Journal of
Marine Science and Engineering

MDPI

Article

On the Non-Gaussianity of the Height of Sea Waves

Alicia Nieto-Reyes †

Department of Mathematics, Statistics and Computer Science, University of Cantabria, 39005 Santander, Spain; alicia.nieto@unican.es
† Current address: Faculty of Science, University of Cantabria, Avd. Los Castros s/n, 39005 Santander, Spain.

Abstract: The objective of this paper is to prove that the sea wave height is not a Gaussian process. This is contrary to the common belief, as the height of a sea wave is generally considered a Gaussian process. With this aim in mind, an empirical study of the buoys along the US coast at a random day is pursued. The analysis differs from those in the literature in that we study the Gaussianity of the process as a whole and not just of its one-dimensional marginal. This is done by making use of random projections and a variety of tests that are powerful against different types of alternatives. The study has resulted in a rejection of the Gaussianity in over 96% of the studied cases.

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

check for updates

Citation: Nieto-Reyes, A. On the Non-Gaussianity of the Height of Sea Waves. *J. Mar. Sci. Eng.* **2021**, *9*, 1446. https://doi.org/10.3390/jmse9121446

Received: 4 October 2021
Accepted: 15 December 2021
Published: 17 December 2021

1. Introduction

The height of sea waves has been broadly studied in the literature, not only from an engineering and statistical perspective [1] but also a probabilistic one [2]. In fact, as commented in [3], the distribution of wave height (sea states) has received significant interest over the years [4–8], which has resulted in the proliferation of models and the analysis of their accuracy [9,10]. The interest also lies in that these models allow one to simulate wave heights [11]. The distribution of the sea wave height is generally considered a strictly stationary Gaussian process when measured with respect to a particular spatial point [12], for instance, the landmark of a buoy. Despite the stationarity being commonly rejected by most tests [13,14] when the associated time series is recorded for long periods of time, the stationary Gaussian model is the most common in the literature [15]. This is in part due to this Gaussian structure allowing for interpretable models of the sea surface that result in the crest height following a particular type of Gaussian distribution: The Rayleigh distribution [16,17]. Sometimes in a more general case, the Weibull distribution is also considered [18]. Other well-known models are based on Forristall distribution [1], Naess distribution [19], Boccotti distribution [20], Klopman distribution [21], van Vledder distribution [22], Battjes–Groenendijk distribution [23], Mendez distribution [24], and LoWiSh II distribution [25]. Having this knowledge on the distribution of sea waves is necessary for a reliable design and analysis of ships and marine installations [26,27]; for instance, offshore structures [26], such as oil platforms or wind turbines.

Although it is considered that certain Gaussian models are appropriate for the distribution of sea wave height when the resulting crest heights are high or the seas have a narrow band spectrum, it is already known that this is not the appropriate model in some cases. In particular, it is not appropriate in cases in which there are steep waves in deep waters or as the water depth decreases [28]. This has already being studied in the literature and, for instance, for non-Gaussian cases [29,30] a methodology to estimate the wave crest height distribution has been developed. The hypothesis of this work is that sea wave height is mainly not Gaussian. Thus, this work goes beyond the existing literature and has the objective to empirically prove that the distribution of the sea wave height is not necessarily Gaussian. The findings are important as the cases that were considered

Gaussian are the more numerous ones. In fact, according to [31], very large waves might be much more frequent than commonly assumed. The main scientific relevance of the study is in that it implies that Gaussian models should not be used in general when modeling the height of sea waves.

In proving the non-Gaussianity, it is here demonstrated that the non-Gaussian cases correspond to non-Gaussian processes with non-Gaussian one-dimensional marginals. This would go in line with the observations in [32] where the authors propose a model that is able to explain the first order probability structure of the process. It occurs that the tests commonly used in the literature have no power against alternatives with one-dimensional Gaussian marginals. These tests are based on the process' characteristic function [33], on its skewness and kurtosis [34,35], or on Stein's characterization of a Gaussian distribution [36], for instance. However, the proposal here is to apply a methodology that is able to check the Gaussianity of the process as a whole. Thus, the Gaussianity hypothesis can be discarded for the heights of the sea waves either if the non-Gaussianity comes from the one-dimensional marginal or from a higher order marginal. The test applied here is known as the random projection test [37] and consists in applying tests that check the one-dimensional Gaussian marginal distribution of the process to a series of processes the result of performing random projections to the original process.

The structure of the paper is as follows. Section 2 is dedicated to present the studied real dataset and Section 4 to analyze it. The methodology applied there is summarized in Section 3. The conclusions are provided in Section 5.

2. Dataset

The studied dataset consists of the sea wave heights measured by buoys run by the Coastal Data Information Program (https://cdip.ucsd.edu, (accessed on 4 October 2021)). The data at this website is freely available and contains raw measurements of surface elevations of buoys located alongside the US coasts. That is, of measurements that have not been preprocessed. To obtain the sea state, wave height, the dataset is here preprocessed by means of the zero-down crossing methodology. Particularly, the study focuses on the variable xyzZDisplacement that is downloaded from the web page https://thredds.cdip.ucsd.edu/thredds/catalog/cdip/realtime/catalog.html, (accessed on 4 October 2021). The data available to be downloaded depends on the date.

The data studied in this manuscript was downloaded on 14 November 2021 and is available from the author upon request. In the downloaded dataset, there were 64 available stations (buoys), which are labeled by an identification number that can be observed, in ascending order, in Table 1. Note that the studied stations have been selected because of their availability and not because they satisfy certain conditions in relation to their location relative to the shoreline, the sea depth, or the prevailing wind directions. However, in case the reader is interested in this information, Table 1 includes information on the location of the 64 buoys used, latitude, longitude, and depth. To provide an example of wind directions, Figure 1 shows the wind direction in degrees (local magnetic variation (deg): 12 E) for Station 073, Scripps Pier, La Jolla CA, on 21 January 2021 for an hour from 16:20 to 17:19 UTC, with a total of 3600 measurements. The range of degree value registered is 0.7–359.7.

Table 2 displays the length of the raw time series associated to each of the 64 buoys. The dataset understudy is here restricted to a time series of length of 500,000 for each of the 54 buoys with a length larger than that. As commented above, the downloaded dataset consists of un-preprocessed raw data, a consequence being that it contains unknown values. After removing them, the resulting time series length recorded by each station is also displayed in Table 2, with the label *studied*. It is worth commenting that in buoy 244, the whole 500,000 first recorded time points have unobserved data. Thus, this buoy is not kept in the study. This is represented in Table 2 by a line in the cell corresponding to the label *studied*. On the contrary, there is a bouy for which the 500,000 first recorded time points fully consists of observed data which is buoy 92.

Table 1. Identification number (first column), name (second column), depth (third column), latitude (fourth column), and longitude (fifth column) of the 64 *buoys* included in this study.

