Available online at www.sciencedirect.com
SCIENCE DIRECT®
Coastal Engineering 52 (2005) 63–77
Coastal Engineering
An International Journal for Coastal, Harbour and Offshore Engineers
www.elsevier.com/locate/coastaleng

Extreme low and high water levels

Rodney J. Sobey

Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK

Received 9 December 2003; received in revised form 16 September 2004; accepted 17 September 2004
Available online 5 November 2004

Abstract

A methodology for the prediction of low end and high end extremes in sustained water level is established. The observational data base is a long-duration sequence of monthly high and low extremes. The data identifies a deterministic trend attributable to Mean Sea Level rise and the nominal 19-year forcing in the astronomical tide. The data is pre-conditioned to remove these trends, defining a net data series suitable for extreme value analysis. Context-specific issues in the extreme value analysis are identified and resolved. These include probability model compatibility with elevation datums, rational estimation of the distribution parameters, and the estimation of confidence limits. The predictions for extreme low and high water levels are both real time and return period dependent.

Keywords: Extreme value analysis; High water level; Low water level; Numerical algorithm; Sea level rise; 19-year cycle

1. Introduction

A rational evaluation of marine climate at any site would include extreme value analyses of sustained water levels. In principle, this is classical extreme value analysis. However, there are some unique aspects of extreme value analysis for sustained water level. The interaction of simultaneous storm and astronomical tides is an important consideration. Storm tides, defined here to include wave setup contributions, may lead to higher than normal tides and also lower than normal tides.

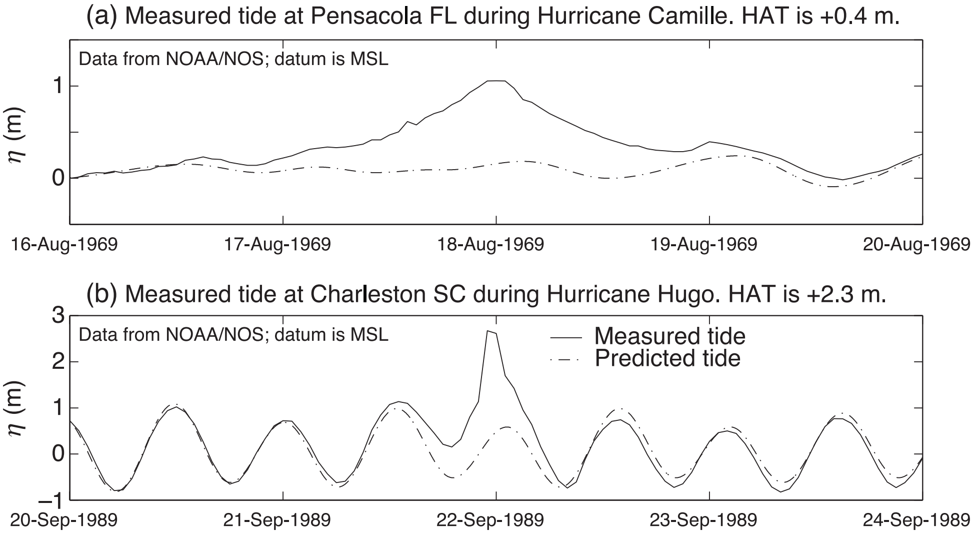
Trends in sea level are a concern over the duration of an historical data sequence. In addition, there is an interest in both the upper and lower extremes of the sustained water level. Investigations of coastal flooding and design and operation of marine facilities require estimates of the highest water level. The lowest water level is an important parameter in the context of navigation and cooling water intakes to power plants.

2. Observational evidence

To a greater or lesser extent, the astronomical tide is a significant part of any observation of sustained water level. Fig. 1a shows the observed and predicted tides at Pensacola FL during Hurricane Camille in August 1969. Here, the observed water level is

E-mail address: r.j.sobey@imperial.ac.uk.

0378-3839/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.coastaleng.2004.09.003



dominated by the storm. Fig. 1b shows the observed and predicted tides at Charleston SC during Hurricane Hugo in September 1989. In this record, the contributions are comparable, and interaction becomes an important concern. Note also in this trace that observed water levels later in the storm fall below the predicted astronomical tide. At sites where the tide range is large, the observed water level is dominated by the astronomical tide.

The tide has a deterministic origin in the mass attraction of the Moon and the Sun on surface waters, and the application of the statistics of extremes to tidal observations alone is not appropriate. The deterministic lower and upper bounds are the Lowest Astronomical Tide (LAT) and Highest Astronomical Tide (HAT), respectively. However, the occurrence of storms has a statistical flavour, so that measured storm tides are suitable for extreme value analysis.

3. Extreme value series

The measurement of tides has been a routine practice at many ports for almost a century. Measured tides include the storm contribution, so that long-term records are available at numerous sites.

While long-duration records are the expectation of classical extreme value analysis, they are an unfamiliar bonus in coastal engineering. In the wave climate context, measured records of perhaps a decade or so are common, and analysis compromise must be introduced. In the popular Partial Duration Series or Peak over Threshold approach, the natural climate-year time base is sacrificed to increase the number of extreme value observations available. The effective time base is dependent of the threshold adopted. An alternative, the Triple Annual Maximum Series approach (Sobey and Orloff, 1995), retains the natural climate-year time base and enhances the data fit by acknowledging annual near-extremes. While the contribution of annual near-extremes is significant for a short duration observational record of order a decade, their value rapidly diminishes with the length of the observational record. The compromises and uncertainties required to accommodate short-duration data series need not be considered here.

It is assumed for this paper that a long-duration observational record of about 100 years is available, and that monthly-extremes of high and low water can be extracted from that data set. Further selecting the highest and lowest monthly sustained water levels defines separate Annual Maximum Series and Annual Minimum Series. The abbreviation AMS will be adopted where both series are implied.

Classical extreme value analysis is applicable, but there are some special concerns. The context requires consideration of

- Extremes of both maximum sustained water level and minimum sustained water level,
- A physical lower bound on maximum sustained water that is not zero,
- A physical upper bound on minimum sustained water that may not be zero,
- Local trends in sea level rise or fall.

These issues are illustrated by the raw AMS series for San Francisco (Fig. 2), with observations beginning in 1898. The periodic nature of the tide will guarantee that each annual maximum observation must exceed at least MHHW (Mean Higher High Water). Similarly, each annual minimum observation must be at least below MLLW (Mean Lower Low Water). Physical bounds of this nature must be assigned in some extreme value models, including Extreme Value II and III and Log Normal.

Note also that numerous annual maximum observations do not exceed Highest Astronomical Tide (HAT) and numerous annual minimum observations do not fall below Lowest Astronomical Tide (LAT). This suggests that the astronomical forcing at the period $2\pi/\Omega_5$=18.61 years (often nominally 19 years), identified as the relative rotation of the lunar and solar orbits or "regression of the lunar nodes" (Doodson, 1921; Schureman, 1940; Doodson and Warburg, 1941), has a significant contribution to the extreme observations. This periodicity is somewhat masked by the definition of contributions to the AMS series from any month in the water year. The 18.61-year periodicity is theoretically accommodated in tidal predictions. Over time scales of hours, the time scale of storm tides, it appears as a slowly varying datum shift. A similar Ω_5 contribution to apparent sea level rise was identified by Töppe (1992) for the North Sea coast of Germany, and by Flick et al. (1999) for numerous US locations.

A definite and seemingly linear trend is also apparent in the Fig. 2 observations, an observation that is consistent with historical datum trends (NOAA/NOS CO-OPS, 2003, Sea Level Trends for 9414290 San Francisco, California) at San Francisco.

