12. ANALYSIS OF THE MEDITERRANEAN DATA

12.1. Introduction

After having presented some of the most important current theories, different aspects of the data are analysed.

As will be illustrated in the current chapter (Section 12.4), the nonlinearities are weak. Therefore, in the first part of the presentation of the analysis, the wave height and wave crest/trough are analysed, omitting the nonlinear theories explained in Chapter 6. For each variable, the significant and maximum height/crest/trough is calculated and compared to the linear theory expectations and, if necessary, to other theories (Chapters 7, 8 and 9).

Finally, the nonlinearities are analysed with the kurtosis and skewness parameter as well as the spectral estimator of the kurtosis: BFI.

Note that often the variables such wave heights or crests are normalized. It always means that they are divided by the standard deviation.

12.2. Wave height

12.2.1. Significant wave height

The observed and estimated normalized significant wave heights, according to the linear theory \( H_{s,m_0} = 4\sqrt{m_0} \), are compared for all records in Figure 2.1. The calculated slope is from the straight line which fits the data best (in red), using the least square method. In the calculation of such a line, the condition of passing through the origin has been established; otherwise it would be nonsense. The black line represents the perfect agreement.
Figure 12.1 Comparison between observed and estimated significant wave height (Rayleigh theory)

As illustrated in Figure 12.1, according to the fitted straight line, $H_{s,obs} \approx 0.931 H_{s,est}$. The observations are 7% lower than the estimations which agrees with the values found by Longuet-Higgins (1980), Holthuijsen (2007) and Forristall (1978) (see Chapter 5). Apart from this result it is important to note that there is little deviation from the straight line, meaning that the scaling factor does not seem to depend on the magnitude of the significant wave height. That may lead to thinking that the use of a constant scale factor for the Rayleigh distribution is quite acceptable:

$$p(H) = \frac{H}{4f^2 m_0} \exp \left( -\frac{H^2}{8f^2 m_0} \right)$$ (12.1)
Figure 12.2 Comparison between scaled Rayleigh distribution (according to the found discrepancy in the observed significant wave height) and the original one. Up: probability density function; down: exceedance probability.

Although the value of the scaling factor found for this data set (0.931) seems to be a more or less “universal” constant compared to Longuet-Higgins (1980), Holthuijzen (2007) and others, it would be better to try to find such a scaling factor from a theoretical background instead of empirical. Remark that the standard deviation of the ratio $H_{s}/H_{n}$ has a standard deviation of 0.0262. The mean of the same ratio is 0.929, very close to the slope of the fitted straight line.
Therefore, it is interesting to compare the observations with the theories explained in Chapter 8, the first accounts for a certain background noise (Modified Rayleigh I; Longuet-Higgins, 1980) and the second accounts for a certain time lag between crest and trough (Modified Rayleigh II; Naess, 1985). In the Figure 12.3 and 12.4, the estimated value for the significant wave height is slightly modified according to such theories.

Figure 12.3 Observed and estimated significant wave height with the Modified I theory

Figure 12.4 Observed and estimated significant wave height with Modified II theory
In the second theory, the minimum of the autocorrelation function has been calculated but if we calculate the autocorrelation function for half mean period (Eq. (4.21)), the results do not differ significantly; the fitted slope is 1.0448 instead of 1.0394, reducing the computational cost.

Both theories, concerning the significant wave height, are better adjusted than the linear theory but the first one is clearly the best (1.00 vs. 1.04) although the scatter is higher. Note that in both cases the involved parameters (respectively, $\nu$ and $r_{\min}$) have been calculated for each record instead of considering the mean values, which are: $\nu = 0.41$ and $r_{\min} = -0.57$.