Buoy	Name	Depth (Feet)	Latitude (ISO 6709)	Longitude (ISO 6709)
28	SANTA MONICA BAY, CA	1270	33.859933°	−118.641100°
29	POINT REYES, CA	1804	37.936675°	−123.462920°
36	GRAYS HARBOR, WA	135	46.856850°	−124.244150°
45	OCEANSIDE OFFSHORE, CA	781	33.177900°	−117.472167°
67	SAN NICOLAS ISLAND, CA	859	33.219278°	−119.872278°
71	HARVEST, CA	1830	34.451650°	−120.779817°
76	DIABLO CANYON, CA	90	35.203815°	−120.859314°
92	SAN PEDRO, CA	1563	33.617933°	−118.316833°
94	CAPE MENDOCINO, CA	1132	40.294870°	−124.731770°
100	TORREY PINES OUTER, CA	1876	32.933000°	−117.390733°
101	TORREY PINES INNER, CA	103	32.925630°	−117.276810°
106	WAIMEA BAY, HI	656	21.670483°	−158.117217°
121	IPAN, GUAM	656	13.354167°	144.788330°
132	FERNANDINA BEACH, FL	51	30.709040°	−81.292080°
134	FORT PIERCE, FL	54	27.551450°	−80.217033°
142	SAN FRANCISCO BAR, CA	56	37.787500°	−122.633100°
143	CAPE CANAVERAL NEARSHORE, FL	32	28.400200°	−80.533450°
144	ST. PETERSBURG OFFSHORE, FL	308	27.344600°	−84.274800°
147	CAPE HENRY, VA	49	36.915000°	−75.722000°
150	MASONBORO INLET, ILM2, NC	52	34.141900°	−77.715045°
153	DEL MAR NEARSHORE, CA	56	32.956583°	−117.279450°
154	BLOCK ISLAND, RI	167	40.967317°	−71.126550°
155	IMPERIAL BEACH NEARSHORE, CA	68	32.569567°	−117.168800°
157	POINT SUR, CA	1210	36.334767°	−122.103900°
158	CABRILLO POINT NEARSHORE, CA	58	36.626300°	−121.907050°
160	JEFFREYS LEDGE, NH	262	42.800000°	−70.170800°
168	HUMBOLDT BAY NORTH SPIT, CA	361	40.896033°	−124.357000°
171	VIRGINIA BEACH OFFSHORE, VA	161	36.611000°	−74.841330°
179	ASTORIA CANYON, OR	595	46.133283°	−124.644450°
181	RINCON, PUERTO RICO	108	18.376580°	−67.279650°
185	MONTEREY BAY WEST, CA	4799	36.700000°	−122.342580°
187	PAUWELA, MAUI, HI	656	21.018567°	−156.421750°
188	HILO, HAWAII, HI	1115	19.780000°	−154.970000°
189	AUNUU, AMERICAN SAMOA	180	−14.273200°	−170.500500°
191	POINT LOMA SOUTH, CA	3444	32.516700°	−117.425200°
192	OREGON INLET, NC	60	35.750350°	−75.330002°
194	ST. AUGUSTINE, FL	77	29.999860°	−81.079960°
198	KANEOHE BAY, HI	266	21.477470°	−157.752620°
200	WILMINGTON HARBOR, NC	42	33.722050°	−78.016420°
201	SCRIPPS NEARSHORE, CA	151	32.868000°	−117.266600°
202	HANALEI, KAUAI, HI	656	22.284717°	−159.574217°
203	SANTA CRUZ BASIN, CA	6200	33.769000°	−119.564700°
204	LOWER COOK INLET, AK	112	59.597500°	−151.829100°
209	BARNEGAT, NJ	84	39.768290°	-73.770370°
213	SAN PEDRO SOUTH, CA	217	33.577667°	−118.182033°
214	EGMONT CHANNEL ENTRANCE, FL	46	27.590300°	−82.931300°
215	LONG BEACH CHANNEL, CA	76	33.700333°	−118.200668°
217	ONSLOW BAY OUTER, NC	98	34.212550°	−76.949000°
220	MISSION BAY WEST, CA	1931	32.751580°	−117.500750°
221	CAPE COD BAY, MA	85	41.840100°	−70.328700°
222	SANTA LUCIA ESCARPMENT, CA	2132	34.767500°	−121.498000°
224	WALLOPS ISLAND, VA	54	37.754166°	−75.325000°
226	PULLEY RIDGE, FL	266	25.700633°	−83.650133°
238	BARBERS POINT, KALAELOA, HI	919	21.323080°	−158.149480°
243	NAGS HEAD, NC	69	36.001330°	−75.420980°
244	SATAN SHOAL, FL	325	24.407166°	−81.966833°
248	ANGELES POINT, WA	265	48.173183°	−123.605217°
249	ARECIBO, PR	105	18.490850°	−66.700517°
250	CAPE HATTERAS EAST, NC	85	35.259250°	−75.286100°
254	POINT SANTA CRUZ, CA	66	36.934397°	−122.033891°
255	TRINITY SHOAL, LA	69	29.086800°	−92.506400°
256	SOUTHWEST PASS ENTRANCE W, LA	147	28.988010°	−89.649270°
430	DUCK FRF 26M, NC	82	36.258808°	−75.592207°
433	DUCK FRF 17m, NC	58	36.199700°	−75.714117°

In Panel A of Figure 2, the surface elevation data for buoy 28 is displayed, which is to be preprocessed to obtain the heights of sea waves. As commented before, this data results from restricting the 21,961,898 observations stored for buoy 28 and taking the ones corresponding to the first 500,000 time points. As it is observable from the plot, the first 500,000 time points contain unobserved data. In fact, as displayed in Table 2, only 412,278 observations have been made. From Panel A of Figure 2, it is noticeable that the unobserved data splits the time series in several clusters. The first one has been reproduced solely in Panel B of Figure 2. The time points reproduced in Figure 2 are measured in coordinated universal times (UTC) and are reported in seconds. As reported in Table 3, the measurements of surface elevation for buoy 28 began recording at time 1,619,719,067, which is the Thursday, 29 April 2021 at 17:00 h 57 min and 47 s in Greenwich mean time (GMT). The last studied measurement of that buoy was recorded at time 1,620,109,691, which is Tuesday, 4 May 2021 at 06:00 h 28 min and 11 s in GMT. The x-axis of the plots in Figure 2 have been computed as:

$$T_t = T_0 + (t-1)/r - d \text{ for } t = 1, \ldots, 500,000$$

where T_t denotes the t-th time point, with T_0 the starting time, r the time sample rate, and d the filter delay. T_0 is provided for each buoy in Table 3. r takes value 1.28 and d takes value 133.3 for all the studied buoys except for buoys:

$$132, 142, 153, 171, 194, 200, 204, 244, 248, 254, 256$$

for which r takes value 2.56 and d value 130. Additionally, for buoys 101 and 155 takes a d value equal to 299.

Table 2. The 64 available *buoys* are labeled by a number, displayed in ascending order. It is also presented the *length* of the associated raw time series. The smallest length value is depicted in bold. Under the *studied* label, it is the length of the raw time series under study after restricting it to the first 500,000 time points, in the 54 cases that the length is larger than this value, and eliminating the unobserved values.