Both the $2\pi/\Omega_5$ periodicity and the apparent linear trend are not random aspects of the observations. They suggest that components of the form

$$\hat{\eta}_{\mathrm{High}}(t) = mt + c_{\mathrm{High}} + a_5\cos(\Omega_5 t + \phi_5)$$

$$\hat{\eta}_{\mathrm{Low}}(t) = mt + c_{\mathrm{Low}} + a_5\cos(\Omega_5 t + \phi_5) \quad (1)$$

respectively should be separated from the observational record. t is time, m is the slope of the linear trend, c_{High} and c_{Low} are datum corrections, and a_5 is the amplitude and ϕ_5 the phase of the 18.61 year periodicity. As the monthly maximum observations and the monthly minimum observations are extracted from the same observational record of the local sustained water level, it is expected that m, a_5, and

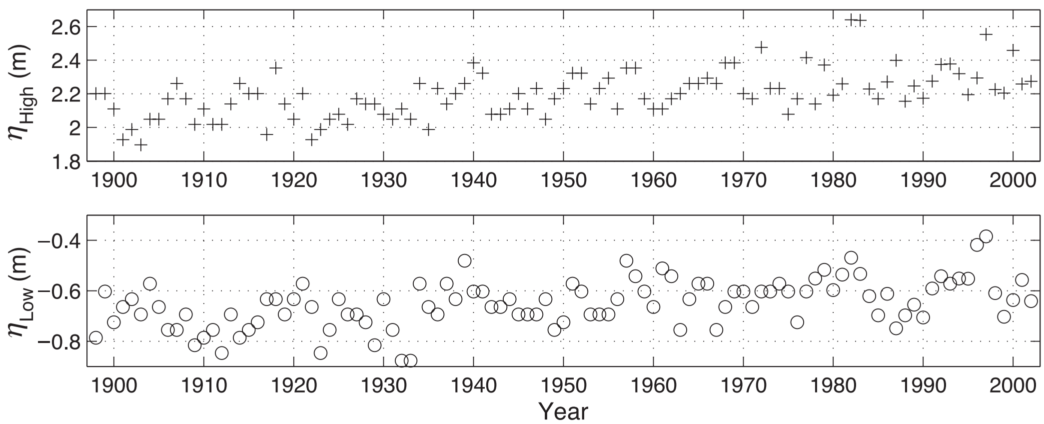


Fig. 2. Raw annual maximum data series for high and low water levels at San Francisco. Data from NOAA/NOS CO-OPS (2003, Monthly-extremes for 9414290 San Francisco, California). Elevations to MLLW datum. MHHW is +1.78 m, HAT is +2.2 m, LAT is −0.6 m.

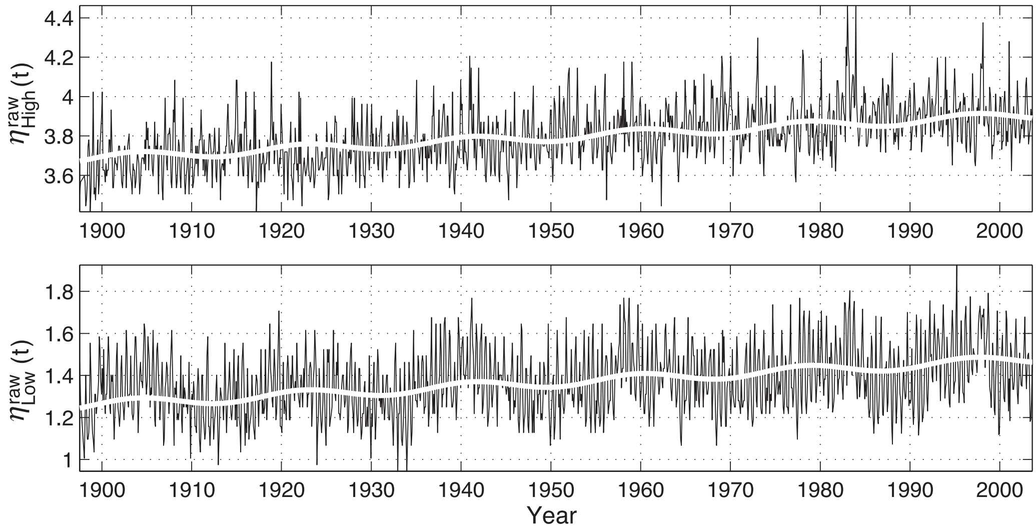


Fig. 3. Monthly-extremes data and trend (thick line) for San Francisco. m=0.00208 m/year, a_5=0.0224 m, and ϕ_5=3.01 radians for zero time at 1 January 1950.

ϕ_5 will be consistent among the observations. The datum for time is arbitrary, but assigning zero time as recent (say 1 January 1950, or 1 January 2000) is convenient.

These trends must be separated from the observational record. A rational approach is a least-squares fit of Eq. (1) to the raw observations, the observational data series of monthly-extremes. Such an algorithm is consistent with the hypothesized separation of the raw observations into a deterministic trend, Eq. (1), and a random residual. Least-squares is also the numerical algorithm implicit in tidal harmonic analysis, the extraction of tidal constituent amplitudes and phases from an observational record, prior to tidal prediction. A suitable least-squares objective function is

$$O\left(m, c_{\mathrm{High}}, c_{\mathrm{Low}}, a_5, \phi_5\right) = \sum_{j}^{12N} \left[\left(\eta^{\mathrm{raw}}_{\mathrm{High}}\left(t_j\right) - \hat{\eta}_{\mathrm{High}}\left(t_j\right) \right)^2 + \left(\eta^{\mathrm{raw}}_{\mathrm{Low}}\left(t_j\right) - \hat{\eta}_{\mathrm{Low}}\left(t_j\right) \right)^2 \right] \quad (2)$$

in which $\eta^{\mathrm{raw}}_{\mathrm{High}}(t_j)$ are the monthly maximum observations, $\eta^{\mathrm{raw}}_{\mathrm{Low}}(t_j)$ are the monthly minimum observa-

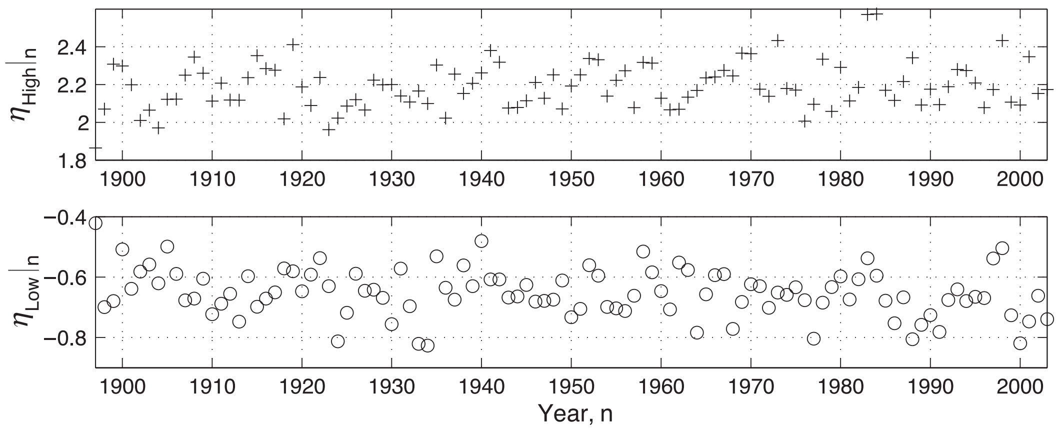


Fig. 4. Net annual-extreme series for San Francisco.

tions, and t_j are the times of the observations. There are 5 unknowns and 12 N monthly observational pairs, over N years. It is appropriate here to use the monthly-extreme observations. Annual-extreme observations have rather erratic time resolution over the 18.61-year period, being selected potentially from any month over each climate year.

The Eq. (1) fit to the monthly-extremes for San Francisco is shown in Fig. 3. m, a_5 and ϕ_5 are listed in the caption. The computed datum trend m matches the Mean Sea Level trend of 0.00213 m/year for the period 1906–1999 reported by NOAA (NOAA/NOS CO-OPS, 2003, Sea Level Trends for 9414290 San Francisco, California)[1] Subtracting this trend, without the datum corrections, from the monthly-extremes data series defines net monthly-extremes data series. Extracting the extreme values over each climate year defines the AMS series, $\eta_{\mathrm{High}}|_n$ and $\eta_{\mathrm{Low}}|_n$, respectively, where n identifies the climate year. For San Francisco, the conditioned AMS series are shown in Fig. 4. These data series are immediately suitable for extreme value analysis.