The scaling produces an enhancement of the probability of the low-mid range whereas the higher waves are reduced (see Figure 12.2). Other parameters than the significant wave height, like mean wave height, root-mean-square wave height and estimated maximum wave height are reduced by the same scaling factor. Therefore, if the scaled Rayleigh distribution was the solution, the same discrepancy would be found in the other observed parameters. Figure 12.5 shows the discrepancy found for $H_{\text{mean}}$ and $H_{\text{rms}}$ compared to the linear theory. Despite not being exactly the same, the differences are not very large. However, there is a certain tendency in which the more the parameter is related to higher waves, the larger the scaling factor becomes. Actually, Massel (1996) stated that the Rayleigh distribution overpredicts the higher waves in a record and the error increases toward the low-probability tail of the distribution. Such a tendency is confirmed in Section 12.2.2, in which it is found that the scaling factor associated to the maximum wave height is even higher.

![Figure 12.5 Discrepancy between Rayleigh theory and observations of: significant wave height, mean wave height, root-mean-square wave height.](image)

Therefore, it becomes appropriate to compare these observed parameters with the crest-to-trough distribution (see Chapter 7) which does not simply rescale the Rayleigh distribution. The discrepancy for the significant wave height is illustrated in Figure 12.6. Although Tayfun (1990)
suggested using the mean value over all records for the autocorrelation as a simplification, here, for each record, the autocorrelation value has been calculated (from the spectrum) and used. The mean value is \( r = 0.62 \), very similar to the one obtained by Tayfun (1990): 0.65; and the median is also higher: 0.64, implying a certain asymmetry of distribution of \( r \) (see Figure 12.7).

![Figure 12.6 Observed and estimated significant wave height with crest-to-trough theory](image1)

![Figure 12.7 Histogram of the autocorrelation factor \( r \) of the crest-to-trough theory](image2)

In conclusion, the found scaling factor with the significant wave height is in agreement with other found in the literature. However, the disagreement is slightly higher as one looks at higher
waves related to low probabilities, suggesting that the overprediction cannot be simply solved by considering a scaled Rayleigh distribution as in Modified Rayleigh distribution I and II (Chapter 8). Therefore, the crest-to-trough theory is used but the observed significant wave height is still lower than the one predicted by this theory.

Note that the possible reasons for the Rayleigh overprediction are:

a) Spectral width, comprising of two aspects:

i) In a wide spectrum, the calculated wave envelope does not vary slowly with time and overpredicts wave crests (see Figure 5.2) and therefore wave heights too. In particular the bimodal character of some spectra is one of the most important factors. Rodríguez et. al (2002) pointed out that the increase of the intermodal distance in a bimodal spectrum gives rise to the generally observed overprediction of large wave heights by the Rayleigh distribution.

ii) Although the wave envelope properly predicted wave crests and troughs, the assumption of estimating the wave height as twice the crest would produce overprediction for the statistics of large wave heights. This assumption, at the same time, assumes: symmetric wave envelopes, the crest having the same value as the trough. In the simplified example below, it is explained why the random asymmetry of waves produces overprediction of large wave heights, whereas the mean is not overpredicted.

<table>
<thead>
<tr>
<th>Height</th>
<th>Crest / Trough</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{\text{mean}} = H + h = 2\eta_{c/t,\text{mean}}$</td>
<td>$\eta_{c/t,\text{mean}} = (H + h)/2$</td>
</tr>
<tr>
<td>$H_{1/2} = H + (H + h)/2 &lt; 2\eta_{c/t,1/2}$</td>
<td>$\eta_{c/t,1/2} = H$</td>
</tr>
</tbody>
</table>

![Figure 12.8 Sketch of random asymmetry in waves](image)

b) The buoy may underestimate extreme waves. In fact, in the WADIC project (1989), they conclude that the WAVERIDER tend to underestimate the wave height.

Errors related with instrumentation probably exist. However, with the present information it is difficult to quantify them and therefore only attention to the first ones will be paid. The crest-to-trough distribution, which is theoretically better than the other analysed above, tries to solve the problem by considering a certain time lag between crest and trough. For i), more research is needed since there is only one fully developed theory relating the wave height and spectral bandwidth: Longuet-Higgins & Cartwright (1956), but it is derived for all local maxima (for the crest)
and not the global maxima that we are interested in. In addition, ε is a very sensitive parameter to the high frequency tail of the spectrum.