Buoy	Length	Studied	Buoy	Length	Studied
28	21,961,898	412,278	188	**2304**	2304
29	82,994,856	467,744	189	9216	6912
36	9,898,154	428,576	191	1705,130	460,831
45	47,402,495	421664	192	57,802,921	479,264
67	38,315,689	334,112	194	331,008	162,047
71	97,159,848	447,008	198	16,644,266	306,294
76	45,186,217	426,410	200	36,864	13,056
92	11,225,258	500,000	201	382,634	366,506
94	44,405,161	412,448	202	12,713,472	2304
100	29,592,746	453,920	203	20,524,202	403,232
101	392,063	276,863	204	18,523,392	6912
106	27,261,324	251,533	209	48,176,809	486,176
121	36,864	29,952	213	23,726,762	474,656
132	92,626,942	14,592	214	9,469,610	410,144
134	184,490	177,578	215	71,599,273	483,872
142	41,028,863	3072	217	33,912,745	343,158
143	4,170,410	463,136	220	43,743,913	476,960
144	290,474	260,522	221	32,689,322	412,448
147	38,223,529	460,832	222	4,370,858	207,222
150	2,732,714	435,488	224	53,720,233	449,312
153	61,614,437	19,968	226	55,466,665	449,312
154	33,799,849	479,264	238	7,206,912	13,824
155	1,276,799	399,819	243	6,806,186	467,744
157	56,572,585	467,574	244	100,154,110	-
158	23,500,970	467,744	248	13,824	10,752
160	40,979,113	414,752	249	16,777,898	474,656
168	96,282,024	469,194	250	11,446,442	410,144
171	76,003,582	190,464	254	3,026,688	20,768
179	47,052,457	458,358	255	13,676,714	398,624
181	8,713,898	481,568	256	29,211,647	74,496
185	5,580,458	453,920	430	25,175,978	272,042
187	33,458,857	315,680	433	32,809,130	428,714

Table 3 displays the starting and ending recording times of the surface elevation for each of the studied stations, in addition to the ones of station 28. The largest time is 1,636,876,799, which represents Sunday, 14 November 2021 at 07:00 h 59 min and 59 s GMT. It is represented in bold in Table 3. This is the end time for the recording of buoy 101, whose starting time is 1,636,570,501, i.e., Wednesday, 10 November 2021 at 18:00 h 55 min and 1 s GMT. Meanwhile, the smallest starting time point is 1,560,970,667, which represents the Wednesday 19 June 2019 at the 18:00 h 57 min and 47 s GMT. This number has also been highlighted in bold in Table 3. It corresponds to the buoy with identification number

071, whose end time point is 1,561,361,291, i.e., Monday, 24 June 2019 at 07:00 h 28 min and 11 s GMT.

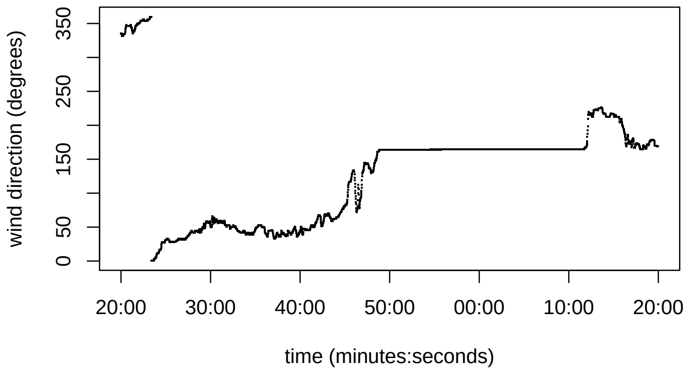

Figure 1. Plot of 3600 wind directions (degrees) with respect to time (minutes:seconds) for the Station 073, Scripps Pier, La Jolla, CA, on 21 January 2021 from 16:20 to 17:19 UTC.

Table 3. Starting and ending time of the recorded surface elevations for each of the studied buoys. The time is measured in seconds in UTC.

Buoy	Start Time (s. UTC)	End Time (s. UTC)	Buoy	Start Time (s. UTC)	End Time (s. UTC)
28	1619719067	1,620,109,691	188	1,635,934,857	1,635,936,656
29	1,572,037,067	1,572,427,691	189	1,634,756,400	1,634,763,599
36	1,629,143,867	1,629,534,491	191	1,635,544,667	1,635,935,291
45	1,599,843,600	1,600,234,224	192	1,591,718,267	1,592,108,891
67	1,606,942,667	1,607,333,291	194	1,636,124,270	1,636,253,570
71	**1,560,970,667**	1,561,361,291	198	1,623,873,467	1,624,264,091
76	1,601,575,067	1,601,965,691	200	1,635,524,870	1,635,539,270
92	1,628,107,067	1,628,497,691	201	1,636,577,867	1,636,876,799
94	1,602,187,067	1,602,577,691	202	1,626,850,800	1,627,241,424
100	1,613,757,467	1,614,148,091	203	1,620,842,267	1,621,232,891
101	1,636,570,501	**1,636,876,799**	204	1,629,496,670	1,629,691,982
106	1,615,579,067	1,615,969,691	209	1,599,238,667	1,599,629,291
121	1,624,278,600	1,624,307,399	213	1,618,340,267	1,618,730,891
132	1,600,205,270	1,600,400,582	214	1,629,478,667	1,629,869,291
134	1,636,732,667	1,636,876,799	215	1,580,939,867	1,581,330,491
142	1,620,435,470	1,620,630,782	217	1,610,384,267	1,610,774,891
143	1,633,618,667	1,634,009,291	220	1,602,701,867	1,603,092,491
144	1,636,649,867	1,636,876,799	221	1,611,338,267	1,611,728,891
147	1,607,014,667	1,607,405,291	222	1,633,463,867	1,633,854,491
150	1,634,741,867	1,635,132,491	224	1,594,907,867	1,595,298,491
153	1,609,953,430	1,610,148,742	226	1,593,543,467	1,593,934,091
154	1,610,470,667	1,610,861,291	238	1,606,899,600	1,607,290,224
155	1,635,879,301	1,636,269,925	243	1,631,559,467	1,631,950,091
157	1,592,679,467	1,593,070,091	244	-	-
158	1,618,516,667	1,618,907,291	248	1,636,534,670	1,636,729,982
160	1,604,861,867	1,605,252,491	249	1,623,769,067	1,623,774,467
168	1,561,658,267	1,562,048,891	250	1,627,934,267	1,628,324,891
171	1,606,739,270	1,606,934,582	254	1,634,927,270	1,635,317,894
179	1,600,117,067	1,600,507,691	255	1,626,191,867	1,626,387,179
181	1,630,069,067	1,630,459,691	256	1,624,062,470	1,624,453,094
185	1,632,517,067	1,632,907,691	430	1,617,209,867	1,617,405,179
187	1,610,737,067	1,611,127,691	433	1,611,244,667	1,611,635,291

Once the dataset is preprocessed, making use of the zero-down crossing methodology to obtain the sea wave heights, the length of each time series decreases, as reported in Table 4. As in Table 2, in Table 4, buoy 244 has a line in place of its length, because of the studied recordings being unobserved. Thus, this buoy is not included in the rest of the analysis. In Figure 3, the sea wave heights of 6 buoys that cover the spectrum of mean heights shown in Table 5 are displayed. It is observable from the plots that they correspond to different sea states. In particular, the plot corresponding to buoy 226 (1st row–1st column) represents wave heights that are generally of less than 1 m while the plots corresponding to buoy 433 and 155 (1st row–2nd column and 2nd row–1st column, respectively) show heights that are generally of less than 3 m and that of buoys 222 and 189 (2nd row–2nd column and 3rd row–1st column) have slightly higher values. Meanwhile, the plot of buoy 106 (3rd row–2nd column) shows much higher values of the sea heights. For the exact values, see Table 5, which represents the range and mean of sea heights and periods for each of the studied 63 buoys. It is observable from these values that the study comprises a variety of sea states. For instance, the maximum mean height recorded is over to 2 m and corresponds to buoy 106. Meanwhile, the time series studied for buoy 226 has a mean value of only 0.17 m. Both values are highlighted in bold in Table 5.

