

Article

# On the Non-Gaussianity of Sea Surface Elevations

Alicia Nieto-Reyes <sup>†</sup> 

Department of Mathematics, Statistics and Computer Science, Universidad de Cantabria, 39005 Santander, Spain; alicia.nieto@unican.es

<sup>†</sup> Current address: Faculty of Science, Universidad de Cantabria, Avd. Los Castros s/n, 39005 Santander, Spain.

**Abstract:** The sea surface elevations are generally stated as non-Gaussian processes in the current literature, being considered Gaussian for short periods of relatively low wave heights. The objective here is to study the evolution of the distribution of the sea surface elevation from Gaussian to non-Gaussian as the period of time in which the associated time series is recorded increases. To do this, an empirical study based on the measurements of the buoys in the US coast downloaded at a casual day is performed. This study results in rejecting the null hypothesis of Gaussianity in below 25% of the cases for short periods of time and in over 95% of the cases for long periods of time. The analysis pursued relates to a recent one by the author in which the heights of sea waves are proved to be non-Gaussian. It is similar in that the Gaussianity of the process is studied as a whole and not just of its one-dimensional marginal, as it is common in the literature. It differs, however, in that the analysis of the sea surface elevations is harder from a statistical point of view, as the one-dimensional marginals can be Gaussian, which is observed throughout the study and in that a longitudinal study is performed here.

**Keywords:** Gaussian process; normal distribution; nortsTest R package; random projections; stationarity; time series analysis



**Citation:** Nieto-Reyes, A. On the Non-Gaussianity of Sea Surface Elevations. *J. Mar. Sci. Eng.* **2022**, *10*, 1303. <https://doi.org/10.3390/jmse10091303>

Academic Editors: Angelica Lo Duca and Andrea Marchetti

Received: 10 August 2022

Accepted: 9 September 2022

Published: 15 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Much attention in the literature is dedicated to the study of the sea surface height [1–3], a function of the sea surface elevation which is generally obtained by making use of the zero-up or down crossing methodology. The sea surface height is relevant because of design and analysis of off-shore structures [4] and ships [5] and, therefore, the literature is large in terms of studying its distribution [6–10]. The sea surface height has been modeled, for instance, as

- a Rayleigh distribution [11,12],
- a, more general, Weibull distribution [13],
- a Forristall distribution [1],
- a Naess distribution [14],
- a Boccotti distribution [15],
- a Klopman distribution [16],
- a van Vledder distribution [17],
- a Battjes–Groenendijk distribution [18],
- a Mendez distribution [19], or
- a LoWiSh II distribution [20].

In [21] it is experimentally proved that the sea heights do not follow a Gaussian distribution. This study is dedicated, however, to the analysis of the sea surface elevation, as opposed to the sea heights. The measurements of sea surface elevation are obtained by buoys throughout the sea, which are later preprocessed to obtain the sea heights. Sea surface elevations have been studied from a statistical point of view, studying its distribution [22], the skewness of the distribution [23], and the modellization of the process [24,25]. Consideration has also being given to how to measure [26] and record the data [27]. From an

applied perspective, the literature contains works on sea surface elevations to, for instance, ship motion forecasting [28] and the development of sea surface elevation maps [29]. Furthermore, ref. [30] studies that for certain wave groups the periods of stationarity are short and [31] that the sea surface elevation is only Gaussian in short periods of relatively low wave heights. For many sea states the process is nonlinear due to its second order structure [32], and even to much higher order structures for high waves [33].

This work goes beyond the existing literature and it is dedicated to empirically study how the Gaussianity of the distribution of the sea surface elevation evolves along an increase of the time period of the associated time series. For that, the measurements of 59 buoys along the US coast are studied. It is obtained that over 50% of the cases are non-Gaussian with a length time period corresponding to  $2 \times 10^4$  observations. This results increases to over 95% of the cases if the length time period increases to correspond to  $10^5$  observations. From a statistical point of view, the importance of studying the sea surface elevation is high and lies in that it is a raw measurement. While experimental studies show that the distribution of sea heights are clearly non-Gaussian, having a non-Gaussian one dimensional marginal, the non-Gaussianity of the sea surface elevation is not so obvious; which makes the problem more interesting.

In fact, in proving the non-Gaussianity, it is here demonstrated that some cases that common hypothesis tests consider as Gaussian correspond to non-Gaussian processes with Gaussian one-dimensional marginals. It is worth saying that out of a 22.03% rejection rate for a length time period corresponding to  $10^3$  observations, a 6.78% correspond to non-Gaussian processes with Gaussian one-dimensional marginals, the Gaussianity of which would not have been rejected but for the use of the methodology applied here. This methodology is based on the random projection test [34], a goodness of fit test that checks the Gaussianity of the process as a whole and not just of a finite order marginal, as other established test in the literature do; see, for instance [35,36]. The obtained findings are important due to the cases that the literature considered as Gaussian are numerous. These cases include very large waves and, in fact, according to [37], very large waves might be much more frequent than commonly assumed.

The rest of the manuscript includes: The description of the studied dataset in Section 2 and of the applied methodology in Section 3. The results of the analysis are described in Section 4. The analysis makes use of the `nortsTest` package of the R software. Section 5 contains the obtained conclusions.

## 2. Datasets

The Coastal Data Information Program (<https://cdip.ucsd.edu> (accessed on 20 August 2022)) contains surface elevations measured by buoys that are along the cost of the US. For the present study, these measurement were downloaded on the 20 August 2022 from the web page <https://thredds.cdip.ucsd.edu/thredds/catalog/cdip/realtime/catalog.html> (accessed on 20 August 2022). In particular, the variable downloaded is that named *xyzZDisplacement*. The set of data used here differs from that in [21] and it has not being used in the literature before.

There are a total of 59 datasets, each corresponding to the collected time series of a station (buoy). Each buoy has an identification number, which is displayed in the first and fourth columns of Table 1. The latitude and longitude coordinates of these buoys are displayed in Figure 1 and, rounded to 2 decimal values, also included in the table. The top plot of Figure 1 contains a world map where the coordinates have been drawn. The bottom plot is a zoom of the top one that contains only the coordinate of the buoys that are close to the US mainland coast.

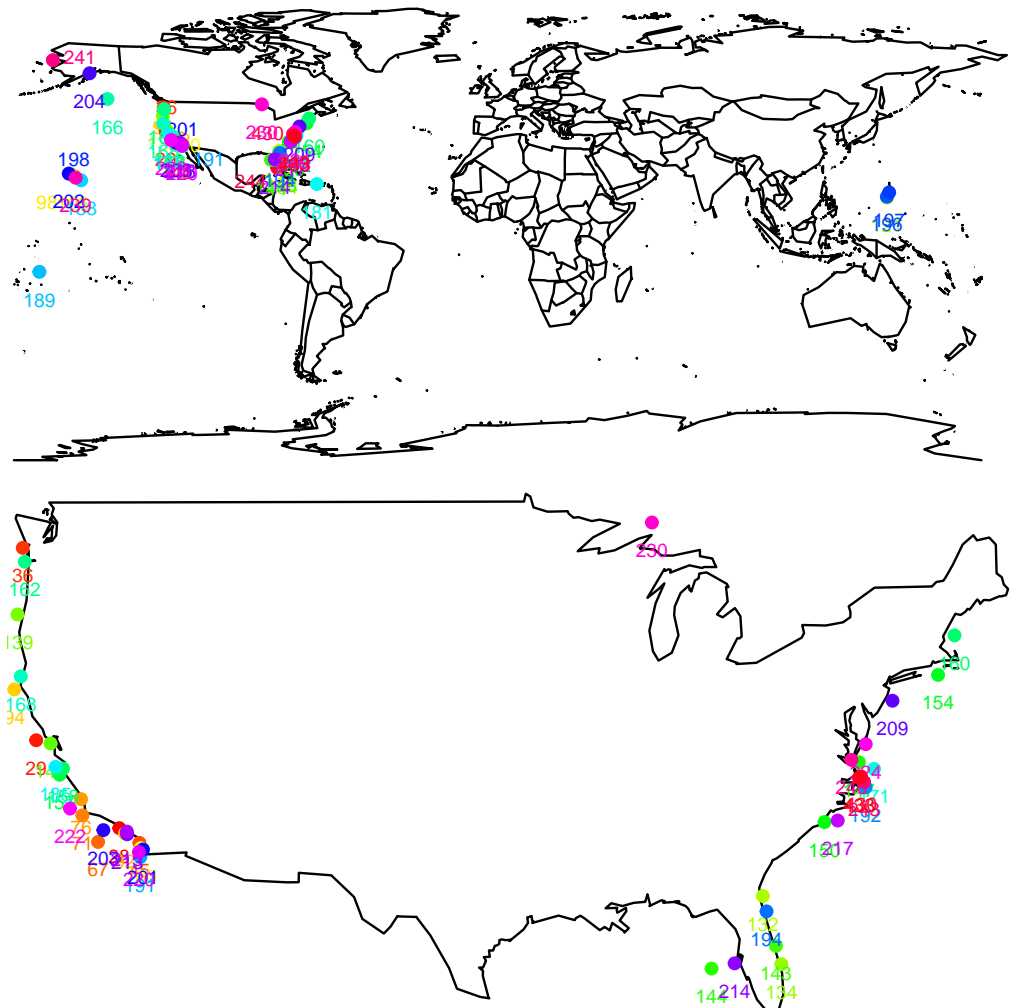
In Table A1, in Appendix A, it can be observed the time period of the time series associated to each of the 59 buoys, under the columns designated with the names Start Time (columns 2 and 3) and End Time (columns 4 and 5). The shortest time period is depicted in bold, which is that of station 244. This is the buoy with the shortest time period because the length of the associated time series is the smallest among that of the 59 time series. This

information is presented in Table A2, also in Appendix A, under the columns designated with the name *length*.