4. Extreme value analysis for maximum values

Extreme value analysis must choose the probability model for interpolation and extrapolation. The literature for annual maximum events is extensive. Suitable distributions have probability density functions (PDF's) that are asymptotic to a tail at the high end. Mostly the literature designation is "largest value", to distinguish from distributions that are asymptotic to a tail at the low end, which are designated "smallest value". The exception is the Extreme Value III distribution, where the "largest (smallest) value" refers to a required upper (lower) bound.

The common two-parameter candidate distributions are listed in Table 1 (Benjamin and Cornell, 1970; Ang and Tang, 1984), for the random variable y. Also listed in the table are the normalized variable Y, the distribution parameters, the range of the random variable y, and distribution mean μ and variance σ^2 for the normalized variable.

The Extreme Value II and III and Log Normal models explicitly require that all possible values of y be positive. For some geophysical variables, including wave height, this is physically and numerically appropriate. For extreme water levels, it is neither physically nor numerically appropriate. It is not physically consistent with tidal theory and observations, where no entry in the annual maximum series will fall below at least MHHW. In addition, typical station datums are at or near MLLW, so that the physical lower bound on annual maximum observations will always be finite and positive. It would be appropriate to adopt an analysis datum shift, from y_0=0 to y_0=MHHW, so that the observations are consistent with the mathematical expectations of the probability model. Note that the Extreme Value I model will assign finite probabilities to extreme high water levels below MHHW.

The datum level, MHHW, is a physical parameter, and quite distinct from the statistical location parameters u and v identified in column 4 of Table 1. There is potential confusion in the designation "Mean" Higher High Water. It perhaps suggests that MHHW is a statistical parameter, so that the Extreme Value II, Extreme Value III and Log Normal distributions in Table 1 should be regarded as three and not two parameter distributions.

There are five recognized periodicities in the Earth–Moon–Sun system, with periods of 24.8 h, 27.3 days, 365.2 days, 8.85 years and 18.6 years. MHHW is the average of the Higher High Waters each lunar day over the longest 18.6-year period. The period of averaging is fixed by definition. It is an average, but it is an average of a deterministic variable, the astronomical tide prediction. MHHW is a deterministic and physical parameter.

MHHW values published in Tide Tables are defined by national tidal authorities. Even if observed, rather than predicted, Higher High Waters were used in the averaging process, the contribution of storm tides to the observed HHWs would be compensatory, some positive and some negative. More significantly, the average will be overwhelmingly dominated by the

[1] The NOAA presentation identifies a datum shift down at the time of the massive 1906 San Francisco earthquake, which prompted their exclusion of data prior to 1906. Flick et al. (1999) used data for the period 1855–1999, before and after the 1906 datum shift. Their Mean Sea Level trend estimate, 0.00475 ft/year (0.00156 m/year), is accordingly biased low.

Table 1
Two-parameter extreme value distributions for maximum values

Probability model	Normalized variable	CDF PDF range	Parameters mean variance
Extreme Value I (Gumblel)	$Y=y-v/u$	$F(Y)=\exp[-\exp(-Y)]$	u, v
		$f(Y)=\exp[-Y-\exp(-Y)]$	$\mu=0.5772$
		$-\infty<y<+\infty$	$\sigma^2=\pi^2/6$
Extreme Value II	$Y=y/u$	$F(Y)=\exp(-Y^{-\alpha})$	u, α
		$f(Y)=\alpha Y^{-\alpha-1}\exp(-Y^{-\alpha})$	$\mu=\Gamma(1-1/\alpha)$
		$y\geq 0$	$E(Y^2)=\Gamma(1-2/\alpha)$
Extreme Value III (Weibull)	$Y=y/u$	$F(Y)=1-\exp(-Y^{-\alpha})$	u, α
		$f(Y)=\alpha Y^{\alpha-1}\exp(-Y^{\alpha})$	$\mu=\Gamma(1+1/\alpha)$
		$y\geq 0$	$E(Y^2)=\Gamma(1+2/\alpha)$
Log Normal	$Y=\log(y)-\alpha/2^{1/2}\beta$	$F(Y)=1/2[1+\mathrm{erf}(Y)]$	α, β
		$f(Y)=\pi^{-1/2}\exp(-Y^2)$	$\mu=0$
		$y>0$	$\sigma^2=0.5$

$F(y)=F(Y)$, $f(y)=f(Y)(dY/dy)$; $\sigma^2=E(Y^2)-\mu^2$.

astronomical contributions, each lunar day for the more than 6500 lunar days over the 18.6-year averaging period.

There is a physical synergy also in the adoption of MHHW as a deterministic lower bound (the mean HHW over the 18.6-year period) and in the re-introduction on the 18.6-year periodicity as a deterministic trend.

Given the selection of a probability model from Table 1, the analysis steps are accordingly

(1) Datum shift the raw AMS series as required by the probability model.
The Extreme Value II, Extreme Value III and Log Normal definition for y to be positive is accommodated by an analysis datum shift to a data lower bound at $y=y_0$. For annual maximum series of water surface elevations, y_0=MHHW would be physically appropriate.
The Extreme Value I model requires no datum shift, so that $y_0=0$.
The net AMS series for extreme value analysis would be $y-y_0$.
The balance of the extreme value analysis follows the practice for extreme wave heights (e.g. Sobey and Orloff, 1995).

(2) Estimate the distribution parameters by maximizing the sample likelihood (strictly log-likelihood) function

$$L_1(p_1,p_2)=\frac{1}{N}\sum_{n=1}^{N}\ln[f_Y(y_n;p_1,p_2)] \tag{3}$$

in which p_1 and p_2 are the distribution parameters (e.g. u and v for Extreme Value I).
Maximizing L_1 is a problem in nonlinear optimization, for which appropriate algorithms are available in scientific subroutine libraries. Algorithms require specification of the objective function L_1 and reasonable initial estimates of the distribution parameters; gradients are estimated internally by finite differences. The initial estimates are most conveniently provided by the method of moments.
Of the established procedures for estimation of distribution parameters from observational data, only this method of maximum likelihood is based on the PDF of the candidate probability model. Other methods are based on the CDF, the integrated form of the candidate probability model. The CDF, because of its integrated definition, is significantly less sensitive to the observational data.

(3) Estimate the confidence limits, L_P and U_P, on the probability model, such that

$$Pr(L_P<y_P<U_P)=1-\alpha \tag{4}$$

where there is a $(1-\alpha)$ probability that the interval L_P to U_P contains the point estimate $\bar{y}_P=F_Y^{-1}(P)$ for $\bar{y}$ at cumulative probability level P. ($F_Y^{-1}(P)$ denotes the inverse function, such that $F_Y(\bar{y}_P)=P$.) A 95% confidence interval, $1-\alpha=0.95$, is the common practice.
For the long-duration data series in the present context, pragmatic estimates are provided by the

asymptotic normal approximation (Mood et al., 1974) with mean $\bar{y}_P$ and standard deviation

$$S_{F_Y}(\bar{y}_P) = S_P = \left[\frac{P(1-P)}{N f_Y^2(\bar{y}_P)}\right]^{1/2} \quad (5)$$

From the Normal distribution, the lower and upper 95% confidence limits at cumulative probability P will be $L_P = \bar{y}_P - 1.96S_P$, and $U_P = \bar{y}_P + 1.96S_P$.

(4) Estimate event magnitudes for an average recurrence interval of T_r years.

$$T_r = \frac{\Delta t_Y}{1 - F_Y(y)} \quad (6)$$

where Δt_Y is the uniform time interval from which the samples y_n of the random variable Y are selected; for AMS series, Δt_Y is 1 year. Point estimates of $F_Y(y)$, and its confidence limits, are completely defined in steps 2 and 3, respectively.