12.2.2. Maximum wave height

Apart from the significant wave height, the expected maximum wave height is also an important parameter. Note that the maximum is not a parameter indirectly involved in the distribution of the wave height but a random variable itself. Therefore, in order to properly compare the observed value with the expected value of the linear theory, an averaging process is necessary. Firstly, in a way analogous to the other parameters, the observed maximum wave height is normalized by the standard deviation of each record. This maximum only depends on the number of waves in a record. All the records have the same order of magnitude of waves (250-300 waves) and they have been concatenated to create longer records with higher number of waves.

Considering that there are approximately 40,000 records, the total amount of waves \( N_T \) is about 10 million. For example, for calculating the observed maximum height which corresponds to \( n \) waves, firstly, the records are concatenated in such a way that each group has approximately \( n \) waves. The maximum of each group is taken. There are approximately \( \frac{N_T}{n} \) maximum heights (associated to \( n \) waves) and the average of them is the final associated maximum to \( n \) waves, obtained from observations.

The fact of having a large amount of data here plays an important role. In general, having more observations give more reliability to the empirical results. Moreover, this particular analysis allows the studying what happens when considering large values of \( N \). According to linear theory, the values of the maxima increase only slowly with increasing \( N \) because of this dependency:

\[
E_N \{ H_{\text{max}} \} \approx 2 \left( 1 + \frac{0.29}{\ln N} \right) \sqrt{2 \ln N}
\]

Figure 12.9 illustrates the above results, using a logarithmic scale for \( N \). Once more, as for the significant wave height, the estimations of linear theory overpredict the observations but now (for the maximum wave height) the scaling factor is noticeably lower: 0.863 instead of 0.931. The tendency of a lower scaling factor (higher discrepancy) for higher heights is corroborated (see Section 12.2.1). Forristall (1978) found a slightly higher value for the maximum wave in \( N = 1000 \) of 116 hours of hurricane waves generated in the Gulf of Mexico: 0.907. On the contrary, Earle (1975) found that the maximum wave height was well predicted by the linear theory but he only analysed records of 200 waves and, instead of considering the expected value of the maxima, he calculated the 50 percentile value, which for 200 waves is approx. 2% lower than the expected one and he used the observed significant (7% lower than the Rayleigh one). So, in fact, Earle (1975) did not exactly use the Rayleigh distribution. However, this 9% (roughly 2% plus 7%) difference does still not explain why the present discrepancy is higher (14%). Sobey et al (1990) also compared the highest wave heights in a standard 20 minute record (approx. 300 waves) from tropical Cyclone Victor using linear theory. His data showed a systematic overprediction of 10%, practically the same as the discrepancy obtained in the analysis of 1000 waves of hurricane-generated waves in the Gulf of Mexico (1978). Despite 10% being higher than the 7% commonly found value for the significant wave height, it is still lower than the present discrepancy for the maximum wave height: 14%. The maximum wave height is 8.52 m.
Using the crest-to-trough theory, Massel (1991) calculated the reduction of the maximum wave height (compared to linear theory) for various sample sizes considering the crest-to-trough distribution. It is about 0.93: a reduction of 7% is not high enough to explain the discrepancy found in the observed mean maximum wave height. For the Modified Rayleigh I and II distribution, the discrepancies for the maximum wave height are, respectively, 7% and 4%.

As explained in Chapter 9, one can include the effect of the spectral band width in the calculation of the maxima by considering a reduced effective number of waves (Cartwright, 1958). The spectral band width used is the mean of all the observations: $\varepsilon = 0.69$. The reduction of the expected maximum wave height is practically constant (varying from 2.5 to 1%). Remind in this analysis $N > 500$ and that the reduction was more variable for $N < 100$, see Chapter 9.

Then, one could roughly approximate the estimated overprediction for the maximum wave height, in linear theory, as the sum of the two obtained previously (from crest-to-trough distribution and from considering an effective number of waves) achieving a discrepancy of 8-10.5%. This discrepancy agrees with the results of Sobey et al. (1990) and Forristall (1978) presented above but it does not explain the 14% discrepancy in the Mediterranean data.