**Table 1.** Buoys identification number, with associated longitude and latitude coordinates, whose surface elevation measurements constitute the datasets analyzed in this paper.

Buoy	Latitude	Longitude	Buoy	Latitude	Longitude
028	33.86°	−118.64°	185	36.7°	−122.34°
029	37.94°	−123.46°	188	19.78°	−154.97°
036	46.86°	−124.24°	189	−14.27°	−170.5°
045	33.18°	−117.47°	191	32.52°	−117.43°
067	33.22°	−119.87°	192	35.75°	−75.33°
071	34.45°	−120.78°	194	30°	−81.08°
076	35.2°	−120.86°	196	13.68°	144.81°
092	33.62°	−118.32°	197	15.27°	145.66°
094	40.29°	−124.73°	198	21.48°	−157.75°
098	21.41°	−157.68°	201	32.87°	−117.27°
100	32.93°	−117.39°	202	22.28°	−159.57°
106	21.67°	−158.12°	203	33.77°	−119.56°
121	13.35°	144.79°	204	59.6°	−151.83°
132	30.71°	−81.29°	209	39.77°	−73.77°
134	27.55°	−80.22°	213	33.58°	−118.18°
139	43.77°	−124.55°	214	27.59°	−82.93°
142	37.79°	−122.63°	215	33.7°	−118.2°
143	28.4°	−80.53°	217	34.21°	−76.95°
144	27.34°	−84.27°	220	32.75°	−117.5°
147	36.92°	−75.72°	222	34.77°	−121.5°
150	34.14°	−77.72°	224	37.75°	−75.33°
154	40.97°	−71.13°	230	48.03°	−87.73°
157	36.33°	−122.1°	239	20.75°	−157°
158	36.63°	−121.91°	240	37.02°	−76.15°
160	42.8°	−70.17°	241	64.47°	−165.48°
162	46.22°	−124.13°	243	36°	−75.42°
166	50.03°	−145.2°	244	24.41°	−81.97°
168	40.9°	−124.36°	430	36.26°	−75.59°
171	36.61°	−74.84°	433	36.2°	−75.71°
181	18.38°	−67.28°			

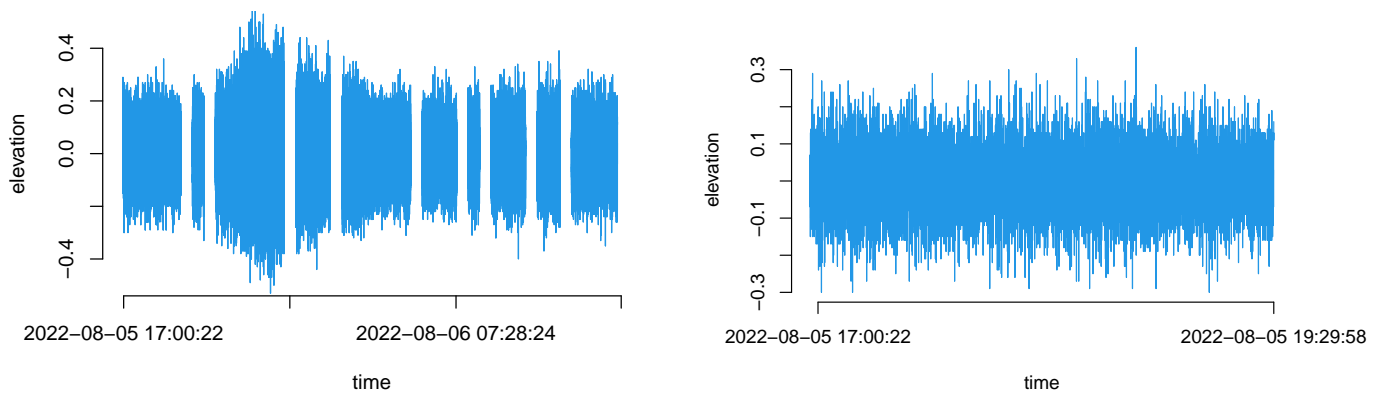
Each of the 59 datasets under study is restricted to a time series of length  $5 \times 10^5$ . The first  $5 \times 10^5$  are the ones selected. The datasets consist of raw data, which contains unknown, unobserved, values. After taking out those from the first  $5 \times 10^5$  selected, the length of the time series associated to each of the 59 buoys can also be observed from Table A2, in Appendix A. It is designated by the name *studied*. It can be observed from Table A2, in Appendix A, that there is one buoy for which the whole  $5 \times 10^5$  first elements have been observed: buoy 249. There, it can also be observed that there are two buoys, 188 and 202, that have the minimum value under the label *studied*.



**Figure 1. Top panel:** World map with the coordinates of the 59 buoys whose measurements constitute the datasets analyzed in this paper. **Bottom panel:** Zoom of the top panel that shows the US mainland with the coordinates of the buoys close to it. The identification number is that in Table 1.

In the left panel of Figure 2, it is displayed part of the recorded data for buoy 433. The displayed data results from restricting the 1,622,186 observations stored for buoy 433 and taking the ones corresponding to the first  $10^5$  time points. As it is obvious from the plot, the first  $10^5$  time points contain unobserved elements. Indeed, 410,144 observations have been made out of the  $5 \times 10^5$  selected for the study (see Table A2, in Appendix A). From the left panel of the figure, it is also observable that the unobserved data splits the time series in ten parts. The right panel of Figure 2 is a zoom of the left panel containing solely the part to the left of the time series.

Greenwich mean time (GMT) in the format year-month-day and hours:minutes:seconds is used for the x-axis of the plots in the figure.



**Figure 2.** **Left panel:** representation of the  $10^5$  first recordings, including unobserved (missing) data, of the time series associated to buoy 433. The nine observed voids represent unobserved data. **Right panel:** representation of the first segment of the time series in the left panel. Sea surface elevation in meters and time in seconds GMT with the format year-month-day and hours:minutes:seconds.

### 3. Methodology

Given  $X_t$  a real valued random variable for each  $t \in \mathbb{Z}$ ,

$$X := \{X_t\}_{t \in \mathbb{Z}}$$

is a stochastic process [38]. Most common hypotheses on stochastic processes are those of stationarity [39] and Gaussianity [40].  $X$  is stationary if

- $E[X_t] = E[X_{t+k}]$  for all  $k, t \in \mathbb{Z}$ , where  $E$  denotes the expectation function,
- $\text{Cov}(X_t, X_k) = \text{Cov}(X_{t-k}, X_0)$  for all  $k, t \in \mathbb{Z}$ , where  $\text{Cov}$  denotes the covariance function and
- $\text{Var}[X_t] < \infty$  for all  $t \in \mathbb{Z}$ , where  $\text{Var}$  denotes the variance.

$X$  is Gaussian if

$$(X_{t_1}, \dots, X_{t_n}) \text{ is a Gaussian random vector for all } n \in \mathbb{N}.$$

It occurs that a stationary Gaussian process is strictly stationarity.  $X$  is strictly stationary if

$$(X_{t_1}, \dots, X_{t_n}) \text{ and } (X_{t_1+k}, \dots, X_{t_n+k})$$

are equally distributed for all  $n \in \mathbb{N}$  and  $k, t_1, \dots, t_n \in \mathbb{Z}$ . Consequently, given a stationary process  $X$ , it is Gaussian if

$$(X_t, \dots, X_t) \text{ is a Gaussian random vector for all } t \in \mathbb{N}. \tag{1}$$

#### 3.1. Tests for Stationarity

This manuscript is about testing the Gaussianity of stochastic processes. Typically, those tests assume that the process is stationary. Thus, this assumption has to be previously checked. For that, the most common tests in the literature are

1. Ljung-Box test [41],
2. Augmented Dickey-Fuller test [42],
3. Phillips-Perron test [43] and
4. kpps test [44].

For the first three tests, the tests can be simplified as contrasting the null hypothesis

$$H_{0,1} : X \text{ is non stationary} \tag{2}$$

against the alternative

$$H_{a,1} : X \text{ is stationary}$$

while the kpps test results in the null hypothesis

$$H_{0,2} : X \text{ is stationary} \tag{3}$$

against the alternative

$$H_{a,2} : X \text{ is non stationary.}$$

The hypotheses are tested in different ways. For instance, Ljung-Box test makes use of the autocorrelation function, which, at lag  $k$  for a stationary process is

$$\frac{\text{Cov}(X_t, X_{t+k})}{\text{Var}(X_t)}.$$

This is observable from its statistic:

$$n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k},$$

where  $\hat{\rho}_k$  denotes the sample autocorrelation at lag  $k$  and  $n$  the sample size. Note that it depends on a constant  $h$ .