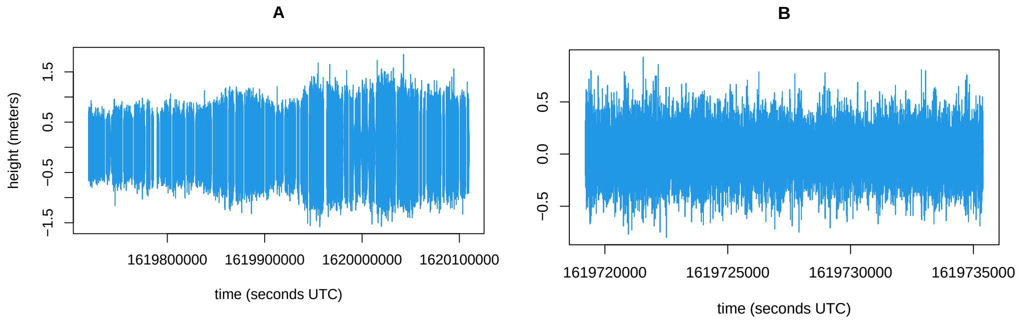

Figure 2. (Panel **A**): Representation of the studied time series for buoy 28. The voids represent unobserved data. (Panel **B**): Representation of the first observed segment of the time series in (Panel **A**). Height in meters and time in seconds UTC.

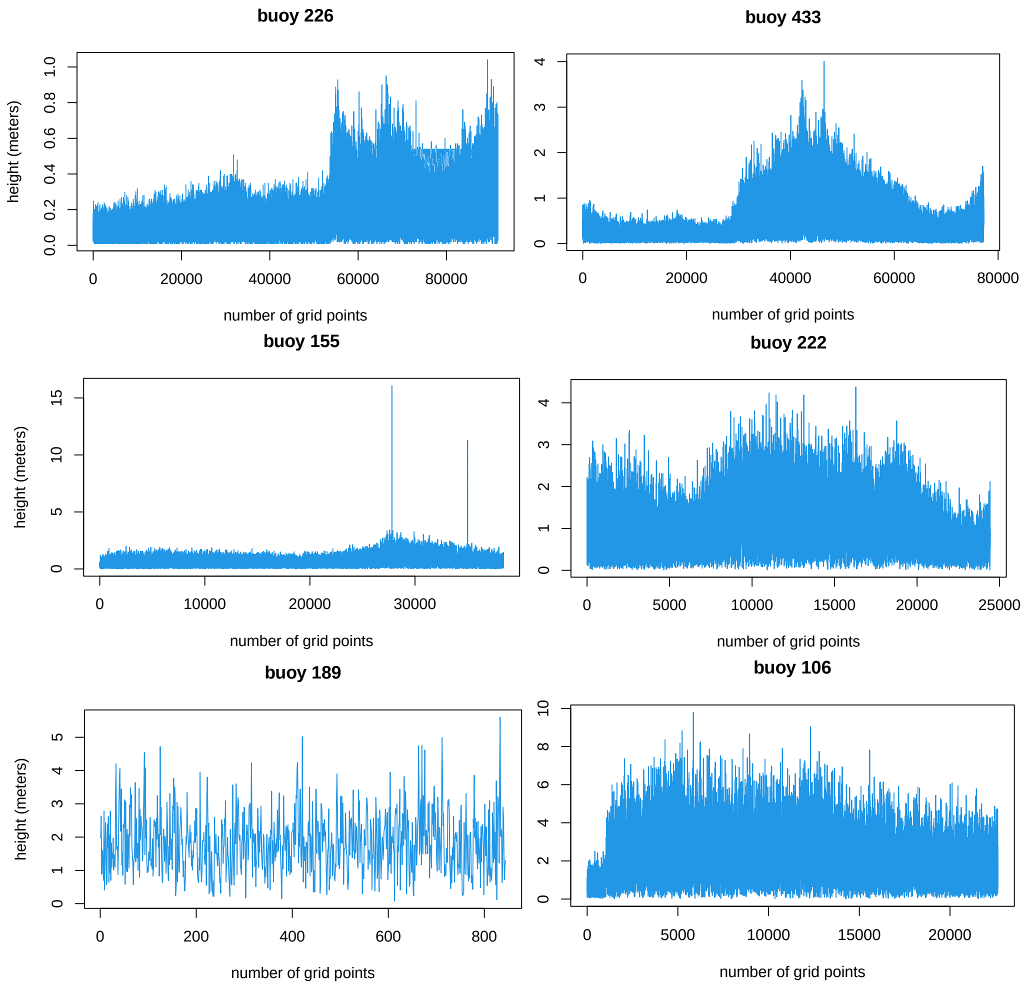

Figure 3. Display of the sea wave height time series for buoys 226, 433, 155, 222, 189, and 106 (from left to right and top to bottom).

Table 4. Length of the 64 studied wave height time series after preprocessing the corresponding surface elevations. The buoy with no associated value corresponds to that in which the surface elevations are unobserved.

Buoy	Length	Buoy	Length	Buoy	Length	Buoy	Length
28	53,055	143	78,727	188	161	220	52,880
29	50,072	144	54,798	189	844	221	85,438
36	50,696	147	88,537	191	51,258	222	24,429
45	63,179	150	92,248	192	96,017	224	82,826
67	30,739	153	927	194	10,578	226	91,653
71	45,184	154	85,352	198	50,000	238	697
76	45,324	155	38,422	200	850	243	87,981
92	69,255	157	56,255	201	48,102	244	-
94	45,071	158	67153	202	274	248	477
100	53,103	160	73,135	203	49,098	249	73,017
101	33,875	168	65,148	204	567	250	64,811
106	22,640	171	11,208	209	81,533	254	927
121	3718	179	57,563	213	71,486	255	87,275
132	1136	181	88,800	214	86,138	256	4787
134	24,806	185	43,133	215	69,636	430	45,559
142	201	187	27,017	217	60,739	433	77,214

Table 5. For each of the 63 buoys of positive length, minimum, maximum, and mean wave height and period of the studied time series.

