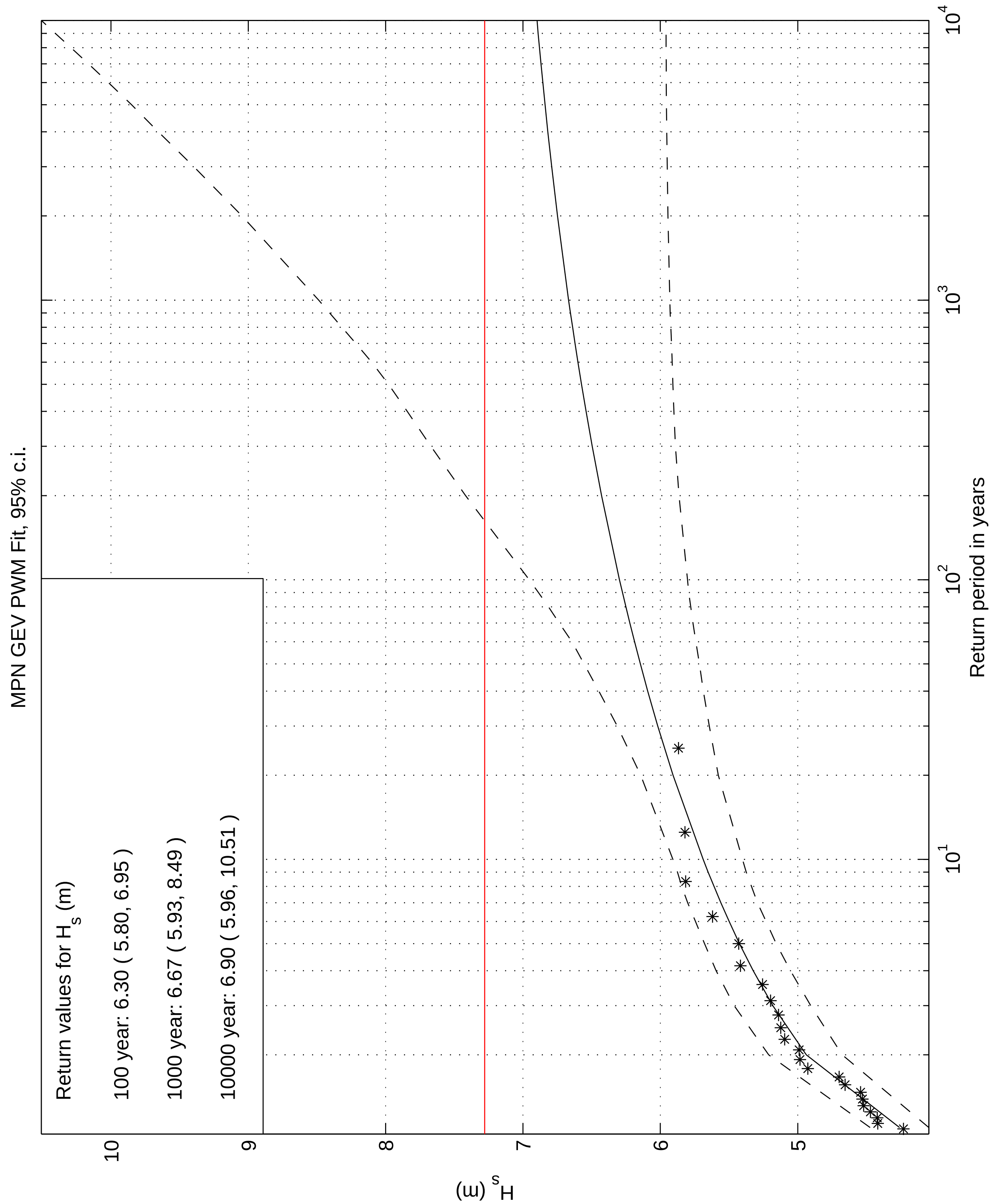
GEV fit buoy: ELD	H4803.30	gw1eldwa
WL | DELFT HYDRAULICS		Fig. 4.4b



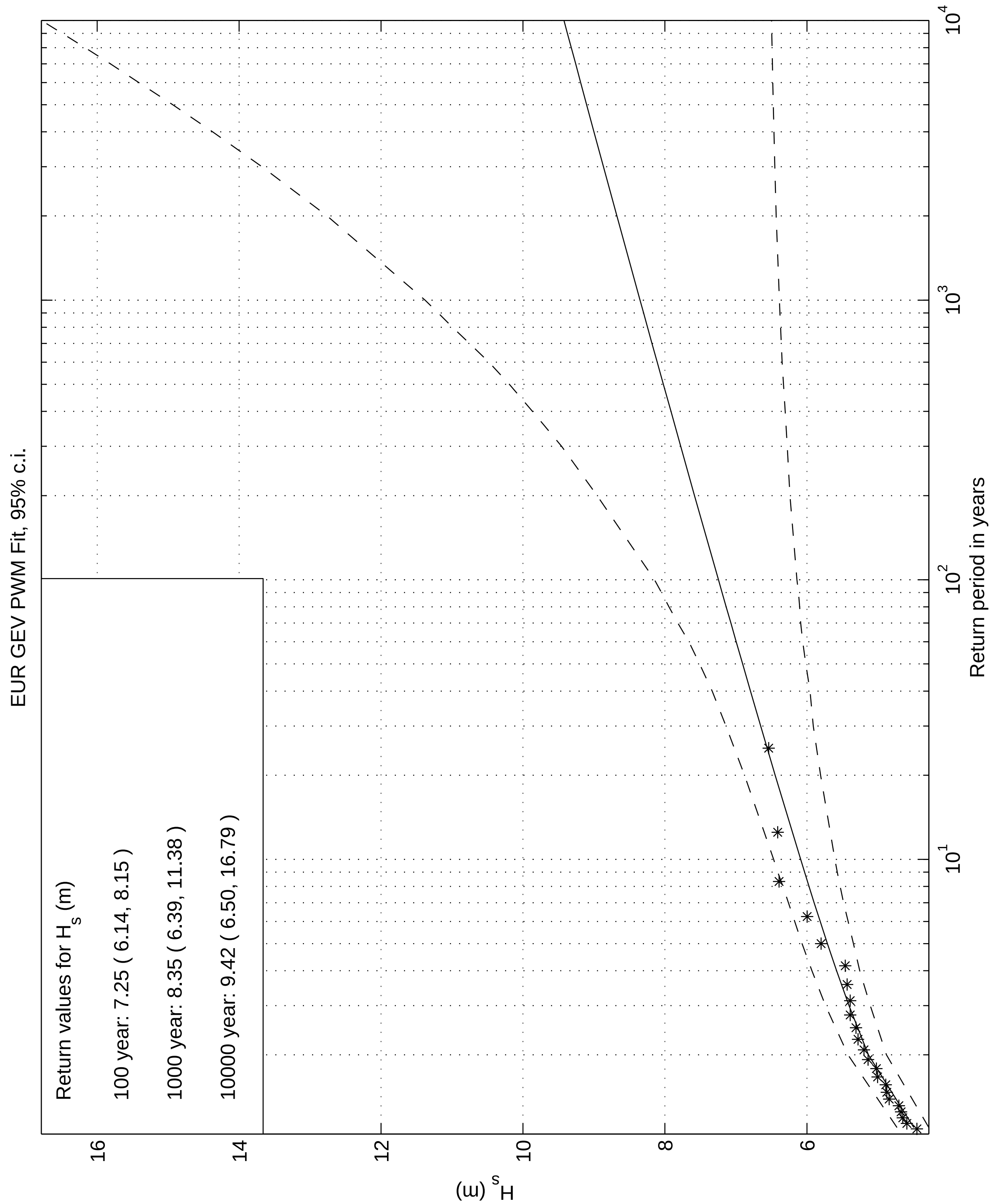