(5) Reintroduce the analysis datum shift y_0.
The event magnitude corresponding to an assigned T_r and an observational Δt_Y will be

$$\bar{y}|_{T_r} = y_0 + F_Y^{-1}(1 - \Delta t_Y / T_r) \quad (7)$$

For an AMS series and an average recurrence interval of 100 years, $\bar{y}|_{100} = y_0 + F_Y^{-1}(0.99)$. Similarly, the confidence bands become

$$L_P = \bar{y}_P - 1.96S_P, \text{ and } U_P = \bar{y}_P + 1.96S_P. \quad (8)$$

in which $\bar{y}_P = y_0 + F_Y^{-1}(P)$.

In ideal application, all four of the two-parameter distributions in Table 1 would be attempted. Maximum likelihood estimates of the distribution param-

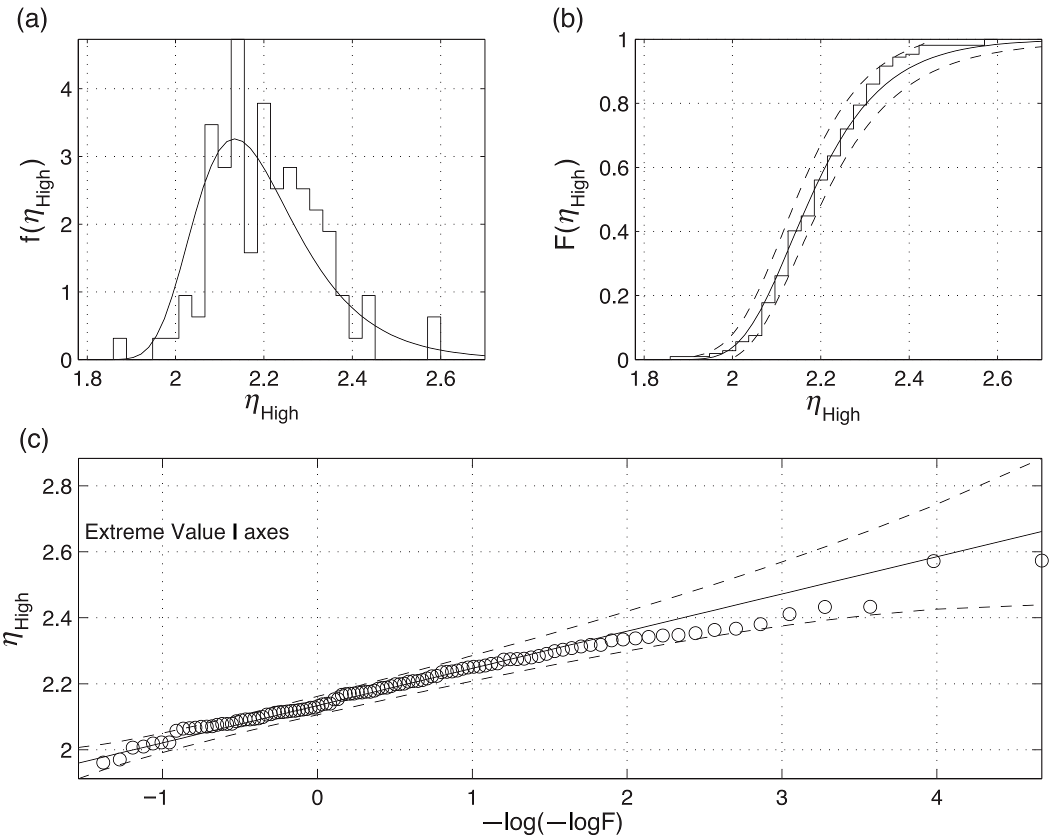


Fig. 5. Extreme maximum analysis—Extreme Value I model. u=0.113 m, v=2.13 m, L_1=0.666, $\bar{\eta}_{High}|_{100}$=2.65 m, S_{100}=0.11 m.

eters, together with the 95% confidence bands would be established. The most suitable distribution would be adopted, with guidance from

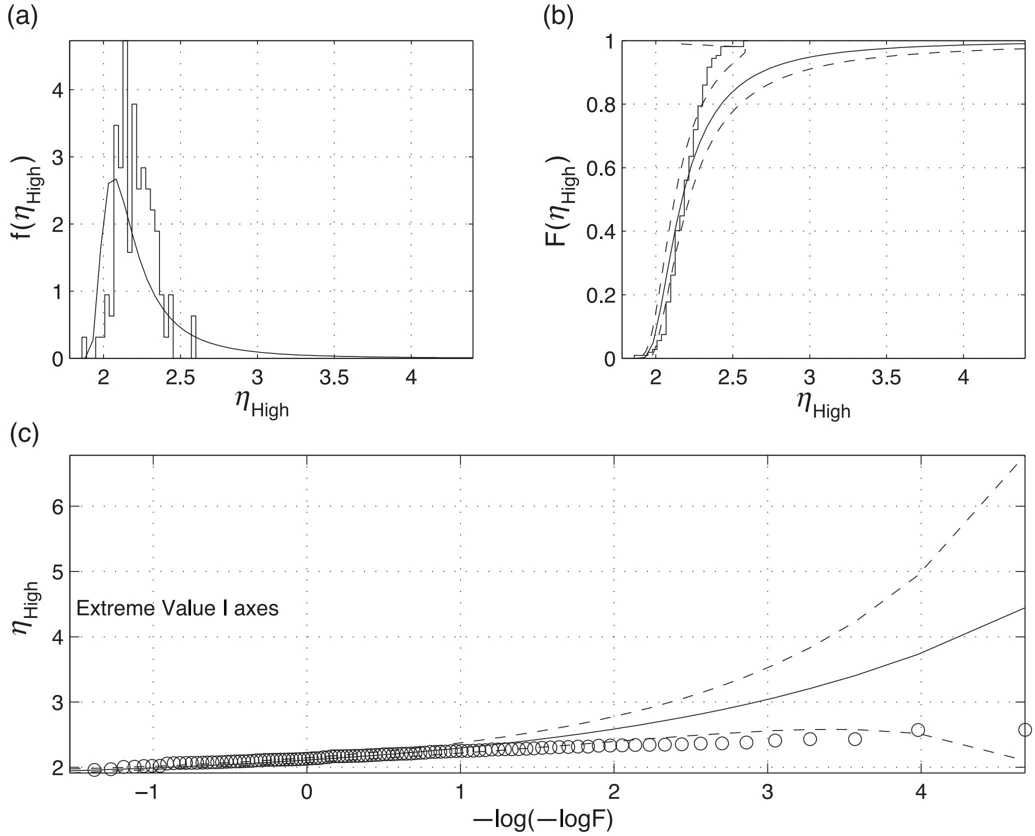
(i) the visual goodness-of-fit of the PDF over the entire range of the observational data,
(ii) the magnitude of likelihood function L_1,
(iii) the width of the confidence bands,
(iv) the significance of any data outliers, and
(v) any overriding regulatory requirements.

In each application, model selection is a matter of engineering judgement.

Figs. 5–8 show extreme maximum analysis results for each of the four models listed in Table 1. The respective distribution parameters, the sample likelihood function L_1, the event magnitude $\eta_{\mathrm{High}}|_{100}$ at an average recurrence interval of 100 years and the estimated standard deviation S_{100} on this event, following Eq. (5), are listed in the captions.

These four figures collectively illustrate the challenges of model selection, but having four options provides the opportunity for some rational choice. The data fit to the CDF (part b) appears almost equally satisfactory for all four models. In contrast, the PDF context (part a) highlights both the data variability and the differing shapes of the distributions. The enhanced sensitivity to the PDF emphasizes the value of the parameter fitting by the method of maximum likelihood. Popular alternative methods of parameter fitting, such as probability plotting and the method of moments, both adopt the CDF description of the probability model.