### 3.2. Tests for Gaussianity

Most tests for Gaussianity of stochastic processes assume the process is stationary and test whether a finite marginal distribution of the process is Gaussian, generally, the one-dimensional marginal. That is, instead of testing whether (1) is satisfied, these tests contrast the null hypothesis

$$H_{0,3} : X_t \text{ is a Gaussian random variable} \tag{4}$$

against the alternative

$$H_{a,3} : X_t \text{ is not a Gaussian random variable}$$

by checking whether  $X_t$  is a Gaussian random variable. Let us reflect that, because of the stationarity, the distribution of  $X_t$  is the same for all  $t \in \mathbb{Z}$ ; that is, it is independent of  $t$ .

Common tests to check the Gaussianity of a real valued random variable require a sample of independent and identically distributed random variables [45]. As this work deals with stochastic processes, the independence assumption is not verified. However, there are also many tests for this situation. Here, it is made use of the Epps test [35], which checks that the characteristic function of the one-dimensional distribution of the process is that of a Gaussian distribution, and of the Lobato and Velasco test [36], which checks that the third and fourth order moments of the one-dimensional distribution of the process are those of a Gaussian distribution.

If the null hypothesis  $H_{0,3}$  is rejected, with the above mentioned tests, the null hypothesis

$$H_{0,4} : X \text{ is a Gaussian process} \tag{5}$$

is rejected against the alternative

$$H_{a,4} : X \text{ is not a Gaussian process.}$$

However, it may occur that

$$X_t \text{ is a Gaussian random variable}$$

while

$$X \text{ is not a Gaussian process.}$$

The above mentioned tests are at nominal level again this type of alternatives. For this, it is used here the random projection test [34], which test the Gaussianity of the whole distribution of the process and not just of a finite dimensional marginal. For elaboration on it, see Section 3.2.

Random Projection Test

The random projection test was introduced in [34] as a tool to test the Gaussianity of stationary processes that is able to reject the null hypothesis of Gaussianity (5) against alternatives with Gaussian finite-dimensional marginals. The procedure is based on a result in [46] that implies that if

$$\langle \{X_j\}_{j \leq t}, d \rangle,$$

with  $d$  drawn from a Dirichlet distribution [47], is Gaussian, then

$$\{X_j\}_{j \leq t}$$

is Gaussian. Note that due to the stationarity assumption, the Gaussianity of  $\{X_j\}_{j \leq t}$  is equivalent to (1). In what follows, the procedure is explained in detail.

Let

$$\lambda_1, \lambda_2 > 0$$

be two parameters. Making use of the following stick-breaking procedure, a Dirichlet distribution is considered:

1. Let

$$\beta(\lambda_1, \lambda_2)$$

denote a beta distribution with parameters  $\lambda_1, \lambda_2$ .

2. Let  $d_0$  be drawn from the distribution  $\beta(\lambda_1, \lambda_2)$ . Note that

$$d_0 \in [0, 1].$$

3. For any  $k \in \mathbb{N}$ , the natural numbers, let  $d_k$  be the result of multiplying

$$1 - \sum_{i=0}^{k-1} d_i$$

and an element drawn independently from the distribution  $\beta(\lambda_1, \lambda_2)$ . Note that

$$d_k \in [0, 1 - \sum_{i=0}^{k-1} d_i].$$

Let  $X$  be a stationary process. The associated projected process based on  $\{d_k\}_{k \in \mathbb{N}}$  is

$$Y := \{Y_t\}_{t \in \mathbb{Z}}$$

with

$$Y_t := \sum_{i=0}^{\infty} d_i X_{t-i}.$$

Then, making use of this randomly projected process, it suffices to apply to it a test for the null hypothesis of Gaussianity (4).

The selection of the parameters  $\lambda_1, \lambda_2$  is important. It is explained in [34] that values such as

$$\lambda_1 = 100 \text{ and } \lambda_2 = 1$$

result in an projected process  $Y$  similar to  $X$ . However, values such as

$$\lambda_1 = 2 \text{ and } \lambda_2 = 7$$

result in projected processes different from  $X$  while providing an effective method.

### 3.3. False Discovery Rate

When multiple tests are performed, the multiplicity has to be taken into account. For that it is used here the false discovery rate [48]. The false discovery rate aims at controlling the expected proportion of falsely rejected hypothesis. It was first introduced in [49] to take into account the multiplicity of independent tests. In [48] it was established that the definition in [49] remains valid for certain types of dependency. However, for general dependent cases [48] has to be applied.

## 4. Results of the Analysis

This section analyzes whether each of the 59 datasets provided in Section 2, one per buoy, is drawn from a Gaussian process as the length of the time series increases. For that, the tests described in Section 3.2 are used here. As commented in Section 2, our datasets have been restricted to  $5 \times 10^5$  observations that include missing data; with the amount of non-missing observations being recorded in Table A2 of Appendix A under the label *studied*. The length of the time series along which this longitudinal study is performed, makes use of the following values

$$Length \in \{10^3, 10^4, 2 \times 10^4, 4 \times 10^4, 6 \times 10^4, 8 \times 10^4, 10^5\} \quad (6)$$

and is computed by selecting that amount of non-missing observations from the buoys that have them.

To relate the above length quantities to time, it is important to take into account that the time in seconds UTC (Coordinated universal time) associated to an observation  $t \in \mathbb{N}$  is computed as

$$T_t := T_0 + \frac{t - 1}{r} - d$$

where  $T_0$  is the time at which the recording starts,  $r$  is the sample rate and  $d$  is the filter delay.  $r$  takes value 1.28 and  $d$  133.3 but for buoys in

$$B := \{132, 142, 171, 194, 204 \text{ and } 244\}$$

that  $r$  takes value 2.56 and  $d$  value 130. Thus, the amount of recorded time,  $T_t - T_0$ , for  $t = 10^3$  observations corresponds to 647.1688 s in UTC time in general and to 260.2344 for the buoys in  $B$ . This results in GMT time in 10.79 min and 4.34 min, respectively. Table 2 displays these results, rounded to two decimal values, for each  $t$  equal to the length in (6). The table also includes the translation of seconds UTC into GMT time. Thus, it is observable from the table that the analysis we pursue here is of time series that have been recorded for a period that varies between 4 min and 21 h. In fact, in general, for a length of  $10^3$  the periods are close to 11 min while the are just above 2 h for a length of  $10^4$  and of more than 4 h for a length of  $2 \times 10^4$ .

Let us select, for instance, a time period of 2.13 h in general and of 1.05 h for the buoys in  $B$ . Then, for each of the buoys (but for buoys 188, 189, 202 and 204), the first non-missing observations corresponding to those time periods would be selected. Buoys 188, 189, 202 and 204 are not studied for these time periods because, as reported in Table A2 under the label *studied*, their non-missing observations are less than  $10^4$ . Thus, 55 out of the 59 buoys are studied in this scenario. This is not the case, however, when we study buoys recorded for 4.34 min as the time series duration associated to all the buoys is larger than that time period.