GEV fit buoy: K13	H4803.30	gw1k13wa
WL | DELFT HYDRAULICS		Fig. 4.4c



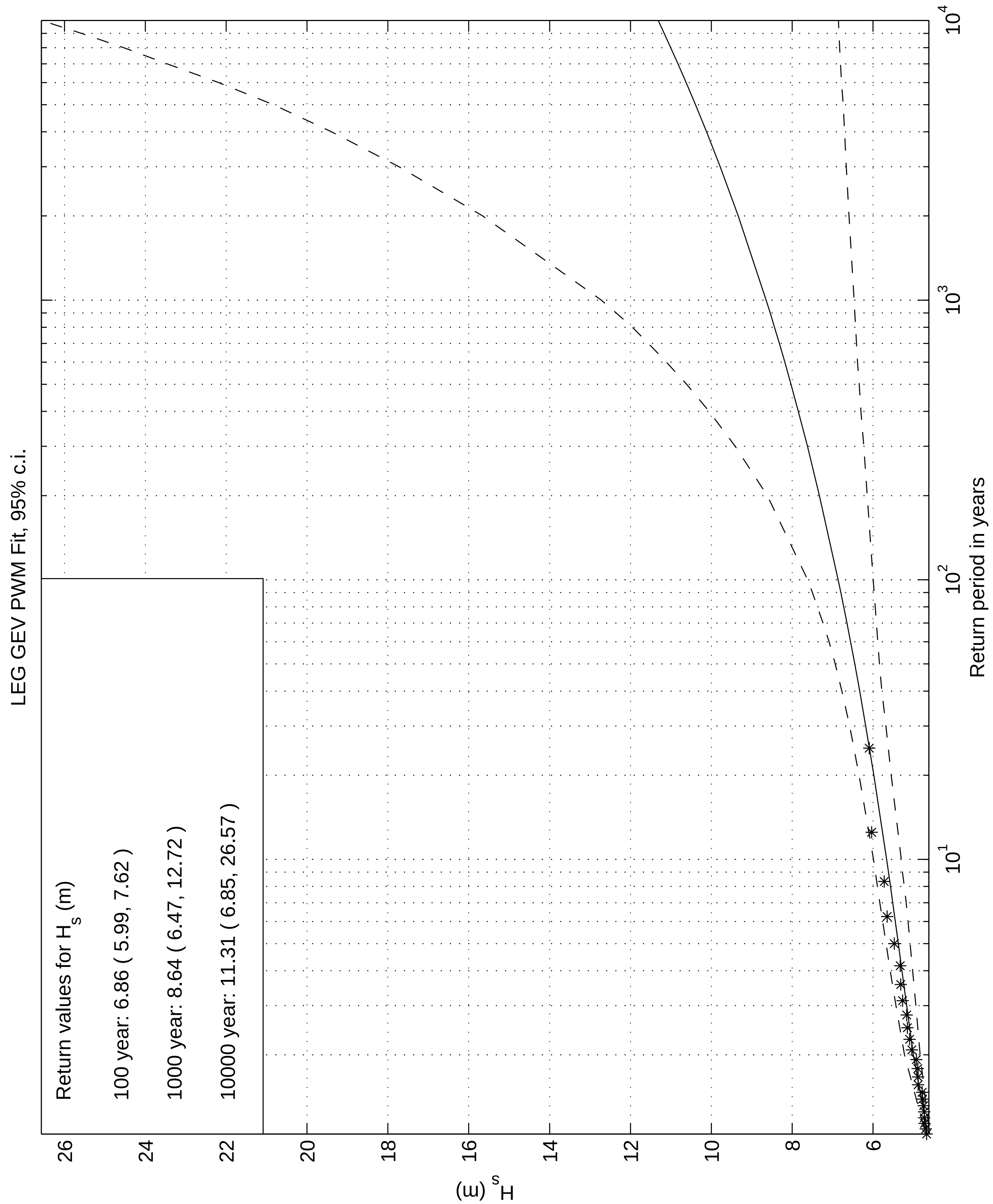
GEV fit buoy: YM6	H4803.30	gw1ym6wa