Finally, it is important to note that the quality control explained in Chapter 2 has been crucial in order to have reliable results. More precisely, in the analysis of the maximum wave height the effects are noticeable. For example, without rejecting the data with significant wave height lower than half meter, the observations show a large enhancement for large $N$ (see Figure 12.10). The effects for low $N$ are practically null, probably because of the averaging process. The reason is not well-known.
12.2.3. Summary of results

In Table 12.1, the observations/expectations ratios for some of the theories considered in the present study (those which assume surface elevation Gaussian distributed) are summarized:

Table 12.1 Relation observations/expectations and % of discrepancy, according to different theories

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Rayleigh</th>
<th>Modified Rayleigh I (Longuet-Higgins)</th>
<th>Modified Rayleigh II (Naess)</th>
<th>Crest-to-trough</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_{mean} )</td>
<td>0.96</td>
<td>-4 %</td>
<td>1.04</td>
<td>+4 %</td>
</tr>
<tr>
<td>( H_{max} )</td>
<td>0.95</td>
<td>-5 %</td>
<td>1.02</td>
<td>+2 %</td>
</tr>
<tr>
<td>( H_s )</td>
<td>0.93</td>
<td>-7 %</td>
<td>1.00</td>
<td>0 %</td>
</tr>
<tr>
<td>( H_{max} )</td>
<td>0.86</td>
<td>-14 %</td>
<td>0.93</td>
<td>-7 %</td>
</tr>
<tr>
<td>( H_{max}(c) )</td>
<td>-12 %</td>
<td>-5 %</td>
<td>-2 %</td>
<td>-5 %</td>
</tr>
</tbody>
</table>

The modified Rayleigh I is the one with less discrepancy for all the studied parameters. In second place, there is the crest-to-trough distribution. However, bear in mind that the Modified Rayleigh I theory also considers a scaling factor for the wave crest distribution. In Section 12.3, it will be shown how the crest does not suffer from the same discrepancy as the height. Therefore, the crest-to-trough, becomes the “best” theory although its results are not convincing. In fact, Tayfun (1990) recognized that, from observations, a systematic overprediction of 4% for the mean wave height was found, just as in linear theory; the same is obtained in the present study.

For the maximum wave height there is an important overprediction for all the theories. With \( H_{max}(c) \) one wants to denote that the estimated values for the maxima are obtained considering...
the influence of $\varepsilon$. Remember that the inclusion of $\varepsilon$ has to do with the fact that the considered wave amplitude in the linear theory is larger than the real amplitude. Such an influence has been roughly calculated, using the effective number of waves (see Chapter 9), by subtracting 2% from all discrepancies of the maxima (in absolute value). However, the remaining discrepancy is still considerable.

12.3. Wave crest/trough

12.3.1. Significant wave crest/trough

The significant crest/trough is computed according to linear theory: 
\[ \eta_{\text{crest, s}} = \eta_{\text{trough, s}} = 2\sqrt{m_0} \]

and then compared with observations, as well as mean and root-mean-square values. Figure 12.11 and 12.12 illustrate the results. Compared to the wave height, both crest and trough show a better agreement with linear theory: the scaling factor for the significant value is, in both cases, around 0.98 (for the crest slightly higher than the trough), and there is no evident tendency of higher scaling factor for statistics related to larger heights.

![Figure 12.11 Comparison between observed and estimated parameter of wave crest (linear theory)](image)
Surprisingly, the discrepancies found for the crest and trough are little, above all in the case of the crest, suggesting that the Rayleigh distribution is quite acceptable. This agrees with the crest-to-trough theory for the wave height (which assumes Rayleigh distribution for, respectively, crests and trough) whereas the Modified Rayleigh I becomes worse.

### 12.3.2. Maximum wave crest/trough

Following the same procedure as in the calculation of the maximum wave height, the maximum crest/trough is computed. In Figure 12.13 the results are compared with linear theory expectations. The scaling factor is, respectively, for the crest: 0.991, and for the trough: 0.962.