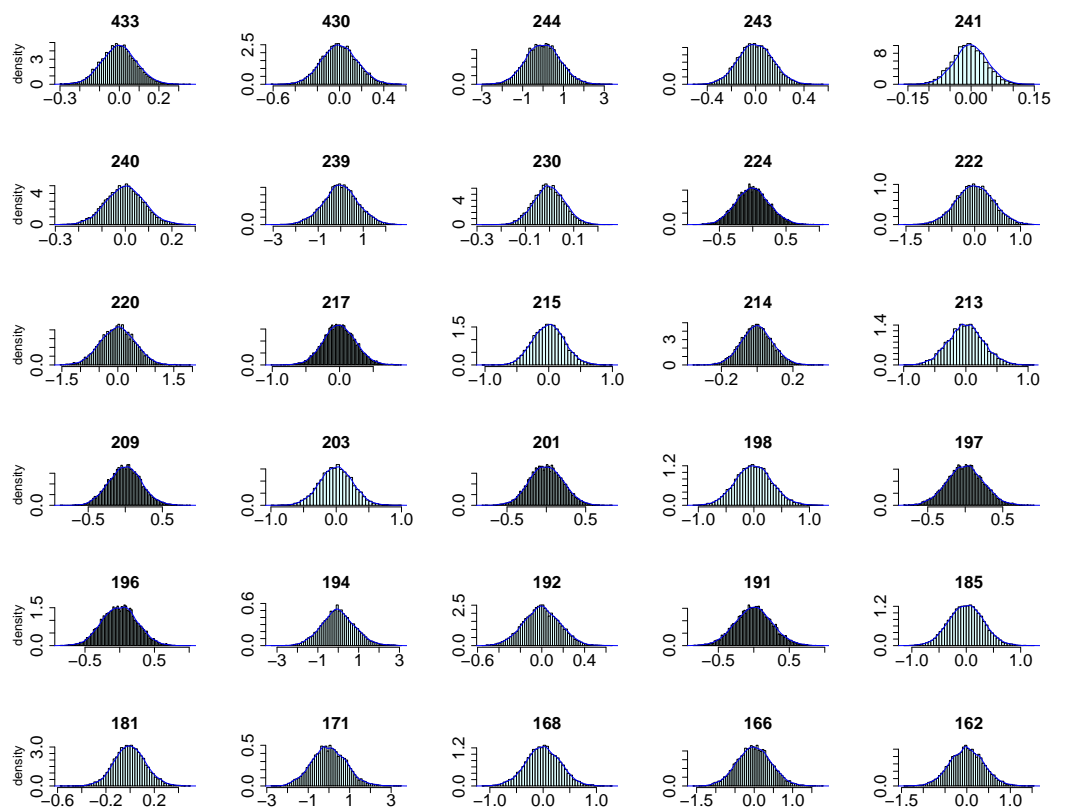
First, despite the time series associated to each buoy consists of non iid (independent identically distributed) drawn observations, in Figure 3 it is plotted the histogram and kernel density estimated curve associated to each of the 55 time series for a time period of 2.13 h in general and of 1.05 h for the buoys in  $B$ . For each histogram 70 cells have been used; but for buoy 241 that is plotted based on 40 cells. A Gaussian kernel is used



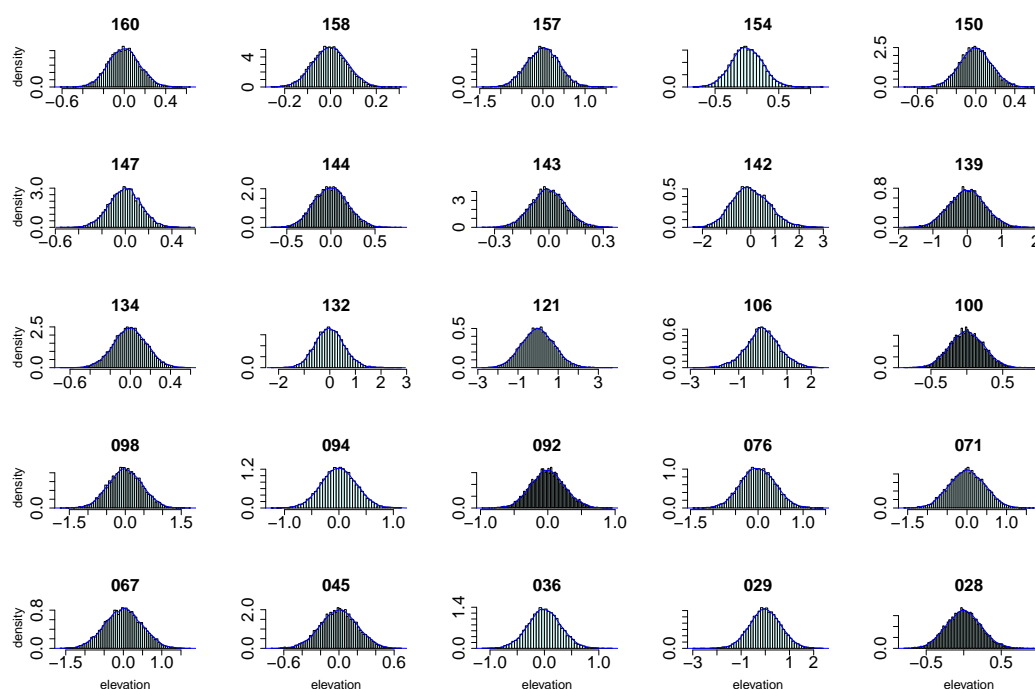
for the density estimations with the bandwidth resulting of applying least squares cross validation [50,51]. The obtained bandwidths are reported in Table A3, in Appendix A. These plots are done to display descriptive information on the datasets. Note that the null hypothesis of Gaussianity in (5) cannot be rejected by just inspecting the plots. This is due to the fact that when testing the null hypothesis of Gaussianity it is assumed  $X$  is a Gaussian process and the objective is to reject such assumption by the evidence provided by the data. This is done at certain significance level. Here, it is used 0.05. By inspection of the plots in Figure 3, one cannot guarantee whether there is enough evidence for such a rejection as it cannot be quantified.

**Table 2.** Time period in UTC and GMT associated to each length in (6). The times are divided in those for the 6 buoys in  $B$  and the rest of 53 buoys, which are labelled by general.

Length	UTC		GMT	
	General	$B$	General	$B$
$10^3$	647.17 s	260.23 s	10.79 min	4.34 min
$2.5 \times 10^3$	1819.04 s	846.1719 s	30.32 min	14.10 min
$10^4$	7678.42 s	3775.86 s	2.13 h	1.05 h
$2 \times 10^4$	15,490.92 s	7682.11 s	4.30 h	2.13 h
$4 \times 10^4$	31,115.92 s	15,494.61 s	8.64 h	4.30 h
$6 \times 10^4$	46,740.92 s	23,307.11 s	12.98 h	6.47 h
$8 \times 10^4$	62,365.92 s	31,119.61 s	17.32 h	8.64 h
$10^5$	77,990.92 s	38,932.11 s	21.66 h	10.81 h



**Figure 3.** Cont.



**Figure 3.** Histogram, based on 70 cells (but for buoy 241 on 40 cells), and kernel density estimated curve associated to each of the 55 buoys (time series) analyzed for a time period of 2.13 h in general and of 1.05 h for the buoys in *B*.

As the tests for testing the null hypothesis of Gaussianity reported in Section 3.2 require the stationarity assumption for the process, making use of the tests provided in Section 3.1, it is first checked whether each of the datasets, when making use of each of the time periods in Table 2, is drawn from a stationary process. The results obtained from checking the stationarity are displayed in Table 3. Only one result is provided by test and time period because the maximum of the obtained *p*-values is reported for the Augmented Dickey-Fuller, Phillips-Perron and Ljung-Box tests while the minimum is reported for the kpps test. Note that in the first three tests the null hypothesis of non-stationarity is tested, as in (2), and in the fourth it is the null hypothesis of stationarity, as in (3). Thus, *p*-values smaller than 0.01 are obtained for the Augmented Dickey-Fuller test and the Phillips-Perron test and larger than 0.1 for the kpps test.

*p*-Values that are close to zero are obtained for the Ljung-Box test, but for time periods of 17.32 h and 21.66 h in general, and of 8.64 h and 10.81 h for buoys in *B*. For a time period of 17.32 h in general, and of 8.64 h for buoys in *B*, the minimum *p*-value is 0.18 in the Ljung-Box test and is obtained for buoy 241. The other *p*-values of that test for that time period are technically zero. It is also for buoy 241 that the Ljung-Box test gives a *p*-value of 0.36 when the time period is of 21.66 h in general, and of 10.81 h for buoys in *B*. For that time period and test, the *p*-values associated to the rest of buoys are smaller than 0.01. When considering low time periods, the ultimate case being 4.34 min, all the tests provides the aimed result. Thus, it can be assumed (with the possible exception of buoy 241 for time periods of 17.32 h and 21.66 h in general and of 8.64 h and 10.81 h for buoys in *B*) that the studied datasets are drawn from stationary processes, and check their Gaussianity under the mentioned assumption.

**Table 3.** Summary of the obtained *p*-values, for different time periods, for each of the 59 studied datasets under four different stationarity tests: *Augmented Dickey-Fuller* (first column), *Phillips-Perron* (second column), *Ljung-Box test* (third column) and *kpps* (fourth column). The null hypothesis is of stationarity for the *kpps* test and of non-stationarity for the other three. For each of the lengths, the minimum *p*-value value over the studied buoys is displayed for the *kpps* test and the maximum in the other three cases.