WL | DELFT HYDRAULICS		Fig. 4.4d



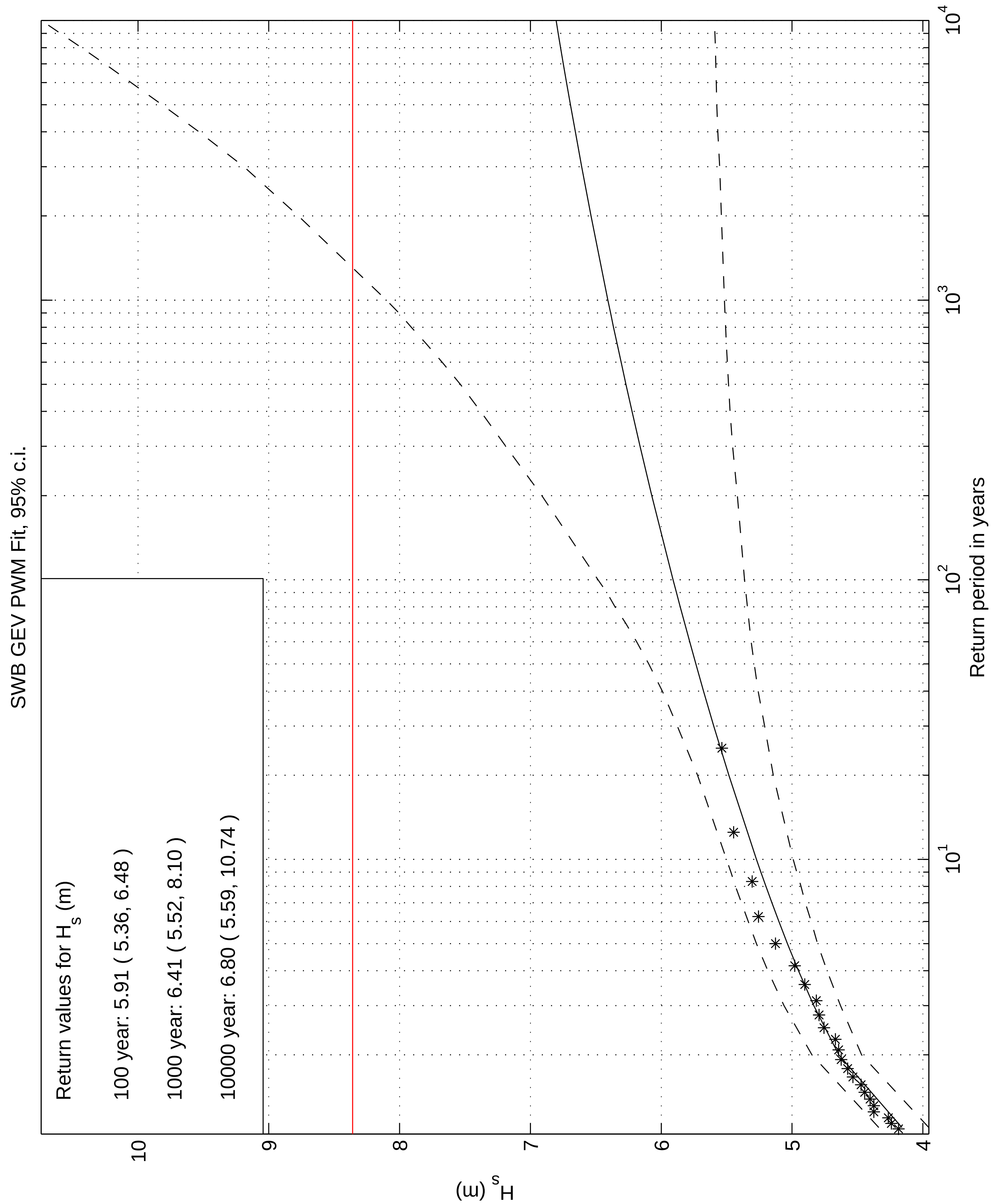
GEV fit buoy: MPN	H4803.30	gw1mpnwa
WL | DELFT HYDRAULICS		Fig. 4.4e