Buoy	Height			Period		
	Min	Max	Mean	Min	Max	Mean
28	0.01	3.14	0.69	1.56	19.53	6.07
29	0.01	9.50	1.48	1.56	21.88	7.30
36	0.01	4.67	1.21	1.56	17.97	6.60
45	0.01	1.58	0.40	1.56	19.53	5.21
67	0.01	4.54	1.08	1.56	28.91	8.49
71	0.01	4.18	1.06	1.56	18.75	7.73
76	0.01	3.44	0.84	1.56	19.53	7.35
92	0.01	2.27	0.49	1.56	17.97	5.64
94	0.01	5.74	1.11	1.56	19.53	7.15
100	0.01	3.25	0.65	1.56	19.53	6.68
101	0.01	1.86	0.41	1.56	21.09	6.38
106	0.01	9.79	**2.34**	1.56	17.97	8.68
121	0.03	6.39	1.92	1.56	12.50	6.29
132	0.04	4.11	1.32	0.78	11.33	5.01
134	0.01	1.85	0.47	1.56	15.62	5.59
142	0.10	4.56	1.56	0.78	11.72	5.88
143	0.01	1.39	0.33	1.56	18.75	4.60
144	0.01	3.67	0.41	1.56	10.94	3.71
147	0.01	4.41	0.61	1.56	14.84	4.07
150	0.01	1.57	0.34	1.56	17.19	3.69
153	0.01	2.28	0.76	0.78	20.70	8.39
154	0.01	7.99	0.68	1.56	16.41	4.39
155	0.01	16.08	0.69	1.56	24.22	8.13
157	0.01	4.26	0.87	1.56	19.53	6.49
158	0.01	2.12	0.44	1.56	19.53	5.44
160	0.01	2.75	0.47	1.56	17.19	4.43
168	0.01	2.05	0.49	1.56	15.62	5.63
171	0.02	10.24	2.29	0.78	14.06	6.64
179	0.01	4.88	0.87	1.56	15.62	6.22
181	0.01	1.66	0.31	1.56	12.50	4.24
185	0.01	9.93	1.16	1.56	21.88	8.22
187	0.01	8.83	2.23	1.56	27.34	9.13
188	0.09	3.76	1.79	1.56	17.97	10.97
189	0.08	5.60	1.83	1.56	14.06	6.38
191	0.01	40.08	0.74	1.56	19.53	7.02
192	0.01	2.60	0.52	1.56	16.41	3.90
194	0.01	8.16	2.14	0.78	14.06	5.98
198	0.01	2.89	0.65	1.56	14.06	4.79
200	0.01	6.04	1.98	0.78	10.55	5.99
201	0.01	2.06	0.41	1.56	17.97	5.95
202	0.08	4.58	1.77	1.56	11.72	6.52
203	0.01	2.74	0.63	1.56	21.88	6.42
204	0.02	3.98	1.21	0.78	8.98	4.74
209	0.01	2.40	0.46	1.56	13.28	4.66
213	0.01	2.27	0.52	1.56	19.53	5.19
214	0.01	1.10	0.21	1.56	15.62	3.72
215	0.01	1.73	0.32	1.56	21.09	5.43
217	0.01	3.03	0.58	1.56	15.62	4.41
220	0.01	1.94	0.51	1.56	20.31	7.05

Table 5. *Cont.*

Buoy	Height			Period		
	Min	Max	Mean	Min	Max	Mean
221	0.01	4.36	0.59	1.56	10.16	3.77
222	0.01	4.38	1.07	1.56	21.09	6.63
224	0.01	1.94	0.46	1.56	14.06	4.24
226	0.01	1.04	**0.17**	1.56	10.94	3.83
238	0.01	5.15	1.91	1.56	22.66	15.47
243	0.01	4.20	0.50	1.56	13.28	4.15
248	0.05	3.67	1.39	1.17	16.02	8.78
249	0.01	2.87	0.77	1.56	13.28	5.08
250	0.01	4.12	0.83	1.56	11.72	4.94
254	0.02	5.56	1.85	0.78	18.36	8.73
255	0.01	1.58	0.35	1.56	9.38	3.57
256	0.01	7.09	2.04	0.78	13.28	6.08
430	0.01	4.57	0.75	1.56	15.62	4.66
433	0.01	4.01	0.47	1.56	16.41	4.34

3. Methodology

Let $X := \{X_t\}_{t\in\mathbb{Z}}$ be a stochastic process. It is Gaussian if for any $T \subset \mathbb{Z}$ with finite cardinality, $\{X_t\}_{t\in T}$ follows a multivariate normal distribution. Most common procedures to test the Gaussianity of a stochastic process are designed for stationary processes. Thus, in Section 4, the stationarity is checked first. The most commonly used tests for stationarity are the Augmented Dickey–Fuller test [13], the Phillips–Perron test [38], and the Ljung–Box test [39], and so, we apply those. These tests check the null hypothesis of non-stationarity against the alternative of non-stationarity.

Once it can be assumed that X is a stationary process, the aim of this study is to contrast the null hypothesis of X being Gaussian against the alternative that X does not follow a Gaussian process. There are a variety of tests with this aim. However, most of them simply contrast the null hypothesis of the one-dimensional marginal distribution of X being Gaussian against the alternative that this marginal distribution is not Gaussian. Of course, a rejection with those tests will provide a rejection for X being Gaussian. However, these tests are not able to detect non-Gaussian processes with Gaussian one-dimensional marginals. A class of them checks whether the characteristic function of the one-dimensional marginal of the process is that of a Gaussian distribution. From this class, the Epps test is used [33], which compares at certain points the empirical characteristic function of the one-dimensional marginal distribution of the process with that of a Gaussian distribution. Another class is that of the skewness—kurtosis tests, also known as Jarque–Bera tests [40]. These tests check whether the empirical skewness and kurtosis depart from zero, as a Gaussian distribution has zero skewness and kurtosis. From these, the Lobato and Velasco test is used [34].

To test the Gaussianity of the stationary process X completely, and not just of a certain order marginal, the random projection test is used [37]. This test consists of projecting the process X and applying a test for the Gaussianity of the one-dimensional marginal of the projection. The hypotheses required for the application of the random projection test come from the hypotheses required by the test applied to the one-dimensional marginal of the projection. To introduce these hypotheses, the notation

$$\gamma_X(t) := \mathbb{E}[(X_0 - \mathbb{E}[X_0])(X_t - \mathbb{E}[X_0])]$$

is used for any $t \in \mathbb{Z}$. In the case the Epps test is applied, the hypotheses are:

(E1.) X is an ergodic stationary process.
(E2.) The characteristic function of the one-dimensional marginal of X is analytic.
(E3.) $\sum_{t\in\mathbb{Z}} |t|^{\zeta} |\gamma_X(t)| < \infty$, for some $\zeta > 0$.
(E4.) The spectral density matrix of the process:

$$\{(\cos(\lambda_1 X_t), \sin(\lambda_1 X_t), \ldots, \cos(\lambda_N X_t), \sin(\lambda_N X_t))\}_{t\in\mathbb{Z}} \quad (1)$$

at frequency 0 exists and is positive definite. In (1) the $\{\lambda_i\}_{i=1}^N$, with $N > 1$, are drawn at random in such a way that λ_1 and λ_2 are independent and identically distributed with an absolutely continuous distribution.

In the case the Lobato and Velasco test is applied, the hypotheses are:

(L1.) X is an ergodic stationary process.
(L2.) $\sum_{t=0}^{\infty} |\gamma_X(t)| < \infty$.
(L3.) $X_t - \mathbb{E}[X_0] = \sum_{i=1}^{\infty} k(i)\epsilon_{t-i}$, where $\sum_{i=1}^{\infty} |k(i)| < \infty, \sum_{i=1}^{\infty} ik(i) < \infty$ and the $\{\epsilon_t\}_{t\in\mathbb{Z}}$ are independent and identically distributed random variables with $\mathbb{E}[\epsilon_n] = 0$ and $\mathbb{E}[X_0^4] < \infty$.