Time Period		Tests			
General	B	<i>Augmented Dickey-Fuller</i>	<i>Phillips-Perron</i>	<i>Ljung-Box</i>	<i>kpps</i>
10.79 min	4.34 min	<0.01	<0.01	$8.66 \times 10^{-8}$	>0.1
30.32 min	14.10 min	<0.01	<0.01	$6.09 \times 10^{-10}$	>0.1
2.13 h	1.05 h	<0.01	<0.01	$1.48 \times 10^{-6}$	>0.1
4.30 h	2.13 h	<0.01	<0.01	0	>0.1
8.64 h	4.30 h	<0.01	<0.01	$2.60 \times 10^{-7}$	>0.1
12.98 h	6.47 h	<0.01	<0.01	0	>0.1
17.32 h	8.64 h	<0.01	<0.01	0.18	>0.1
21.66 h	10.81 h	<0.01	<0.01	0.36	>0.1

In order to study the Gaussianity of the datasets under study, it is first analyzed the one-dimensional marginal distribution of the process. This is because a rejection of the null hypothesis (4) implies the sought rejection of the whole distribution of the process, in (5). For analyzing the one-dimensional marginal distribution, it is made use of the Epps and Lobato and Velasco tests commented in Section 3. The results are displayed in Table 4 for the particular case of a time period of 2.13 h in general and of 1.05 h for the buoys in B. There, for each dataset, associated to a buoy (columns 1 and 5), it can be observed the *p*-values resulting from applying the Epps test (columns 2 and 6) and the Lobato and Velasco test (columns 3 and 7). As multiplicity has to be taken into account, columns 4 and 8 display the FDR values. It can be observed from the table that 20 of the 55 FDR values are smaller than 0.05. They have been highlighted in bold. If the less conservative FDR introduced in [49] had been used, the number of rejections would have increased. If multiplicity had not been taken into account at all and the null hypothesis (4) were rejected when the minimum of the two *p*-values was smaller than 0.05, the number of rejections would have increased to 22.

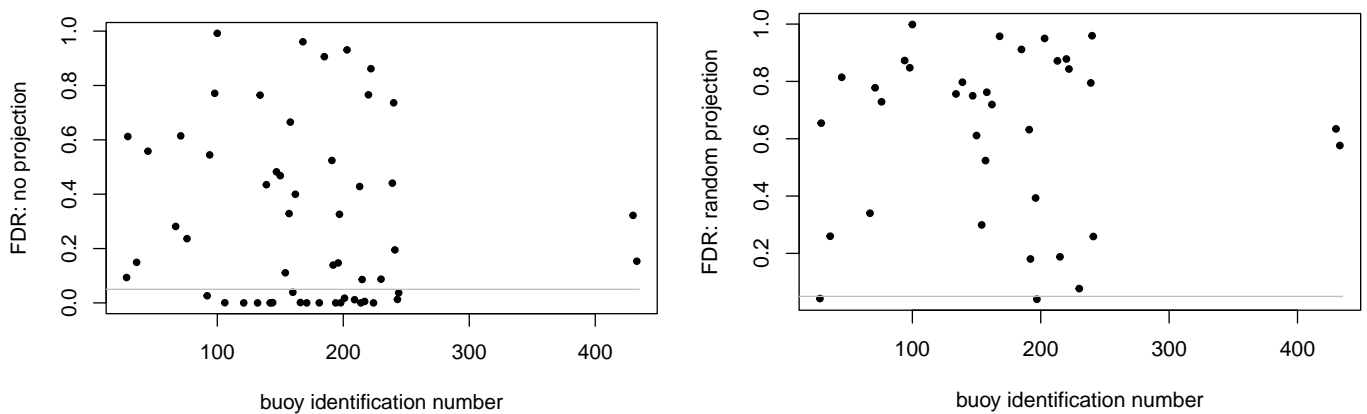
**Table 4.** *p*-Values resulting from applying the Epps test (columns 2 and 6) and the Lobato and Velasco test (columns 3 and 7) per dataset associated to each of the 55 buoys (columns 1 and 5) studied under a time period of 2.13 h in general and of 1.05 h for the buoys in B. FDR (columns 4 and 8) combination, for dependent *p*-values, of the two *p*-values per buoy, with the ones smaller than 0.05 highlighted in bold.

Buoy	Epps	L.-V.	FDR	Buoy	Epps	L.-V.	FDR
028	0.91	0.05	0.09	171	0.66	$7.14 \times 10^{-11}$	$1.43 \times 10^{-10}$
029	0.42	0.61	0.61	181	$1.61 \times 10^{-3}$	$1.54 \times 10^{-4}$	$3.07 \times 10^{-4}$
036	0.15	0.14	0.15	185	0.91	0.6	0.91
045	0.97	0.28	0.56	191	0.52	0.31	0.52
067	0.28	0.15	0.28	192	0.88	0.07	0.14
071	0.47	0.61	0.61	194	$8.07 \times 10^{-3}$	$2.76 \times 10^{-9}$	$5.53 \times 10^{-9}$
076	0.20	0.24	0.24	196	0.07	0.35	0.15
092	0.09	0.01	<b>0.03</b>	197	0.81	0.16	0.33
094	0.87	0.27	0.54	198	0.49	$1.58 \times 10^{-4}$	$3.17 \times 10^{-4}$
098	0.39	0.79	0.77	201	0.01	0.31	<b>0.02</b>

**Table 4.** Cont.

Buoy	Epps	L.-V.	FDR	Buoy	Epps	L.-V.	FDR
100	0.58	0.99	0.99	203	0.93	0.74	0.93
106	$3.29 \times 10^{-4}$	0.02	$6.58 \times 10^{-4}$	209	0.73	0.01	<b>0.01</b>
121	0.41	$1.43 \times 10^{-6}$	$2.86 \times 10^{-6}$	213	0.43	0.42	0.43
132	0.08	$3.08 \times 10^{-11}$	$6.16 \times 10^{-11}$	214	0.04	$6.35 \times 10^{-6}$	$1.27 \times 10^{-5}$
134	0.76	0.76	0.76	215	0.81	0.04	0.09
139	0.82	0.22	0.43	217	0.03	0.00	<b>0.01</b>
142	$3.56 \times 10^{-4}$	$1.59 \times 10^{-8}$	$3.18 \times 10^{-8}$	220	0.56	0.77	0.77
143	0.02	$1.49 \times 10^{-4}$	$2.98 \times 10^{-4}$	222	0.86	0.60	0.86
144	0.01	$5.51 \times 10^{-4}$	$1.10 \times 10^{-3}$	224	0.01	$3.13 \times 10^{-9}$	$6.27 \times 10^{-9}$
147	0.48	0.36	0.48	230	0.66	0.04	0.09
150	0.33	0.47	0.47	239	0.22	0.77	0.44
154	0.63	0.06	0.11	240	0.74	0.64	0.74
157	0.16	0.45	0.33	241	0.10	0.20	0.19
158	0.60	0.67	0.67	243	0.06	0.01	<b>0.01</b>
160	0.18	0.02	<b>0.04</b>	244	0.51	0.02	<b>0.04</b>
162	0.92	0.20	0.40	430	0.17	0.32	0.32
166	0.68	$6.87 \times 10^{-4}$	$1.37 \times 10^{-3}$	433	0.19	0.08	0.15
168	0.72	0.96	0.96				

To better illustrate the findings, the results of Table 4 are summarized in the left plot of Figure 4. The *x*-axis represents the buoy’s identification number while the *y*-axis displays the obtained FDR for dependent tests. A grey line at  $y = 0.05$  is drawn to show what buoys have a FDR above or below that value, which result in a rejection of the null hypothesis (4). It can observe that there are three FDR that are just above 0.05. They correspond to buoys 028, 215 and 230.



**Figure 4.** Left panel: FDR values corresponding to the studied buoys in Table 4. Right panel: FDR values corresponding to the buoys studied in Table 5. The line  $y = 0.05$  is displayed in both panels in color grey.

In what follows it is pursued a further study in the 35 buoys for which there is yet no evidence to reject the null hypothesis of Gaussianity, displayed in (5). This further study consists in applying the random projection test based on the Epps and Lobato and Velasco tests with parameters (100, 1) and (2, 7). The results of applying the random projection test are reported in Table 5. There it can be observed that the random projection test is able to reject the null hypothesis of Gaussianity in 2 out of the 35 buoys, which results in a total of 22 rejections out of 55 (the 40%). The FDR values that result in a rejection are highlighted in bold.

**Table 5.** FDR values (column 6) resulting of applying the random projection test for each buoy (column 1) with a FDR adjusted  $p$ -value larger than 0.05 in Table 4 (time period of 2.13 h in general and of 1.05 h for the buoys in B). The parameters and the one-dimensional tests used in performing the random projection test are included for each buoy.