Therefore, in contrast to the results of the maximum wave height, the maximum crest and trough are reasonably well predicted by the linear theory. In fact, the maximum wave crests corresponds well with the results of Cartwright (1958, see Figure 11.1) Although he considered smaller values of \( N \) (\( N < 10^4 \)). In order to further appreciate quantitatively the agreement with linear theory, the 95% confidence limits have been plotted. The crest, except for low waves, is comprised inside the interval whereas the trough is, until \( N = 5 \times 10^3 \), below it. The confidence interval has been calculated as 1.96 the standard sampling error \( S \) (assumed as being Gaussian distributed); the standard variation divided by the size of the sample. The variance of the maximum crest has been approximated using the asymptotic expression (Cartwright, 1958):

\[
S = \frac{\sqrt{Var(\eta_{max\ crest})}}{\sqrt{n}}, \quad \text{where} \quad n = \frac{\text{num total waves}}{N}
\]

\[
Var(\eta_{max\ crest}) = E(\eta_{max\ crest}^2) - E^2(\eta_{max\ crest}) \approx \frac{1}{2 \ln(N)} \left(1.6449 - \frac{2.1515}{\ln(N)}\right)
\]
Although the general agreement with linear theory, the trough is slightly lower than the crest, leading to us to think that nonlinear effects are present and therefore the crest is more peaked and higher than the trough. However, these effects seem to be weak.

Including the effect of the spectral band width, according to Cartwright (1958), the obtained discrepancy is the same as in the wave height: 1 - 2.5% (see Figure 12.14). The estimated maximum wave amplitude with the inclusion of the spectral width lays between the maximum crest level and maximum trough level, suggesting that, in general, the linear theory slightly overpredicts the wave amplitude (both crests and troughs). However, the crest seems to be quite well predicted due to the possible presence of weak nonlinear effects. In the opposite sense, the wave trough appear to be affected by such nonlinearities.
12.4. Nonlinearities

12.4.1. Introduction

From the above showed results in which neither the wave height nor crest/trough are underpredicted by the linear theory, one may suspect that the surface elevation can be almost considered Gaussian distributed. In the present section, one justifies with the observed skewness and kurtosis parameter (of the surface elevation) that nonlinearities can be neglected in this case. Note that skewness is related to the asymmetry in the surface elevation and the kurtosis with the enhancement of wave heights. Finally, the BFI is considered which, theoretically, in the narrow-band case is related to the kurtosis.

12.4.2. Kurtosis and Skewness

In Chapter 6 the nonlinear effects were discussed and it was concluded that they can be explained, in part, with the kurtosis \( k \) and skewness \( s \) of the surface elevation. Therefore, although they are not spectral parameters, it is interesting to analyse them and see if they differ from the Gaussian ones. The Gaussian distribution (linear theory) has \( k = 3 \) and \( s = 0 \). Sometimes, the kurtosis is "normalized" by subtracting 3 although here the first definition is used.

Figure 12.15 illustrates the agreement of one “mean” record with the Gaussian distribution (in red). With “mean” record, one means a record with kurtosis and skewness similar to the mean values of all the records (see Table 12.2 and 12.3).
Both kurtosis and skewness are computed for all records and then the mean and the standard deviation are calculated. In Figure 12.16, the distribution of each parameter is illustrated. For more detailed results see Table 12.2 and 12.3 in which a distinction for each buoy is made.

<table>
<thead>
<tr>
<th>kurtosis</th>
<th>Roses</th>
<th>Blanes</th>
<th>Llobregat</th>
<th>Tortosa</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.9853</td>
<td>2.9936</td>
<td>2.9877</td>
<td>3.0361</td>
<td>3.0100</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.1510</td>
<td>0.1575</td>
<td>0.1583</td>
<td>0.1674</td>
<td>0.1625</td>
</tr>
<tr>
<td>Maximum</td>
<td>4.1033</td>
<td>4.0175</td>
<td>3.9500</td>
<td>4.0219</td>
<td>4.1033</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.4455</td>
<td>2.4667</td>
<td>2.5175</td>
<td>2.5390</td>
<td>2.4455</td>
</tr>
</tbody>
</table>
Table 12.3 Distribution of the skewness for each buoy