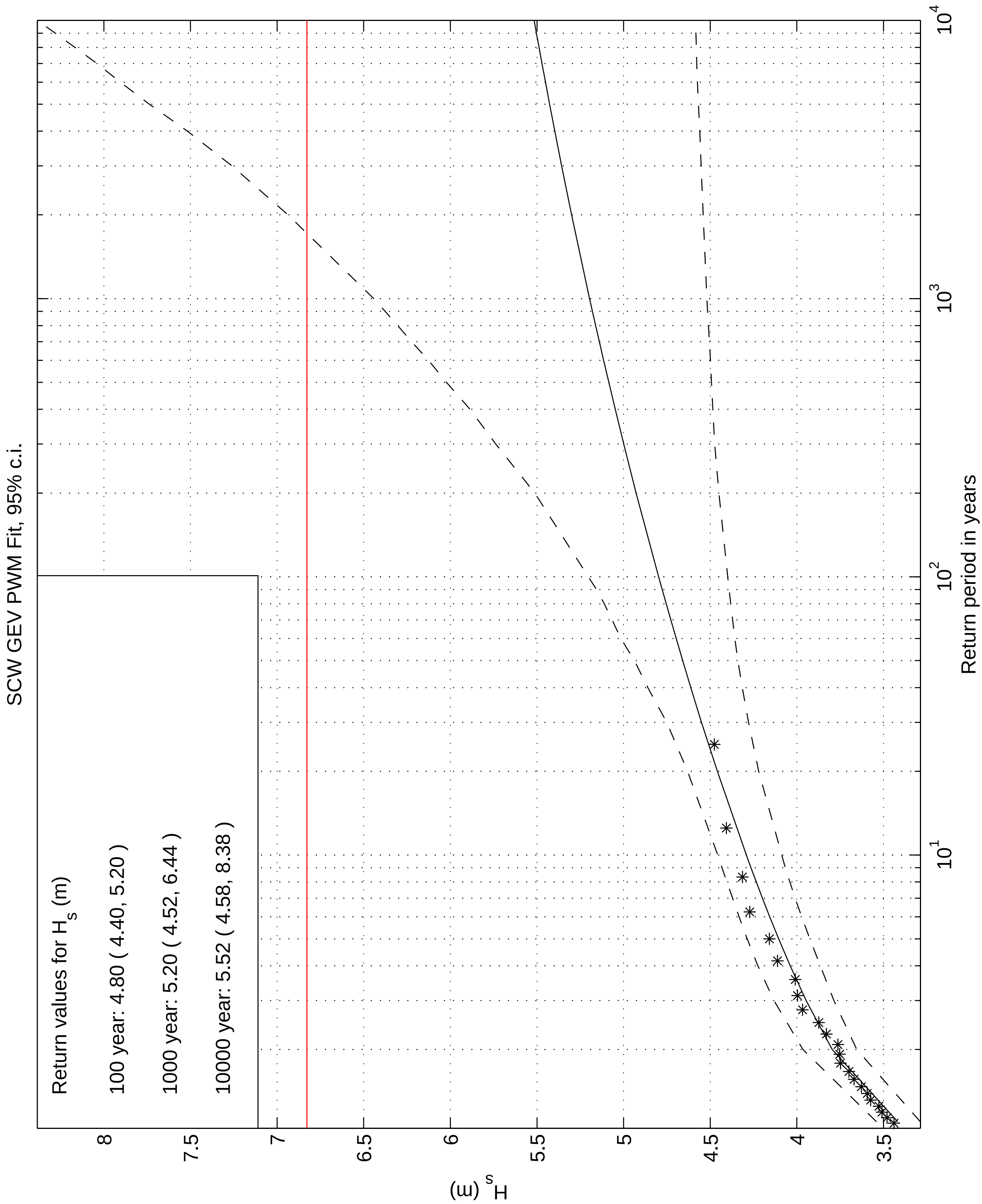
GEV fit buoy: EUR	H4803.30	gw1eurwa
WL | DELFT HYDRAULICS		Fig. 4.4f



GEV fit buoy: LEG	H4803.30	gw1legws
WL | DELFT HYDRAULICS		Fig. 4.4g



GEV fit buoy: SWB	H4803.30	gw1swbwr
WL | DELFT HYDRAULICS		Fig. 4.4h



GEV fit buoy: SCW	H4803.30	gw1scwwr
WL | DELFT HYDRAULICS		Fig. 4.4i