Buoy	Epps (100,1)	Epps (2,7)	L.-V. (100,1)	L.-V. (2,7)	FDR
028	0.89	0.62	0.04	0.01	<b>0.04</b>
029	0.42	0.46	0.61	0.65	0.65
036	0.20	0.26	0.16	0.26	0.26
045	0.98	0.98	0.27	0.26	0.81
067	0.34	0.26	0.15	0.16	0.34
071	0.46	0.53	0.62	0.78	0.78
076	0.21	0.73	0.24	0.62	0.73
094	0.87	0.81	0.27	0.76	0.87
098	0.39	0.56	0.79	0.85	0.85
100	0.62	0.84	1.00	1.00	1.00
134	0.72	0.24	0.76	0.32	0.76
139	0.80	0.56	0.24	0.39	0.80
147	0.54	0.75	0.36	0.55	0.75
150	0.19	0.53	0.48	0.61	0.61
154	0.60	0.50	0.07	0.29	0.30
157	0.19	0.16	0.43	0.52	0.52
158	0.59	0.51	0.69	0.76	0.76
162	0.91	0.98	0.19	0.24	0.72
168	0.73	0.76	0.96	0.87	0.96
185	0.91	0.77	0.60	0.60	0.91
191	0.50	0.47	0.33	0.63	0.63
192	0.69	0.09	0.08	0.08	0.18
196	0.10	0.32	0.36	0.45	0.39
197	0.58	0.07	0.14	0.01	<b>0.04</b>
203	0.93	0.95	0.74	0.72	0.95
213	0.43	0.65	0.42	0.87	0.87
215	0.67	0.50	0.05	0.11	0.19
220	0.60	0.88	0.78	0.35	0.88
222	0.84	0.62	0.59	0.64	0.84
230	0.79	0.83	0.02	0.08	0.08
239	0.22	0.29	0.77	0.79	0.79
240	0.78	0.93	0.76	0.96	0.96
241	0.11	0.18	0.26	0.11	0.26
430	0.21	0.41	0.34	0.63	0.63
433	0.31	0.72	0.14	0.65	0.58

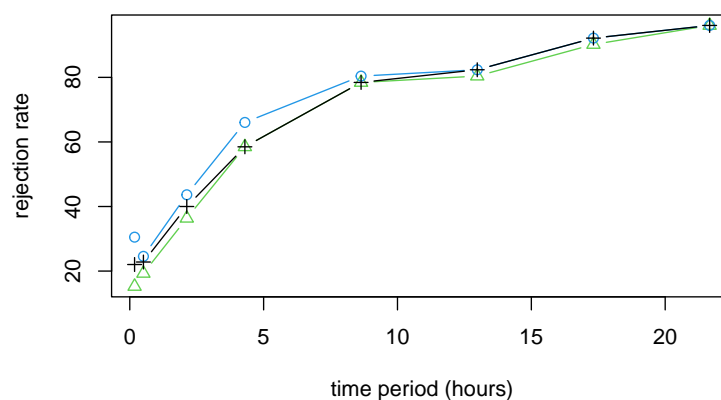
The results in Table 5 have been summarized in the right plot of Figure 4. There, the FDR values larger and smaller than 0.05 can be clearly observed; and that there is a  $p$ -value just above 0.05, the one corresponding to buoy 230.

This procedure that we have exemplified through the use of a time period of 2.13 h in general and of 1.05 h for the buoys in *B*, has been performed for all the time periods displayed in Table 2. This has resulted in a rejection rate displayed in Table 6 with the label *with projection*. Thus, if it is used a time period of 4.34 min for the buoys in *B* and of 10.79 min in general, the Gaussianity of 22.03% of the processes is rejected. This rate increases to 58.49% and 82.35% for time periods of 2.13 h and 4.30 h in general and of 1.05 h and 2.13 h for buoys in *B*, respectively. It reaches 96.08% for a time period of 21.66 h in general and 10.81 h for buoys in *B*.

**Table 6.** Rejection rates along different length periods when no projections are used (first row), when the proposed projection procedure is used (second row) and when no multiplicity is taken into account (third row), i.e., the minimum of the *p*-values is used for the rejection. m stands for minutes and h for hours.

Time Period	General <i>B</i>	10.79 m 4.34 m	30.32 m 14.10 m	2.13 h 1.05 h	4.30 h 2.13 h	8.64 h 4.30 h	12.98 h 6.47 h	17.32 h 8.64 h	21.66 h 10.81 h
no projection		15.25	19.30	36.36	58.49	78.43	80.39	90.20	96.08
with projection		22.03	22.81	40.00	58.49	78.43	82.35	92.16	96.08
minimum		30.51	24.56	43.64	66.04	80.39	82.35	92.16	96.08

If the random projection was not applied, that is, if only the Epps and Lovato and Velasco tests were applied, the rejection rates would have been just below for time periods of 10.79 min, 30.32 min, 2.13 h, 12.98 h and 17.32 h in general and of 4.34 min, 14.10 min, 1.05 h, 6.47 h and 8.64 h for buoys in *B*. The corresponding rejection rates are in Table 6 with the label *no projection*. Meanwhile, if the multiplicity would have not been taken into account altogether and the minimum *p*-value would have been selected, the rejection rates would also have been similar, just a bit over as the time period decreases, separately in general and in *B*. These values are displayed in Table 6 with the label *minimum*. To better illustrate all these rejections, we have plotted them in Figure 5. There, the rejections rates of the proposed procedure are plotted with a back plus symbol, those obtained without using the random procedure are plotted with green triangles and the ones where the minimum of the *p*-values is computed are plotted with blue circles. It can be observed from the figure that the three values get closer as the time period increases.



**Figure 5.** Display of the rejection rates in Table 6: the values labelled by *no projection* are in green triangles, those by *with projection* in back plus symbols and those by *minimum* with blue circles. The time period in the x-axis corresponds to the buoys in general. For the buoys in *B*, it is 4.34 min, 14.10 min, 1.05 h, 2.13 h, 4.30 h, 6.47 h, 8.64 h and 10.81 h.

## 5. Conclusions

For the analysis, the sea surface elevations provided by the measurements of the buoys along the US coast have been analyzed along different time periods. It has been obtained in the analysis section that the rejection rates increase when the time periods increase. Thus, while less than 25% of the studied processes can be considered as non-Gaussian when the time period consists of less than 11 min, it reaches over the 95% of the cases when the time period is larger than 21 h in general and of 10 h for the buoys in *B*. Thus, it is clear the non-Gaussianity of the sea surface elevation and that only enough information (in terms of the time period) is need for its rejection. If we had only used non-randomized tests, as it is common in the literature, the obtained rejection rates would have been a bit lower, with the larger differences being obtained as the time period decreases. For the obtained results, a more complex and new goodness of fit test has been required. This is that known as random projection test and requires of the selection of a distribution used to draw a vector in which to project the data and a goodness of fit test for the one-dimensional marginal distribution of the process to apply on the projected data. This leads to the conclusion that the sea surface elevations are generally non-Gaussian despite that their one-dimensional distribution is Gaussian in a few of the cases. The obtained results are significant as they indicate two important aspects:

- The sea surface elevations are generally not Gaussian as the time period increases, from time series recorded in just 4 to 10 min to time series recorded in 10 to 21 h.
- In a few of the cases the one-dimensional marginal is Gaussian for small and moderate time periods. For instance, for time series recorded in just 4 to 10 min, in 14 to 30 min and those recorded in 1 to 2 h.

These two facts result in that sea surface elevations cannot be modeled as Gaussian processes but in a few of the cases can be modeled as non-Gaussian processes with one-dimensional Gaussian marginals. Note that this is not a contradiction due to, as explained in Section 3, there a non-Gaussian processes with one-dimensional Gaussian marginals.

In testing the non-Gaussianity, multiple tests are performed and consequently multiplicity has to be taken into account. However, if it was not to be taken into account, the same results would have been obtained when all the buoys are studied for time periods of 6.47 h and over; with the rejection rates getting larger than those based on the FDR rate when the length period decreases.

**Funding:** A.N.-R. was supported by grant MTM2017-86061-C2-2-P funded by MCIN/AEI/10.13039/501100011033 and “ERDF A way of making Europe”.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Coastal Data Information Program at <https://thredds.cdip.ucsd.edu/thredds/catalog/cdip/realtime/catalog.html> (accessed on 20 August 2022).

**Conflicts of Interest:** The author declares no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

FDR	False discovery rate
UTC	Coordinated universal time
iid	independent identically distributed

### Appendix A

**Table A1.** Start and end times for which the measurements have been recorded by each of the 59 buoys. These are GMT times which are reported in the format year-month-day and hours-minutes-seconds. In bold, the buoy whose measurements have been recorded for a smaller period of time.