Hypothesis (E1.) and (L1.) coincide. Note that, although stationarity is a requirement for the tests applied, no regular wave shape is assumed. However, requiring stationarity automatically implies, as commented in the introduction, that the time series recorded for long periods of time cannot be analyzed under these hypothesis tests. A reason behind this can be seen in [41,42], where the significance of the intra-seasonal and intra-annual variability of the wave distribution is studied. Note that X is stationary if and only if $\{X_t\}_{t\in T}$ and $\{X_{t+k}\}_{t\in T}$ are identically distributed for any $k \in \mathbb{Z}$ and $T \subset \mathbb{Z}$, which does not allow the difference in variability seen in those papers when long periods of time are studied. Furthermore, (E3.) and (L2.) are related. In fact, if t takes non-negative values (L2.), it implies that (E3.). (E2.), (E4.), and (L3.) are also connected, in the sense that (E2.) and (E4.) are assumptions related to the fact that the Epps test makes use of characteristic functions while (L3.) relates to the Lobato and Velasco test, which is based on the distribution moments. Thus, it could be said that there are no big difference in the implications derived from either assuming the Epps or Lovato and Velasco set of assumptions.

In Nieto-Reyes et al. [37], it is proposed to make use of a stick-breaking procedure to draw the sequence in which to project the time series. The procedure makes use of a beta distribution. Two sets of parameters are proposed there: (100,1) when the objective is to obtain a projected process similar to the original one and (2,7) when the aim is that the sequence does not contain many zero values. Therefore, if the aim is to apply only one projected test, the proposal here is to make use of the (100,1) parameters when the one-dimensional distribution before being projected is close to being non-Gaussian and of the (2,7) parameters when it is far from it. For the selection of the test to check the one-dimensional distribution of the projected process, the idea is to use the Lobato and Velasco test when the departure from Gaussianity seems to come from a non-zero skewness or kurtosis and to use the Epps test otherwise. Thus, rejecting the null hypothesis has different implications under the four different scenarios:

- Parameters (100,1) with the Lobato and Velasco test: The non-Gaussianity of the process is related to the third and/or fourth order moment of a small dimensional distribution of the original non-projected process.
- Parameters (2,7) with the Lobato and Velasco test: The non-Gaussianity of the process is related to the third and/or fourth order moment of the one-dimensional distribution of the projected process.
- Parameters (100,1) with the Epps test: The non-Gaussianity of the process is related in general to a small dimensional distribution of the original non-projected process.
- Parameters (2,7) with the Epps test: tThe non-Gaussianity of the process is related in general to the one-dimensional distribution of the projected process.

4. Data Results

This section is dedicated to study whether the sea height time series presented and pre-processed in Section 2 have been drawn from Gaussian processes. The analysis is performed using R and, in particular, the nortsTest package, which is aimed at assessing normality of stationary processes. Thus, the normality tests applied in this section are intended for stationary processes. Then, it is first checked that this assumption is satisfied. For this, three different tests are applied: Augmented Dickey–Fuller, Phillips–Perron, and

Ljung–Box. As commented in Section 3, the aim is to reject the null hypothesis. As it is observable from Table 6, which summarizes the obtained results, this is what occurs for each of the 63 studied buoys. In particular, for each of the buoys, a p-value smaller or equal than 0.01 is obtained for the Augmented Dickey–Fuller and the Phillips–Perron tests and smaller than that value for the Ljung–Box test. Note that the null hypothesis of non-stationarity is tested against the alternative hypothesis of stationarity. This leads us to the conclusion that stationarity can be assumed and, therefore, normality tests for stationary processes can be applied to these data.

Table 6. Summary of the obtained p-values when applying, to each of the 62 buoys, three different tests to check for stationarity. Each of the 62 p-values is smaller or equal than 0.01 for the *Augmented Dickey–Fuller* test (first column) and the *Phillips–Perron* test (second column). Each of them is smaller than 0.01 for the *Ljung–Box* test (third column). The null hypothesis is of non-stationarity.

	Augmented Dickey-Fuller	Phillips-Perron	Ljung-Box
p-value	$\leq$0.01	$\leq$0.01	<0.01

To check the Gaussianity of the process associated to each of the different buoys, as well as the Epps and Lobato and Velasco tests are first applied. As commented in Section 3, these tests check the Gaussianity of the one-dimensional marginal distribution of the process. Thus, a rejection with these tests implies the rejection of the Gaussianity of the whole process. The results appear in Table 7. By looking separately to the resulting p-values of the Epps or Lobato and Velasco tests, the Gaussianity of the process associated to 60 out of the 63 buoys could be rejected when making use of the Epps test and of 61 out of 63 when making use of the Lobato and Velasco test. The buoys for which the Epps test does not have enough evidence to reject the null hypothesis of Gaussianity are buoys 142, 202, and 238. Meanwhile, the buoys for which the Lobato and Velasco test does not have enough evidence to reject the null hypothesis are buoy 188 and, again, buoy 238. The corresponding p-values are highlighted in bold in Table 7. Note that in the five buoys, at least one of the two obtained p-values is smaller than 0.05, reflecting that the process does not follow the Guassian distribution in terms of the characteristic function (Epss test) or in terms of the skewness and kurtosis (Lobato and Velasco test).

To be on the safe side, it is here taken into account the multiplicity of having run two tests. For that, the false discovery rate (FDR) is applied for dependent tests [43]. It is clear that in this case, there is dependency as both tests have been applied to the same dataset. Although it could be argued that the FDR for independent tests [44] also allows for some type of dependency and that this might be of that type, the most restrictive FDR is used here to be cautious. The other would have given smaller values in columns 4 and 8 of Table 7. Using the FDR for dependent tests, the Gaussianity of the processes associated to 61 out of the 63 studied buoys is rejected. Note that this would result in a rejection rate of 96.83%. The FDR values larger than 0.05 are highlighted in bold in Table 7.

A further study is pursued, which consists in applying the random projection test. To do so, four p-values are computed per buoy in a way that the time series is projected two times using the parameters (2,7) and two times using the parameters (100,1). Then, the Epps test is applied two times: One to a projected time series resulting from the parameters (2,7) and the other to one resulting from (100,1). It is the same for the Lovato and Velasco test on the other two projected time series. The results are illustrated in Figure 4 where these p-values are plotted against the corresponding ones reported in Table 7. Thus, the x-axis represents the p-values obtained when computing the Epps (left column) and the Lobato and Velasco test (right column) of the sea height time series without performing a projection. Note that those p-values illustrate whether the one-dimensional marginal of the process follows a Gaussian distribution. Meanwhile, the y-axis displays the obtained p-values when computing the corresponding test on the projected process. The top row plots make use of the parameters (100,1) for the projection and the bottom row plots of the parameters (2,7). A grey line at $y = 0.05$ is drawn to differentiate the p-values that are above or below it. It can be observed that in the two top plots, the p-values obtained with and without a projection are similar. Remember that, as commented in Section 3, making use of the (100,1) parameters results in a projected time series similar to the original one. However, in the case that the (2,7) parameters are used, the p-values obtained when projecting are larger or equivalent to those when the test is applied without performing a projection.