Nevertheless, there is value in the probability-plotting style of CDF presentation (part c), especially



its visual focus on the data fits throughout the active range. For comparative purposes, all presentations have adopted the Extreme Value I scaling. Confidence limits are best displayed in this format.

The Extreme Value I result, Fig. 5, is promising, but the model appears to predict high for the more extreme events. Extreme Value II does not appear a viable option. Much of the data at the high end is outside the 95% confidence limits, and the confidence limits are especially wide at this extreme. The 100-year event magnitude is also very high in comparison to the other models, and well outside the expectations from the observational data. Extreme Value III is another promising result, but the model appears to predict low for the more extreme events. The two highest observed events are beyond the 95% confidence limits. The Log Normal result appears the most satisfactory for this data set. The confidence band is relatively narrow and all the observational data, especially at the high end, falls within the confidence limits.

It is important to note however that these are interpretations appropriate for the San Francisco observational data set. Model performance and selection will be data set and hence site dependent.

5. Trend adjustment

The final step in the prediction of extreme maximum water level is the re-introduction of the deterministic trend, Eq. (1). The predicted extreme maximum water level will be

$$\bar{\bar{\eta}}_{\mathrm{High}}(t;T_{\mathrm{r}}) = \bar{\eta}_{\mathrm{High}}|_{T_{\mathrm{r}}} + mt + a_5\cos(\Omega_5 t + \phi_5) \qquad (9)$$

depending on the average recurrence level T_{r} and on real time t.

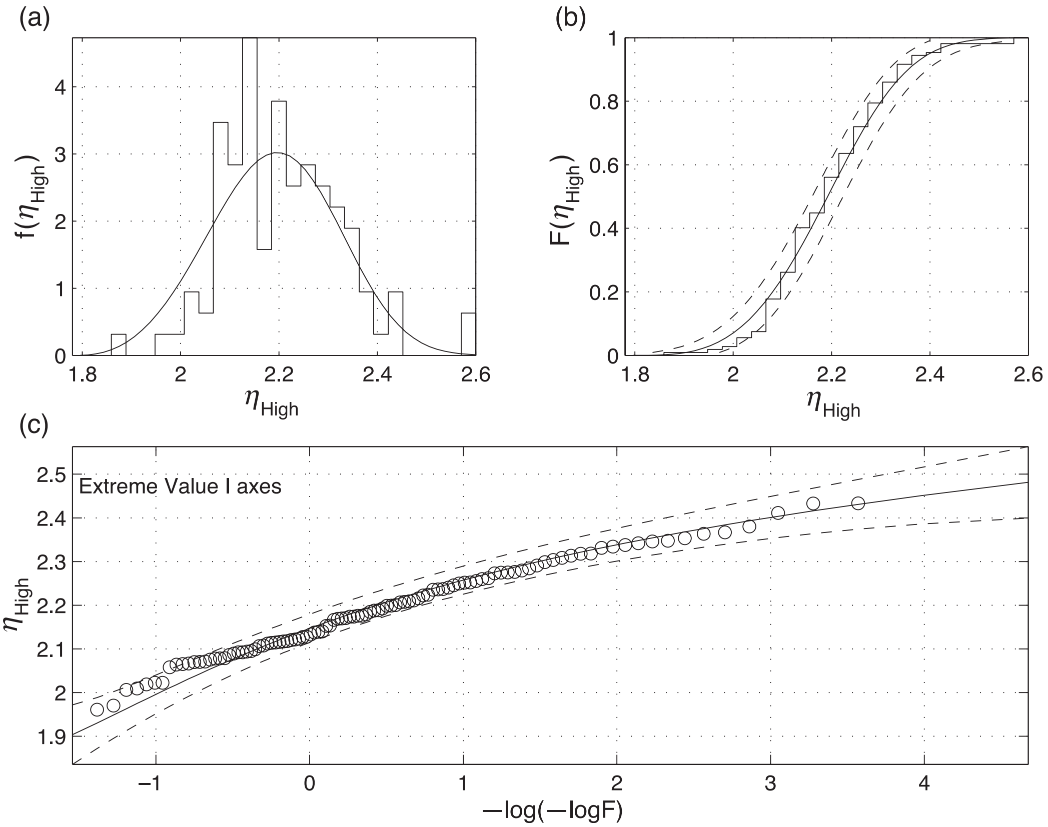
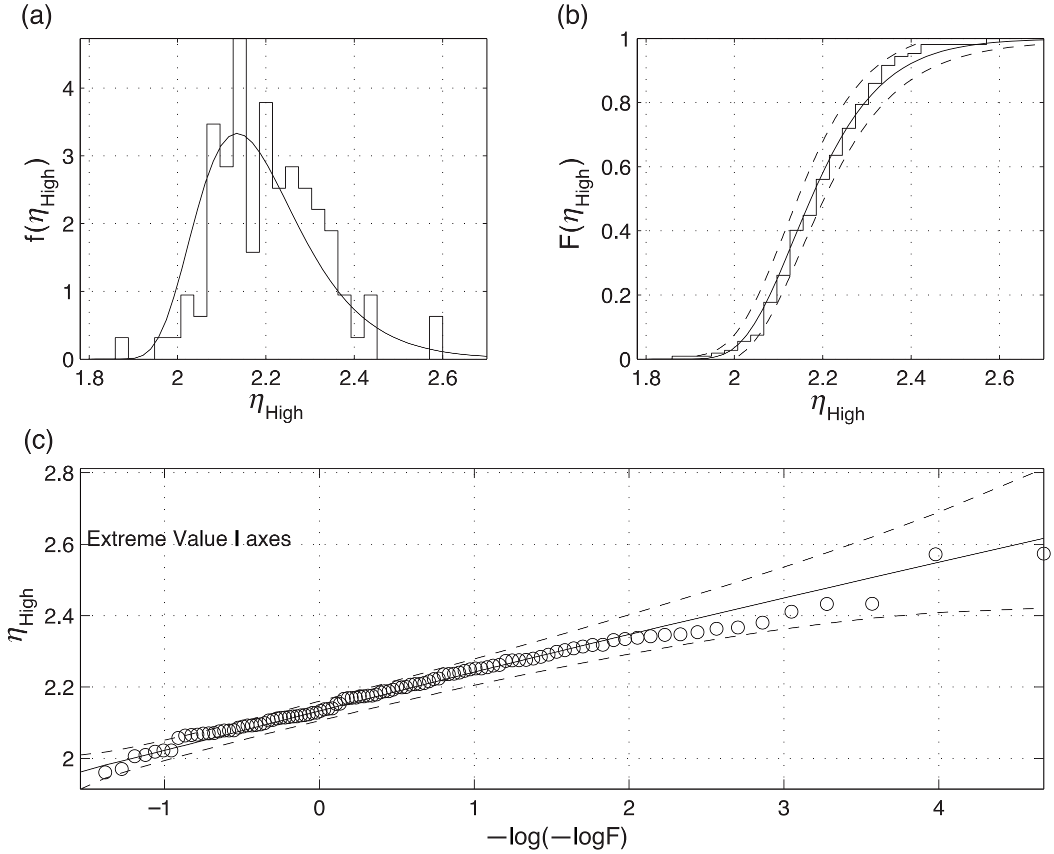


Fig. 7. Extreme maximum analysis—Extreme Value III model. u=3.59 m, α=0.456, L_1=0.678, $\bar{\eta}_{\mathrm{High}}|_{100}$=2.48 m, S_{100}=0.04 m.