Buoy	Start Time (GMT)		End Time (GMT)	
	Date (yyyy-mm-dd)	Time	Date (yyyy-mm-dd)	Time
028	2021-04-29	19:00:00	2022-08-20	08:59:57
029	2022-01-26	21:00:00	2022-08-20	08:59:58
036	2022-05-25	23:00:00	2022-08-20	08:59:58
045	2022-03-01	20:00:00	2022-08-20	08:59:58
067	2020-12-02	21:00:00	2022-08-20	08:59:57
071	2022-08-02	19:32:13	2022-08-20	08:59:58
076	2022-06-02	18:00:00	2022-08-20	08:59:58
092	2021-08-04	21:00:00	2022-08-20	08:59:57
094	2020-10-08	21:00:00	2022-08-20	08:59:57
098	2022-02-04	21:00:00	2022-08-20	08:59:58
100	2021-02-19	18:00:00	2022-08-20	08:59:57
106	2022-01-20	23:00:00	2022-08-20	08:59:58
121	2021-06-21	13:32:13	2022-04-10	22:59:58
132	2020-09-15	22:30:00	2022-03-01	01:22:48
134	2022-01-15	15:00:00	2022-08-20	08:59:58
139	2021-11-14	20:00:00	2022-08-20	08:59:58
142	2022-03-02	15:00:00	2022-05-20	08:22:49
143	2022-08-16	23:00:00	2022-08-20	08:59:58
144	2021-11-11	17:00:00	2022-08-20	09:29:58
147	2022-03-21	15:00:00	2022-08-20	08:59:58
150	2021-11-18	15:00:00	2022-08-20	08:59:58
154	2021-01-12	17:00:00	2022-06-27	21:59:57
157	2020-06-20	20:00:00	2022-08-20	08:59:57
158	2022-07-06	02:02:13	2022-08-20	08:59:58
160	2022-05-02	21:02:13	2022-08-20	08:59:58
162	2021-11-20	18:00:00	2022-08-20	08:59:58
166	2022-05-16	21:00:00	2022-05-18	03:29:58
168	2022-03-25	20:00:00	2022-08-20	08:59:58
171	2022-01-14	16:00:00	2022-05-12	07:52:49
181	2022-08-11	21:02:13	2022-08-20	08:59:58
185	2022-08-06	02:00:00	2022-08-20	09:29:58
188	2021-11-03	10:23:10	2022-08-20	09:20:55
189	2021-10-20	20:02:13	2022-08-20	08:59:58
191	2021-12-13	20:00:00	2022-08-20	08:59:58
192	2022-08-04	17:00:00	2022-08-20	08:59:58
194	2021-11-05	15:00:00	2022-02-06	23:52:49
196	2022-07-19	04:02:13	2022-08-20	09:29:58
197	2022-07-19	04:02:13	2022-08-20	08:59:58
198	2021-06-16	21:00:00	2022-08-20	08:59:57
201	2021-11-10	21:00:00	2022-08-20	08:59:58
202	2021-07-21	08:02:13	2022-05-04	23:59:58
203	2021-05-12	19:00:00	2022-08-20	08:59:57
204	2021-08-20	23:00:00	2022-05-10	02:52:49
209	2022-06-10	18:32:13	2022-08-20	08:59:58
213	2022-03-15	17:00:00	2022-08-20	08:59:58
214	2022-06-28	04:02:13	2022-08-20	08:59:58
215	2022-04-14	19:00:00	2022-08-20	08:59:58
217	2022-08-08	23:02:13	2022-08-20	08:59:58
220	2022-04-26	22:00:00	2022-08-20	08:59:58
222	2021-11-15	15:00:00	2022-08-20	08:59:58
224	2020-07-16	15:00:00	2022-08-20	08:59:57
230	2022-05-19	17:00:00	2022-08-20	08:59:58
239	2022-07-17	19:05:10	2022-08-20	09:03:56
240	2022-07-22	19:02:13	2022-08-20	08:59:58
241	2022-08-02	00:00:00	2022-08-20	08:29:58
243	2021-09-13	20:00:00	2022-08-20	08:59:57
<b>244</b>	<b>2022-03-31</b>	<b>04:30:00</b>	<b>2022-03-31</b>	<b>10:12:49</b>
430	2022-07-15	18:02:13	2022-08-20	08:59:58
433	2022-08-05	17:00:00	2022-08-20	08:59:58



**Table A2.** The 59 available *buoys* are labelled by an identification number in descending order. The *length* of the associated time series is also reported, the smallest being highlighted in bold. The *studied* label represents the length of the studied time series after selecting the first  $5 \times 10^5$  time points and eliminating the unobserved, missing, values.

Buoy	Length	Studied	Buoy	Length	Studied
028	52,817,065	412,278	185	1,583,018	216,438
029	22,722,218	423,968	<b>188</b>	32,062,463	<b>2304</b>
036	9,557,162	458,528	189	33,569,279	6912
045	18,966,698	488,480	191	27,592,874	419,360
067	69,170,857	334,112	192	1,732,778	451,754
071	1,942,272	394,016	194	20,652,288	162,047
076	8,695,466	317,984	196	3,564,288	396,288
<b>092</b>	42,080,425	<b>500,000</b>	197	3,561,984	463,136
094	75,258,025	412,448	198	47,499,433	306,294
098	21,726,890	444,704	201	31,237,802	476,960
100	60,447,913	453,920	<b>202</b>	31,813,631	<b>2304</b>
106	23,376,554	145,322	203	51,379,369	403,232
121	32,447,231	29,952	204	57,986,303	6912
132	117,484,797	14,592	209	7,808,256	426,272
134	23,966,378	474,656	213	17,432,234	446,838
139	30,800,042	347,766	214	5,884,416	301,856
142	17,403,648	10,752	215	14,109,866	481,855
143	378,026	336,554	217	1,262,592	403,232
144	31,147,946	447,008	220	12,768,938	449,142
147	16,777,898	476,960	222	30,712,490	430,880
150	30,380,714	412,448	224	84,575,400	449,312
154	58,742,953	479,264	230	10,248,362	481,568
157	87,427,752	467,574	239	3,714,125	352,544
158	5,008,896	479,264	240	3,161,088	375,584
160	12,109,824	467,744	241	2,029,994	456,224
162	30,145,706	435,488	243	37,661,353	467,744
166	140,714	126,890	<b>244</b>	<b>52,992</b>	29,952
168	16,312,490	453,920	430	3,939,840	426,272
171	26,015,999	205,088	433	1,622,186	410,144
181	940,032	493,088			

**Table A3.** Bandwidth, per buoy, rounded to 2 decimal values used for the kernel density estimated curves in Figure 3.

<b>buoy</b>	028	029	036	045	067	071	076	092	094	098	100
<b>bandwidth</b>	0.04	0.12	0.05	0.04	0.06	0.08	0.07	0.04	0.06	0.08	0.04
<b>buoy</b>	106	121	132	134	139	142	143	144	147	150	154
<b>bandwidth</b>	0.1	0.14	0.08	0.03	0.09	0.13	0.02	0.03	0.01	0.03	0.04
<b>buoy</b>	157	158	160	162	166	168	171	181	185	191	192
<b>bandwidth</b>	0.04	0.01	0.03	0.05	0.07	0.05	0.11	0.02	0.06	0.04	0.03
<b>buoy</b>	194	196	197	198	201	203	209	213	214	215	217
<b>bandwidth</b>	0.13	0.04	0.04	0.05	0.04	0.04	0.04	0.03	0.02	0.04	0.04
<b>buoy</b>	220	222	224	230	239	240	241	243	244	430	433
<b>bandwidth</b>	0.08	0.07	0.04	0.01	0.13	0.01	0.01	0.02	0.13	0.03	0.02

**References**

- Forristall, G.Z. On the statistical distribution of wave heights in a storm. *J. Geophys. Res. Ocean.* **1978**, *83*, 2353–2358. [CrossRef]
- Azaïs, J.M.; León, J.R.; Ortega, J. Geometrical characteristics of Gaussian sea waves. *J. Appl. Probab.* **2005**, *42*, 407–425. [CrossRef]
- Karmpadakis, I.; Swan, C.; Christou, M. Assessment of wave height distributions using an extensive field database. *Coast. Eng.* **2020**, *157*, 103630. [CrossRef]
- Haver, S. On the joint distribution of heights and periods of sea waves. *Ocean. Eng.* **1987**, *14*, 359–376. [CrossRef]
- Mendes, S.; Scotti, A. The Rayleigh-Haring-Tayfun distribution of wave heights in deep water. *Appl. Ocean. Res.* **2021**, *113*, 102739. [CrossRef]
- Tayfun, M.A. Distribution of Large Wave Heights. *J. Waterw. Port Coastal Ocean. Eng.* **1990**, *116*, 686–707. [CrossRef]
- Mori, N.; Liu, P.C.; Yasuda, T. Analysis of freak wave measurements in the Sea of Japan. *Ocean. Eng.* **2002**, *29*, 1399–1414. [CrossRef]
- Stansell, P. Distributions of freak wave heights measured in the North Sea. *Appl. Ocean. Res.* **2004**, *26*, 35–48. [CrossRef]
- Stansell, P. Distributions of extreme wave, crest and trough heights measured in the North Sea. *Ocean. Eng.* **2005**, *32*, 1015–1036. [CrossRef]
- Casas-Prat, M.; Holthuijsen, L. Short-term statistics of waves observed in deep water. *J. Geophys. Res. Ocean.* **2010**, *115*. [CrossRef]
- Longuet-Higgins, M.S. On the distribution of the heights of sea waves: Some effects of nonlinearity and finite band width. *J. Geophys. Res. Ocean.* **1980**, *85*, 1519–1523. [CrossRef]