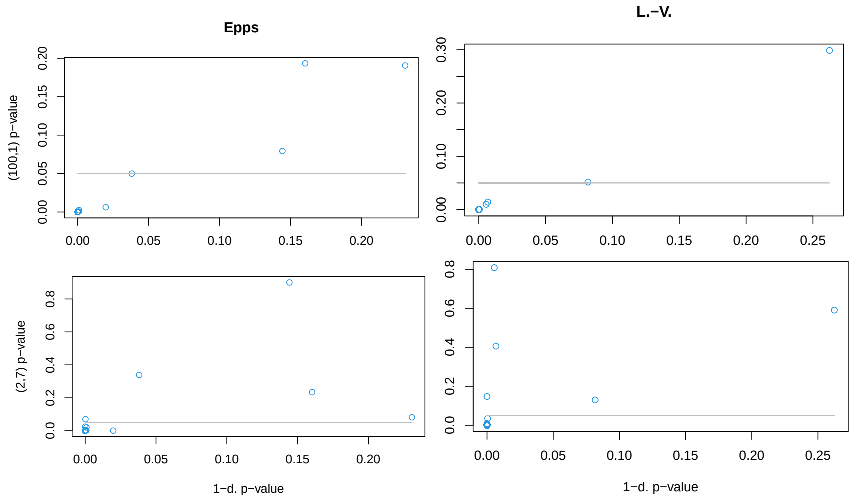

Figure 4. For each of the 63 studied time series, in each of the panels, the p-value is plotted and obtained by applying a Guassianity test on the time series (x-axis) against the one resulting on applying the same test but on the corresponding projected time series (y-axis). The left column makes use of the Epps test and the right of the Lobato and Velasco test. For the projection, top row makes use of the parameters (100,1) and the bottom row of (2,7). The line $y = 0.05$ is displayed in each panel in gray.

Table 7. Gaussianity tests for the one-dimensional marginal of the height of sea waves. For each studied *buoys* (first and fifth columns), it is reported the p-values resulting from applying the Epps (second and sixth columns) and the Lobato and Velasco test (third and seventh columns) and adjusted p-values using the false discovery rate (FDR) for dependent tests (fourth and eighth columns). The bold indicates that the values are larger than 0.05.

Buoy	Epps	L.-V.	FDR	Buoy	Epps	L.-V.	FDR
28	0	1.12×10^{-258}	0	188	3.81×10^{-2}	$\mathbf{2.62} \times 10^{-1}$	$\mathbf{7.62} \times 10^{-2}$
29	2.41×10^{-217}	1.02×10^{-133}	4.81×10^{-217}	189	1.78×10^{-4}	1.56×10^{-11}	3.13×10^{-11}
36	9.61×10^{-277}	1.5×10^{-196}	1.92×10^{-276}	191	0	0	0
45	0	0	0	192	0	1.63×10^{-268}	0
67	0	0	0	194	1.81×10^{-71}	3.75×10^{-117}	7.51×10^{-117}
71	9.33×10^{-303}	0	0	198	0	0	0
76	0	0	0	200	7.52×10^{-4}	6.84×10^{-8}	1.37×10^{-7}
92	0	0	0	201	0	2.21×10^{-234}	0
94	0	4.49×10^{-49}	0	202	$\mathbf{1.6} \times 10^{-1}$	6.74×10^{-3}	1.35×10^{-2}
100	0	4.52×10^{-263}	0	203	0	0	0
101	2.95×10^{-268}	5.97×10^{-210}	5.9×10^{-268}	204	1.98×10^{-2}	6.9×10^{-8}	1.38×10^{-7}
106	2.25×10^{-175}	1.9×10^{-103}	4.5×10^{-175}	209	0	9.03×10^{-318}	0
121	1.25×10^{-23}	6.68×10^{-38}	1.34×10^{-37}	213	0	0	0
132	8.85×10^{-4}	1.62×10^{-18}	3.23×10^{-18}	214	0	0	0
134	1.05×10^{-287}	0	0	215	0	6.3×10^{-309}	0
142	$\mathbf{1.44} \times 10^{-1}$	5.48×10^{-3}	1.1×10^{-2}	217	0	0	0
143	0	0	0	220	0	0	0
144	6.72×10^{-105}	1.71×10^{-175}	3.41×10^{-175}	221	0	1.66×10^{-34}	0
147	2.8×10^{-194}	5.09×10^{-158}	5.59×10^{-194}	222	1.57×10^{-194}	8.51×10^{-105}	3.14×10^{-191}
150	0	1.9×10^{-153}	0	224	0	0	0
153	6.03×10^{-18}	8.73×10^{-12}	1.21×10^{-17}	226	8.42×10^{-251}	1.42×10^{-101}	1.68×10^{-250}
154	0	2.34×10^{-160}	0	238	$\mathbf{2.31} \times 10^{-1}$	$\mathbf{8.17} \times 10^{-2}$	$\mathbf{1.63} \times 10^{-1}$
155	0	0	0	243	1.77×10^{-98}	1.14×10^{-68}	3.55×10^{-98}
157	0	0	0	244	-	-	-
158	0	0	0	248	1.39×10^{-6}	4.47×10^{-4}	2.79×10^{-6}
160	0	3.16×10^{-214}	0	249	0	0	0
168	0	0	0	250	3.97×10^{-315}	2.69×10^{-307}	7.95×10^{-315}
171	5.42×10^{-100}	4.64×10^{-183}	9.28×10^{-183}	254	1.64×10^{-5}	1.58×10^{-6}	3.16×10^{-6}
179	0	2.39×10^{-302}	0	255	0	0	0

Table 7. *Cont.*

Buoy	Epps	L.-V.	FDR	Buoy	Epps	L.-V.	FDR
181	0	2.62×10^{-167}	0	256	1.4×10^{-37}	2.31×10^{-43}	4.62×10^{-43}
185	4.4×10^{-139}	1.85×10^{-198}	3.7×10^{-198}	430	1.79×10^{-195}	1.49×10^{-46}	3.58×10^{-198}
187	1.72×10^{-258}	0	0	433	2.87×10^{-167}	2.05×10^{-54}	5.75×10^{-167}

For each buoy, Table 8 reports the FDR-adjusted p-value for dependent tests resulting of combining the four p-values illustrated in the y-axis of Figure 4. The p-values that does not result in a rejection are highlighted in bold in Table 8. There it can observed that the random projection test is able to reject the null hypothesis of Gaussianity in 60 out of the 63 buoys. This is one case less than the obtained without projecting the different time series.

Table 8. FDR adjusted p-value resulting of combining four p-values to apply the random projection test for each buoy. Values larger than 0.05 is highlighted in bold.