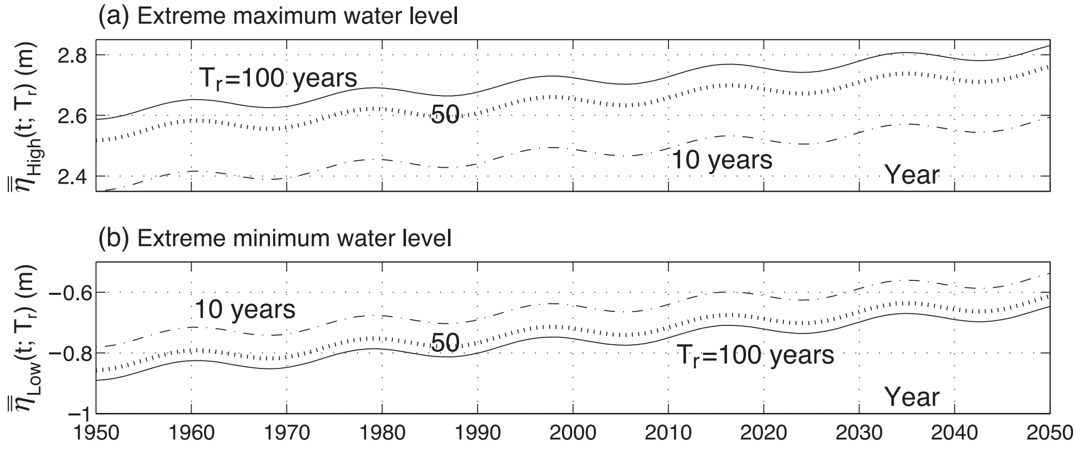


Fig. 9. Trend-adjusted, Log Normal prediction of extreme maximum water level at San Francisco with average recurrence intervals of 10, 50 and 100 years.

Table 2
Two-parameter extreme value distributions for minimum values

Probability model	Normalized variable	CDF PDF range	Parameters mean variance
Extreme Value I	$Y=y-v/u$	$F(Y)=1-\exp[-\exp(-Y)]$	u, v
		$f(Y)=\exp[Y-\exp(-Y)]$	$\mu=0.5772$
		$-\infty<y<+\infty$	$\sigma^2=\pi^2/6$
Extreme Value II	$Y=y/u$	$F(Y)=1-\exp(-Y^{-\alpha})$	u, α
		$f(Y)=\alpha Y^{-\alpha-1}\exp(-Y^{-\alpha})$	$\mu=-\Gamma(1-1/\alpha)$
		$y\leq 0$	$E(Y^2)=\Gamma(1-2/\alpha)$
Extreme Value III	$Y=y/u$	$F(Y)=\exp(-(-Y)^{\alpha})$	u, α
		$f(Y)=-\alpha Y^{\alpha-1}\exp(-(-Y)^{\alpha})$	$\mu=-\Gamma(1+1/\alpha)$
		$y\leq 0$	$E(Y^2)=\Gamma(1+2/\alpha)$
Log Normal	$Y=-\log(-y)-\alpha/2^{1/2}\beta$	$F(Y)=1/2[1+\mathrm{erf}(Y)]$	α, β
		$f(Y)=\pi^{-1/2}\exp(-Y^2)$	$\mu=0$
		$y<0$	$\sigma^2=0.5$

$F(y)=F(Y)$, $f(y)=f(Y)(dY/dy)$; $\sigma^2=E(Y^2)-\mu^2$.

Adopting the Log Normal model (Fig. 8) for extreme maximum water levels at San Francisco, the evolution of the trend-adjusted 10-, 50- and 100-year events are shown in Fig. 9a.

6. Extreme value analysis for minimum values

Adopted probability models must be event-suitable. For annual minimum events, appropriate dis-

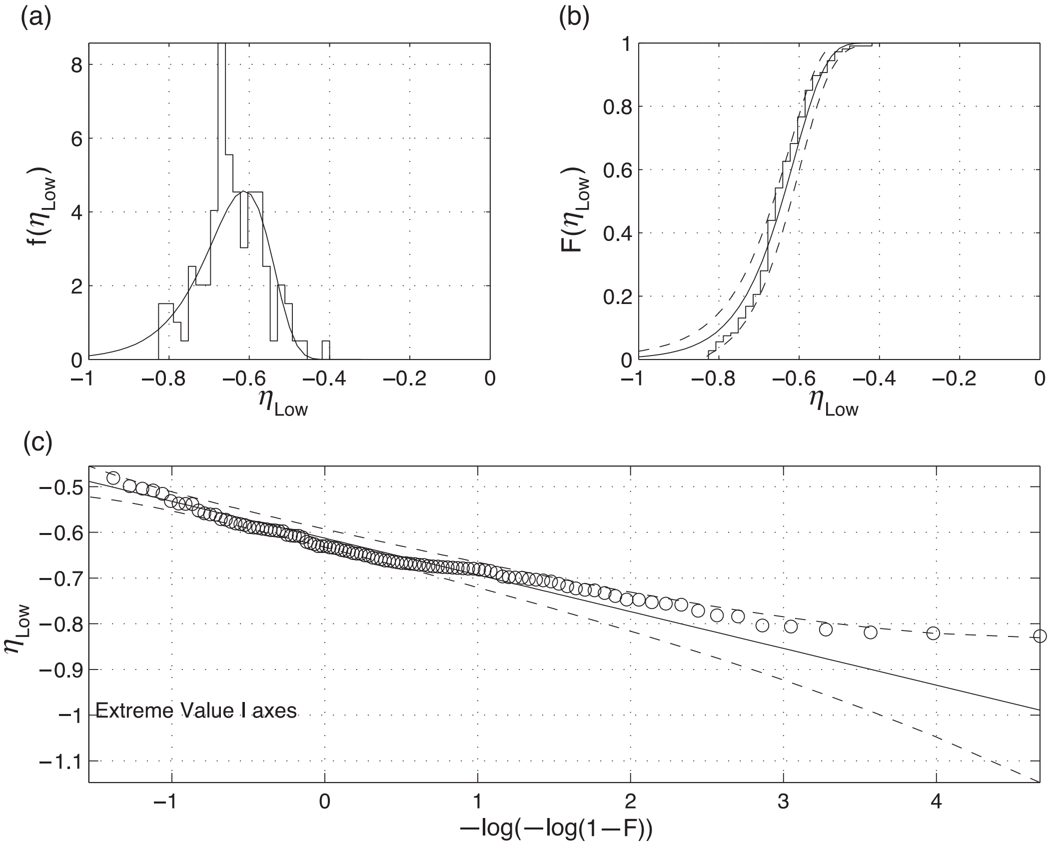


Fig. 10. Extreme minimum analysis—Extreme Value I model. u=0.080 m, v=−0.613 m, L_1=1.03, $\bar{\eta}_{Low}|_{100}$=−0.98 m, S_{100}=0.08 m.

tributions must have a PDF that is asymptotic to a tail at the low end. Each of the common two-parameter candidate distributions for annual maximum series (Table 1) has a variation with a tail at the low end. These models are listed in Table 2 (Benjamin and Cornell, 1970; Ang and Tang, 1984) for the random variable y. Of these, the Log Normal model was not found in the literature, though its definition for minimum values is straightforward.

With the exception of Extreme Value I, the models in Table 2 for minimum values explicitly require that y be negative. For extreme low water levels, an analysis datum shift to y_0=MLLW would be appropriate. This is often implicit because of the almost universal adoption of a station datum identified as a Low Water datum. This datum is commonly at or near to MLLW. The local datum is an issue that must be pursued and confirmed. For the San Francisco data set, y_0=0. Note that the Extreme Value I model will assign finite probabilities to extreme low water levels above MLLW.

An additional departure from the analysis methodology adopted in Section 4 is the definition of average recurrence interval. The focus now is on the low end of the probability distributions, on cumulative (non-exceedance) rather than exceedance probability. The average recurrence interval becomes

$$T_{\mathrm{r}} = \frac{\Delta t_Y}{F_Y(y)}. \tag{10}$$

The event magnitude corresponding to an assigned T_{r} and an observational Δt_Y will be

$$\bar{y}|_{T_{\mathrm{r}}} = y_0 + F_Y^{-1}(\Delta t_Y / T_{\mathrm{r}}). \tag{11}$$

For an AMS series and an average recurrence interval of 100 years, $\bar{y}|_{100}=y_0+F_Y^{-1}(0.01)$.