<table>
<thead>
<tr>
<th></th>
<th>Roses</th>
<th>Blanes</th>
<th>Llobregat</th>
<th>Tortosa</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0062</td>
<td>0.0069</td>
<td>0.0056</td>
<td>0.0448</td>
<td>0.0219</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0562</td>
<td>0.0531</td>
<td>0.0528</td>
<td>0.0701</td>
<td>0.0647</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.3949</td>
<td>0.2748</td>
<td>0.1738</td>
<td>0.3611</td>
<td>0.3949</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.2101</td>
<td>-0.2406</td>
<td>-0.2467</td>
<td>-0.2235</td>
<td>-0.2467</td>
</tr>
</tbody>
</table>

From the statistical viewpoint, the kurtosis has to do with the peakedness of the probability density function of the surface elevation: the higher the kurtosis, the more peaked the distribution. Physically, a higher kurtosis is related to a higher probability of encountering higher (large) wave heights. If the probability density function is more peaked, the probability of mid range wave heights is reduced whereas both surface elevation around zero mean level and at high extremes are enhanced. In Figure 12.17, the maximum and mean wave height per record, both normalized by the standard deviation, are plotted against the kurtosis. Having the Rayleigh-Edgeworth distribution in mind, the tendencies are those expected (see Chapter 11), although the correlation is not so high: the maximum wave height increases with the increase of kurtosis whereas for the mean height is the opposite. In the picture of the maximum height a certain dispersion is expected because the number of waves among all records is slightly different. However, this does not appear to be the main cause because, in the case of the mean height, which does not depend on the number of waves, the dispersion is also present. The red point in the second picture, represents the linear theory.

![Figure 12.17 Normalized maximum and mean wave height vs. Kurtosis.](image)

In order to clearly see the relation between the maximum wave height and the kurtosis parameter, Figure 12.18 shows the dependency of this variable on $N$ and kurtosis simultaneously. In a analogous way as in the Section 12.2.2, the normalized maximum wave height has been calculated for different values of $N$. The difference is now that the records have been split in 12 groups (each group with approximately the same number of records), in ascending order of kurtosis. The mean values of kurtosis for each group are: 2.75, 2.84 (yellow), 2.88, 2.92 (magenta), 2.95, 2.98 (green), 3.01, 3.05 (cyan), 3.08, 3.13 (dark blue), 3.20, 3.35 (black). The stratification of colours clearly shows that a relation exists between kurtosis and encountering higher normalized maximum wave heights.
Figure 12.18 Mean normalized maximum wave height for different $N$. The colors are associated to mean values of kurtosis (see text above).

The mean kurtosis over all records is practically the same as in the case of linear theory but there is a clear dependency of high maximum wave heights for high (larger than 3) values of kurtosis. However, the Rayleigh-Edgeworth distribution, which is the theory which includes the kurtosis parameter (in the definition of BFI), does not seem to be the appropriate probability function to describe this stratification because it does not account for kurtosis lower than the 3 (Gaussian distribution). Point out that a sine wave has a kurtosis of 1.5. Note that, considering the value of $\sqrt{m_0} = 0.3$ m (approximately the mean value of the present data set), the variation of the maximum wave height (see Figure 12.18) can be roughly approximated by $\Delta \approx 2\sqrt{m_0} \approx 0.6$ m. In contrast, with the skewness, there is no stratification (see Figure 12.19). The mean values of skewness for each group are: -0.0877, -0.0486 (yellow), -0.0291, -0.0139 (magenta), -0.0008, 0.0116 (green), 0.0241, 0.0374 (cyan), 0.0523, 0.0705 (dark blue), 0.0966, 0.156 (black). The deviation for large wave heights is, in part, expected due to the random variability.
Figure 12.19 Mean normalized maximum wave height for different $N$. The colors are associated to mean values of skewness.

Figure 12.20 and 12.21 illustrate the time trace of kurtosis and skewness in which the 95% confidence intervals are plotted in dashed lines. Such intervals have been found by simulating with the Monte Carlo method 1000 samples of surface elevation using the Gaussian distribution. Note that the interval of the Tortosa buoy for kurtosis and skewness is wider because the number of data points in the wave record is smaller.