12. Jishad, M.; Yadhunath, E.; Seelam, J.K. Wave height distribution in unsaturated surf zones. *Reg. Stud. Mar. Sci.* **2021**, *44*, 101708. [[CrossRef](#)]
13. Muraleedharan, G.; Rao, A.; Kurup, P.; Nair, N.U.; Sinha, M. Modified Weibull distribution for maximum and significant wave height simulation and prediction. *Coast. Eng.* **2007**, *54*, 630–638. [[CrossRef](#)]
14. Naess, A. On the distribution of crest to trough wave heights. *Ocean. Eng.* **1985**, *12*, 221–234. [[CrossRef](#)]
15. Boccotti, P. On Mechanics of Irregular Gravity Waves. *Atti Accad. Naz. Lincei Mem.* **1989**, *19*, 110–170.
16. Klopman, G. *Extreme Wave Heights in Shallow Water*; Report H2486; WL | Delft Hydraulics: Delft, The Netherlands, 1996.
17. van Vledder, G.P. *Modification of the Glukhovskiy Distribution*; Technical Report H1203; Delft Hydraulics: Delft, The Netherlands, 1991.
18. Battjes, J.A.; Groenendijk, H.W. Wave height distributions on shallow foreshores. *Coast. Eng.* **2000**, *40*, 161–182. [[CrossRef](#)]
19. Mendez, F.J.; Losada, I.J.; Medina, R. Transformation model of wave height distribution on planar beaches. *Coast. Eng.* **2004**, *50*, 97–115. [[CrossRef](#)]
20. Wu, Y.; Randell, D.; Christou, M.; Ewans, K.; Jonathan, P. On the distribution of wave height in shallow water. *Coast. Eng.* **2016**, *111*, 39–49. [[CrossRef](#)]
21. Nieto-Reyes, A. On the Non-Gaussianity of the Height of Sea Waves. *J. Mar. Sci. Eng.* **2021**, *9*, 1446. [[CrossRef](#)]
22. Srokosz, M. On the joint distribution of surface elevation and slopes for a nonlinear random sea, with an application to radar altimetry. *J. Geophys. Res. Ocean.* **1986**, *91*, 995–1006. [[CrossRef](#)]
23. Srokosz, M.; Longuet-Higgins, M. On the skewness of sea-surface elevation. *J. Fluid Mech.* **1986**, *164*, 487–497. [[CrossRef](#)]
24. Hokimoto, T.; Shimizu, K. A non-homogeneous hidden Markov model for predicting the distribution of sea surface elevation. *J. Appl. Stat.* **2014**, *41*, 294–319. [[CrossRef](#)]
25. Pena-Sanchez, Y.; Mérigaud, A.; Ringwood, J.V. Short-term forecasting of sea surface elevation for wave energy applications: The autoregressive model revisited. *IEEE J. Ocean. Eng.* **2018**, *45*, 462–471. [[CrossRef](#)]
26. Schulz-Stellenfleth, J.; Lehner, S. Measurement of 2-D sea surface elevation fields using complex synthetic aperture radar data. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1149–1160. [[CrossRef](#)]
27. Collins, C.O.; Lund, B.; Waseda, T.; Graber, H.C. On recording sea surface elevation with accelerometer buoys: Lessons from ITOP (2010). *Ocean. Dyn.* **2014**, *64*, 895–904. [[CrossRef](#)]
28. Reichert, K.; Dannenberg, J.; van den Boom, H. X-Band radar derived sea surface elevation maps as input to ship motion forecasting. In Proceedings of the IEEE OCEANS’10, Sydney, NSW, Australia, 24–27 May 2010; pp. 1–7.
29. Hessner, K.; Reichert, K.; Hutt, B.L. Sea surface elevation maps obtained with a nautical X-Band radar—Examples from WaMoS II stations. In Proceedings of the 10th International Workshop on Wave Hindcasting and Forecasting and Coastal Hazard Symposium, North Shore, Oahu, HI, USA, 11–16 November 2007; pp. 11–16.
30. Cherneva, Z.; Guedes Soares, C. Non-linearity and non-stationarity of the New Year abnormal wave. *Appl. Ocean. Res.* **2008**, *30*, 215–220. [[CrossRef](#)]
31. Rodriguez, G.; Soares, C.G.; Pacheco, M.; Pe´rez-Martell, E. Wave Height Distribution in Mixed Sea States. *J. Offshore Mech. Arct. Eng.* **2001**, *124*, 34–40. [[CrossRef](#)]
32. Tayfun, M.A. Narrow-band nonlinear sea waves. *J. Geophys. Res. Ocean.* **1980**, *85*, 1548–1552. [[CrossRef](#)]
33. Petrova, P.G.; Aziz Tayfun, M.; Guedes Soares, C. The Effect of Third-Order Nonlinearities on the Statistical Distributions of Wave Heights, Crests and Troughs in Bimodal Crossing Seas. *J. Offshore Mech. Arct. Eng.* **2013**, *135*, 021801. [[CrossRef](#)]
34. Nieto-Reyes, A.; Cuesta-Albertos, J.A.; Gamboa, F. A Random-Projection Based Test of Gaussianity for Stationary Processes. *Comput. Stat. Data Anal.* **2014**, *75*, 124–141. [[CrossRef](#)]
35. Epps, T.W. Testing That a Stationary Time Series is Gaussian. *Ann. Stat.* **1987**, *15*, 1683–1698. [[CrossRef](#)]
36. Lobato, I.; Velasco, C. A simple Test of Normality for Time Series. *Econom. Theory* **2004**, *20*, 671–689. [[CrossRef](#)]
37. Benetazzo, A.; Barbariol, F.; Bergamasco, F.; Torsello, A.; Carniel, S.; Sclavo, M. Observation of Extreme Sea Waves in a Space–Time Ensemble. *J. Phys. Oceanogr.* **2015**, *45*, 2261–2275. [[CrossRef](#)]
38. Coleman, R. What is a Stochastic Process? In *Stochastic Processes*; Springer: Dordrecht, The Netherlands, 1974; pp. 1–5. [[CrossRef](#)]
39. Rozanov, Y.A. *Stationary Random Processes*; Holden-Day: San Francisco, CA, USA, 1967.
40. Kozachenko, Y.; Pogorilyak, O.; Rozora, I.; Tegza, A. 2-Simulation of Stochastic Processes Presented in the Form of Series. In *Simulation of Stochastic Processes with Given Accuracy and Reliability*; Kozachenko, Y., Pogorilyak, O., Rozora, I., Tegza, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 71–104. [[CrossRef](#)]
41. Box, G.; Pierce, D.A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *J. Am. Stat. Assoc.* **1970**, *65*, 1509–1526. [[CrossRef](#)]
42. Said, S.E.; Dickey, D.A. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika* **1984**, *71*, 599–607. [[CrossRef](#)]
43. Perron, P. Trends and Random Walks in Macroeconomic Time Series: Further Evidence From a New Approach. *J. Econ. Dyn. Control.* **1988**, *12*, 297–332. [[CrossRef](#)]
44. Kwiatkowski, D.; Phillips, P.C.; Schmidt, P.; Shin, Y. Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How sure Are We that Economic Time Series Have a Unit Root? *J. Econom.* **1992**, *54*, 159–178. [[CrossRef](#)]
45. D’Agostino, R.B.; Stephens, M.A. Goodness-of-fit techniques. *Qual. Reliab. Eng. Int.* **1986**, *3*, 71. [[CrossRef](#)]
46. Cuesta-Albertos, J.; del Barrio, E.; Fraiman, R.; Matrán, C. The Random Projection Method in Goodness of Fit for Functional Data. *Comput. Stat. Data Anal.* **2007**, *51*, 4814–4831. [[CrossRef](#)]

47. Pitman, J. Combinatorial Stochastic Processes. In *Lectures from the 32nd Summer School on Probability Theory Held in Saint-Flour*; Springer: Berlin/Heidelberg, Germany, 2006.
48. Benjamini, Y.; Yekutieli, D. The Control of the False Discovery Rate in Multiple Testing under Dependency. *Ann. Stat.* **2001**, *29*, 1165–1188. [[CrossRef](#)]
49. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. Methodol.* **1995**, *57*, 289–300. [[CrossRef](#)]
50. Rudemo, M. Empirical Choice of Histograms and Kernel Density Estimators. *Scand. J. Stat.* **1982**, *9*, 65–78.
51. Bowman, A.W. An Alternative Method of Cross-Validation for the Smoothing of Density Estimates. *Biometrika* **1984**, *71*, 353–360. [[CrossRef](#)]