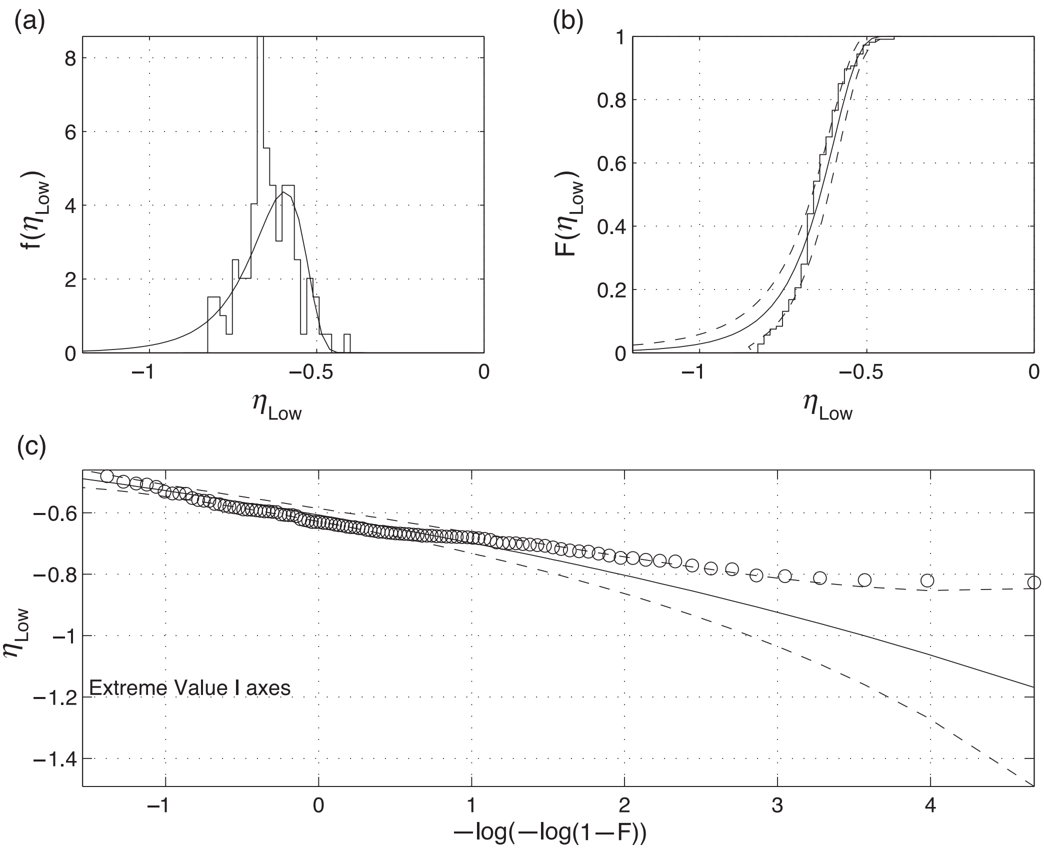


Fig. 11. Extreme minimum analysis—Extreme Value II model. u=7.14 m, α=0.607, L_1=0.941, $\bar{\eta}_{\mathrm{Low}}|_{100}$=−1.16 m, S_{100}=0.16 m.

A final presentation issue is probability plotting. For Extreme Value I scaling, the F axis transformation is $-\log(-\log F)$ for the extreme maximum analysis, but $-\log(-\log(1-F))$ for the extreme minimum analysis.

The five analysis steps outlined in Section 4 for the extreme maximum analysis are otherwise unchanged.

Figs. 10–13 show extreme minimum analysis results for each of the four models listed in Table 2. The respective distribution parameters, the sample likelihood function L_1, the event magnitude $\eta_{\mathrm{Low}}|_{100}$ at an average recurrence interval of 100 years and the estimated standard deviation S_{100} on this event, following Eq. (5), are listed in the captions.

Model selection again follows the guidelines outlined in Section 4. The Extreme Value I result, Fig. 5, is promising, but the model appears to predict high for the more extreme events. Extreme Value II does not appear a viable option. Much of the data at the low end is outside the 95% confidence limits, and the confidence limits are relatively wide at this extreme. The 100-year event magnitude is also quite low in comparison to the other models, and beyond the expectations from the observational data. Extreme Value III is another promising result, but the model appears to predict a little high for the more extreme events. The Log Normal result appears the most satisfactory for this data set. All the observational data, especially at the low end, falls within the confidence limits.

As for the extreme maximum analysis, these are interpretations appropriate for the San Francisco observational data set. Model performance and selection will be data set and hence site dependent. In addition, the model choice for the extreme minimum analysis need not correspond with the choice for the extreme maximum analysis.

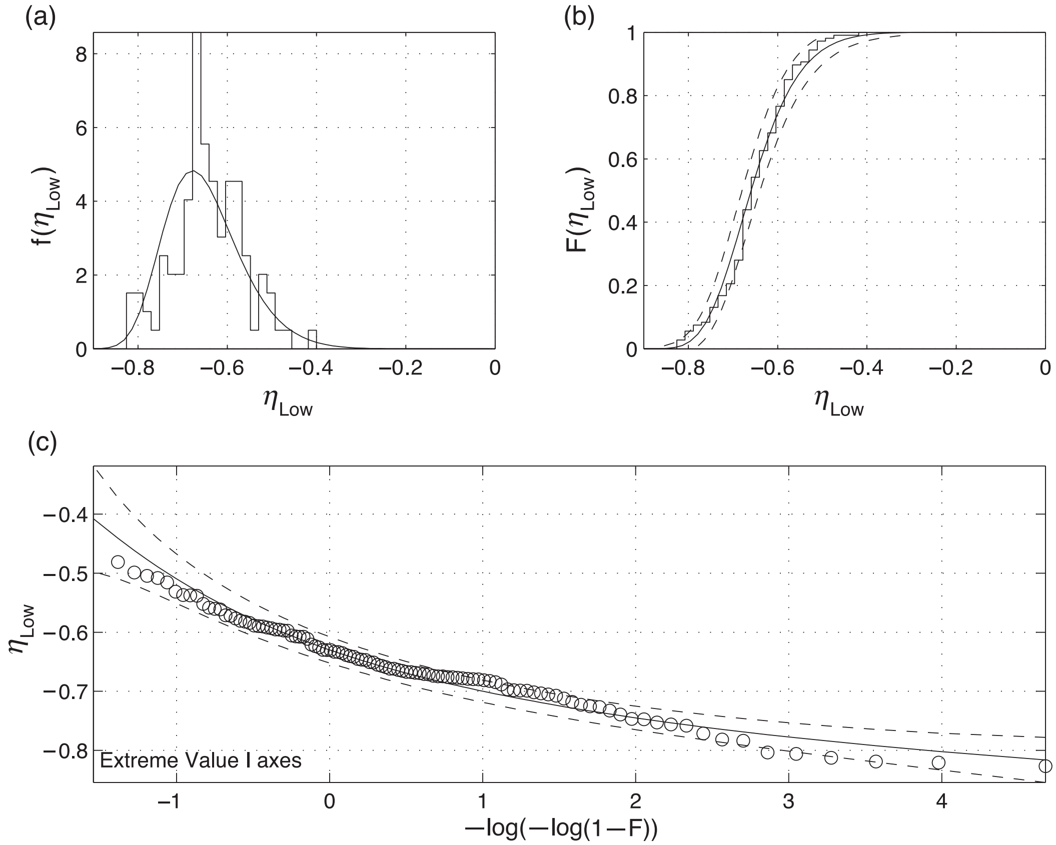


Fig. 12. Extreme minimum analysis—Extreme Value III model. u=8.97 m, α=0.687, L_1=1.09, $\bar{\eta}_{\mathrm{Low}|100}$=−0.81 m, S_{100}=0.02 m.