Figure 12.20 Time trace of kurtosis, concatenating all the records of each buoy.
The skewness measures the asymmetry of the probability density function. For example, a positive skewness means that there is an elongated tail to the higher values. This agrees with the skewed profile of surface elevation, with higher and peaked crests and shallower and rounded troughs, meaning that the crests/troughs are larger/shorter (see Figure 6.2).

In contrast to the kurtosis, the mean value of skewness is relatively higher than the Gaussian one (zero) compared to the kurtosis. However, in this case it is interesting to pay attention to the time trace illustrated in Figure 12.21. It is noticeable that in the last few years, the buoy of Tortosa clearly registered a higher mean compared to the rest of the years and buoys. Although the reason is unknown, it seems related to a registration error because during the same
years, the rest of the buoys do not suffer from this higher mean skewness and, curiously, this "jump" is found at the beginning of 2001, coinciding with the second period of available data in Tortosa (in Chapter 2, it was explained that Tortosa has two periods of available raw data: 1991-1997 and 2001-2006, 2004 is missing). Therefore, at first sight, it seems to be very probable that in the replacement of the buoy, the installed device was not properly moored. Without considering the second part of data of Tortosa, the mean of the skewness drops down from 0.0219 to 0.0036, whereas the mean of the second period is 0.0758. However, the affection to statistics is little, above all for the wave height. The observed maximum wave crest / trough without considering this period of larger skewness, is compared with the linear theory expectations in Figure 12.23. There is little variation, above all in the mid range of number of waves, in which the crest is slightly lower and the trough higher (the wave height is practically not affected).

![Figure 12.23 Observed mean maximum wave crest / trough compared to the linear theory (the period 2001-2006 of Tortosa buoy is not considered)](image)

In conclusion, the mean kurtosis and skewness are practically the same as the Gaussian ones. However, it is seems that for higher kurtosis, the maximum wave height tends to be higher and for higher skewness, the same happens for the crest.

12.4.3. **BFI**

The BFI is a spectral parameter which according to Janssen (2003) is closely related to the kurtosis in the narrow band case. It is important to note that, in general, spectral parameters are preferable to statistical ones because in the first case, one only needs the spectrum which can be predicted. This parameter has been calculated for all records. The mean value is 0.32 (see Figure 12.24) but there is no apparent correlation between kurtosis and BFI (see Figure 12.25), and, therefore, between large wave heights and BFI. That reason is probably due to the narrow assumption made by Janssen (2003). In fact, considering the quarter of records with higher maximum wave height, the mean of the BFI remains almost the same. A similar conclusion was
yield by Olagnon & Magnusson (2004). Alber (1978) demonstrated that the BFI vanishes if the wave spectrum is sufficiently broad. That could be the main reason but also the fact of having mixed states. Rotés (2004) analysed more than 4,000 spectrums in the Mediterranean Sea and found that swell reduces considerably the value of the BFI. Nevertheless, although she calculated the BFI using an alternative method (using the peak frequency and half the spectral width at half the peak energy, see Chapter 6), she found little relation between highest waves and highest BFI, either. In addition, the presence of low frequency noise in the spectrum could increase such a reduction.

Figure 12.24 Histogram of the BFI values of all records

Figure 12.25 Observed kurtosis vs. calculated BFI
12.5. Conclusion

In front of these results, it seems that the overprediction of the linear theory for both the crest and trough is practically negligible and therefore the wave envelope (of the linear theory) itself practically does not overpredict the crests and troughs. On the contrary, the wave height differs considerably. Therefore, the assumption of the wave height being twice the wave amplitude appears to be the most important cause of such an overprediction although other factors are probably also involved. On the other hand, the nonlinearities appear to be weak and, therefore, the assumption of surface elevation being Gaussian distributed appears to be quite acceptable. However, a definitive conclusion cannot be yield only from the analysis of buoy measurements which may tend to avoid high crests.