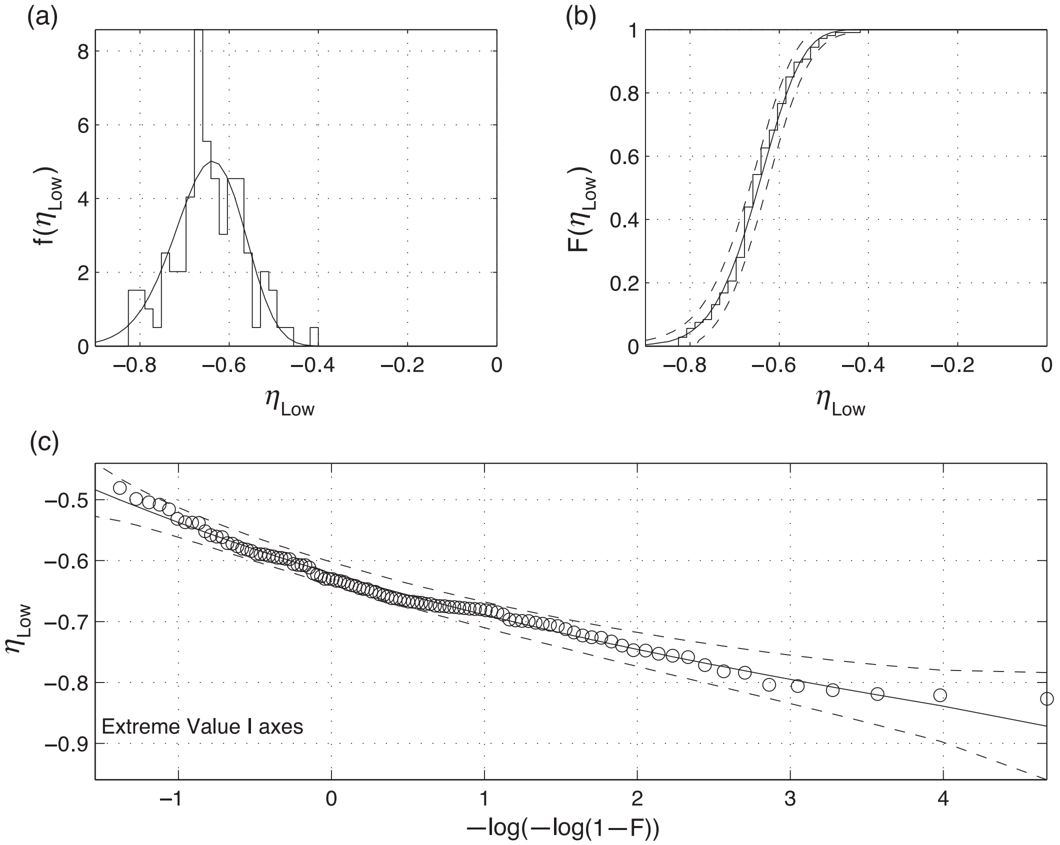


Fig. 13. Extreme minimum analysis—Log Normal model. α=−0.435, β=0.124, L_1=1.11, $\bar{\eta}_{\mathrm{Low}}|_{100}$=−0.87 m, S_{100}=0.04 m.

The final extreme minimum analysis step is the deterministic trend adjustment. The predicted extreme maximum water level will be

$$\bar{\bar{\eta}}_{\mathrm{Low}}(t;T_{\mathrm{r}}) = \bar{\eta}_{\mathrm{Low}}|_{T_{\mathrm{r}}} + mt + a_5\cos(\Omega_5 t + \phi_5) \qquad (12)$$

depending again on the average recurrence level T_{r} and on real time t. Adopting the Log Normal model (Fig. 8) for extreme minimum water levels at San Francisco, the evolution of the trend-adjusted 10-, 50- and 100-year events are shown in Fig. 9b.

7. Conclusions

The dynamics of storm tides can lead to sustained increases in tide elevation and, at different stages of the storm tide evolution, also to sustained decreases in tide elevation. Abnormally high tides are experienced when sustained storm-forced increases correspond with predicted high waters in the local astronomical tide. Similarly, abnormally low tides are experienced when sustained storm-forced decreases correspond with predicted low waters in the local astronomical tide. Long-duration historical observations of tidal elevations are available at many coastal locations, and these observations will include observations of abnormally high and low tides.

The long-duration observational record includes potentially significant contributions from Mean Sea level rise and from the 18.61-year periodicity in the tidal forcing. For the San Francisco observational record, both are present and both are of sufficient magnitude to demand accommodation in extreme value analyses. These are not random influences and should be removed from the record prior to extreme value analysis.

The raw observational record in a form suitable for the prediction of extremes of both high and low water is a record of monthly maximum and minimum water surface elevations. Data conditioning requires the identification of a deterministic trend, Eq. (1), its definition by least-squares from the observational record, and its separation from the record to establish a net data record suitable for extreme value analysis. This net data record is summarized as Annual Maximum and Annual Minimum Series, by selection of the largest (smallest) observation in consecutive climate years.

Table 1 defines the four common two-parameter distributions routinely adopted for extreme maximum analysis. Given a choice of extreme value distribution, a five-step methodology has been outlined for a rational prediction of the extreme maximum event. These steps progressively

(1) datum shift the raw AMS series to y_0 as required by the probability model,
(2) estimate the distribution parameters by the method of maximum likelihood,
(3) estimate the confidence limits on the probability model,
(4) estimate event magnitudes for an average recurrence interval of T_r years, and
(5) reintroduce the analysis datum shift y_0.

For the Extreme Value II and III and Log Normal models, y_0 is assigned as the MHHW, a level below which observations of annual maximum water level will not fall.

Sample extreme maximum analyses are shown in Figs. 5–8. A complete result, including the correction for the datum trend, is shown in Fig. 9a.

The separate extreme minimum analysis identifies those aspects that are unique to events that are asymptotic to a tail at the low end. Issues include the form of suitable two-parameter distributions (Table 2), an analysis datum shift to y_0=MLLW, and the definition of average recurrence interval (Eq. (10)).

Sample extreme minimum analyses are shown in Figs. 10–13. A complete result, including the correction for the datum trend, is shown in Fig. 9b.

The methodology is illustrated with data at San Francisco, but is generally applicable for any site with long-duration tidal observations. Generalization to three-parameter distributions is also straightforward.

References

Ang, A.H.-S., Tang, W.H., 1984. Probability concepts in engineering planning and design. Decision, Risk, and Reliability, vol. II. Wiley, New York.

Benjamin, J.R., Cornell, C.A., 1970. Probability, Statistics and Decision for Civil Engineers. McGraw-Hill, New York.

Doodson, A.T., 1921. The harmonic development of the tide-generating potential. Proceedings of the Royal Society of London, Series A 100, 305–329.

Doodson, A.T., Warburg, H.D., 1941. Admiralty Manual of Tides. Hydrographic Department. Admiralty (HMSO), London.

Flick, R.E., Murray, J.E., Ewing, L.C., 1999. Trends in U.S. tidal datum statistic and tide range. A data report atlas. Reference Series 99-20. Scripps Institution of Oceanography, University of California, San Diego.

Mood, A.M., Graybill, F.A., Boes, D.C., 1974. Introduction to the Theory of Statistic. McGraw-Hill, New York.

NOAA/NOS CO-OPS, August 2003. Retrieve observed water levels and associated ancillary data. http://co-ops.nos.noaa.gov/data_res.html, [web pages, accessed November 2003].

Schureman, P., 1940. Manual of harmonic analysis and prediction of tides. Special Publication, vol. 98. U.S. Department of Commerce, Coast and Geodetic Survey, Washington, DC.

Sobey, R.J., Orloff, L.S., 1995. Triple annual maximum series in wave climate analysis. Coastal Engineering 26, 135–151.

Töppe, A., 1992. Zur Analyse des Meeressoiegelansteiges aus langjährigen Wesserstandsaufzeichnungen an der deutschen Nordseeküxte. Mitteilungen-Leichtweiss-Institut für Wasserbau der Technischen Universität Braunschweig 120, 1–132.