Buoy	Epps		L.-V.		FDR
	(100,1)	(2,7)	(100,1)	(2,7)	
28	0	1.64×10^{-105}	6.64×10^{-105}	8.42×10^{-8}	0
29	3.51×10^{-202}	7.59×10^{-166}	1.58×10^{-144}	2.43×10^{-10}	1.4×10^{-201}
36	9.97×10^{-308}	5.44×10^{-76}	3.71×10^{-154}	1.11×10^{-5}	3.99×10^{-307}
45	0	1.83×10^{-146}	0	0	0
67	0	2.95×10^{-256}	0	3.74×10^{-124}	0
71	6.98×10^{-316}	5.94×10^{-177}	0	1.29×10^{-137}	0
76	0	3.03×10^{-210}	0	8.4×10^{-155}	0
92	0	1.03×10^{-278}	5.85×10^{-315}	6.23×10^{-14}	0
94	5.73×10^{-141}	1.6×10^{-111}	1.73×10^{-30}	5.94×10^{-5}	2.29×10^{-140}
100	0	2.68×10^{-149}	2.36×10^{-169}	1.39×10^{-13}	0
101	4.97×10^{-272}	5.8×10^{-148}	8.39×10^{-131}	5.87×10^{-5}	1.99×10^{-271}
106	7.7×10^{-193}	4.75×10^{-26}	8.91×10^{-84}	7.38×10^{-8}	3.08×10^{-192}
121	5.78×10^{-26}	3.34×10^{-13}	1.03×10^{-36}	8.09×10^{-18}	4.12×10^{-36}
132	2.44×10^{-3}	2.03×10^{-2}	6.56×10^{-17}	4.66×10^{-5}	2.62×10^{-16}
134	5.95×10^{-281}	2.22×10^{-68}	0	6.41×10^{-46}	0
142	$\mathbf{7.94 \times 10^{-2}}$	$\mathbf{9 \times 10^{-1}}$	9.87×10^{-3}	$\mathbf{8.09 \times 10^{-1}}$	3.95×10^{-2}
143	0	8.08×10^{-208}	0	1.73×10^{-144}	0
144	3.48×10^{-96}	1.91×10^{-53}	8.31×10^{-146}	1.87×10^{-28}	3.33×10^{-145}
147	1.42×10^{-182}	2.48×10^{-269}	1.22×10^{-122}	1.55×10^{-15}	9.91×10^{-269}
150	0	3.68×10^{-77}	5.46×10^{-81}	1.41×10^{-5}	0
153	1.82×10^{-17}	9.13×10^{-12}	4.37×10^{-13}	6.95×10^{-8}	7.3×10^{-17}
154	0	0	2.54×10^{-133}	1.45×10^{-36}	0
155	3.8×10^{-308}	8.67×10^{-83}	0	1.89×10^{-140}	0
157	0	2.42×10^{-151}	1.26×10^{-301}	1.36×10^{-13}	0
158	0	6×10^{-162}	9.86×10^{-289}	6.38×10^{-26}	0
160	0	1.24×10^{-175}	8.81×10^{-152}	1.92×10^{-10}	0
168	0	1.49×10^{-210}	0	2.19×10^{-177}	0
171	6.8×10^{-111}	4.02×10^{-45}	2.3×10^{-169}	2.89×10^{-33}	9.2×10^{-169}
179	2.93×10^{-298}	3.69×10^{-167}	2.38×10^{-271}	2.04×10^{-24}	1.17×10^{-297}
181	0	2.48×10^{-70}	6.52×10^{-126}	1.28×10^{-6}	0
185	9.48×10^{-137}	4.31×10^{-58}	1.61×10^{-158}	7.8×10^{-26}	6.42×10^{-158}
187	1.85×10^{-274}	7.41×10^{-76}	0	2.52×10^{-105}	0
188	4.99×10^{-2}	$\mathbf{3.39 \times 10^{-1}}$	$\mathbf{\times 10^{-1}}$	$\mathbf{\times 10^{-1}}$	$\mathbf{2 \times 10^{-1}}$
189	3.99×10^{-4}	$\mathbf{6.99 \times 10^{-2}}$	1.41×10^{-10}	1.07×10^{-3}	5.62×10^{-10}
191	0	0	0	0	0
192	0	4.43×10^{-252}	4.65×10^{-187}	1.37×10^{-18}	0
194	1.02×10^{-72}	6.07×10^{-31}	9.61×10^{-112}	4.3×10^{-34}	3.84×10^{-111}
198	0	6.67×10^{-254}	0	6.04×10^{-41}	0

Table 8. *Cont.*

Buoy	Epps		L.-V.		FDR
	(100,1)	(2,7)	(100,1)	(2,7)	
200	5.27×10^{-4}	1.64×10^{-3}	5.17×10^{-7}	5.22×10^{-3}	2.07×10^{-6}
201	0	5.91×10^{-172}	6.19×10^{-228}	2.74×10^{-18}	0
202	$\mathbf{1.93 \times} 10^{-1}$	$\mathbf{2.34 \times} 10^{-1}$	1.41×10^{-2}	$\mathbf{4.06 \times} 10^{-1}$	$\mathbf{5.62 \times} 10^{-2}$
203	0	9.69×10^{-157}	2.52×10^{-296}	3.65×10^{-213}	0
204	6.17×10^{-3}	4.85×10^{-4}	1.14×10^{-7}	1.6×10^{-3}	4.55×10^{-7}
209	0	1.55×10^{-160}	6.35×10^{-227}	4.8×10^{-21}	0
213	0	2.89×10^{-168}	0	3.93×10^{-41}	0
214	0	2.75×10^{-309}	0	4.2×10^{-17}	0
215	0	3.27×10^{-149}	3.46×10^{-281}	1.67×10^{-11}	0
217	0	6.4×10^{-197}	2.47×10^{-239}	1.69×10^{-27}	0
220	0	1.27×10^{-246}	0	5.2×10^{-126}	0
221	0	9.72×10^{-119}	7.12×10^{-31}	4.02×10^{-4}	0
222	4.17×10^{-194}	3.94×10^{-62}	1.18×10^{-58}	$\mathbf{1.47 \times} 10^{-1}$	1.67×10^{-193}
224	0	5.35×10^{-217}	4.09×10^{-213}	6.44×10^{-12}	0
226	0	2.99×10^{-176}	4.11×10^{-89}	4.17×10^{-10}	0
238	$\mathbf{1.91 \times} 10^{-1}$	$\mathbf{8.15 \times} 10^{-2}$	$\mathbf{5.17 \times} 10^{-2}$	$\mathbf{1.3 \times} 10^{-1}$	$\mathbf{1.91 \times} 10^{-1}$
243	9.55×10^{-97}	2.45×10^{-73}	3.32×10^{-65}	2.35×10^{-8}	3.82×10^{-96}
248	1.12×10^{-6}	2.44×10^{-2}	4.61×10^{-4}	3.51×10^{-2}	4.47×10^{-6}
249	0	0	0	2.81×10^{-144}	0
250	7.49×10^{-301}	1.75×10^{-285}	1.78×10^{-203}	5.94×10^{-10}	3×10^{-300}
254	3.34×10^{-6}	2.55×10^{-3}	1.43×10^{-6}	8.38×10^{-3}	5.71×10^{-6}
255	0	0	0	6.69×10^{-133}	0
256	2.79×10^{-40}	6.22×10^{-32}	2.81×10^{-42}	8.9×10^{-20}	1.13×10^{-41}
430	1.51×10^{-181}	4.04×10^{-129}	7.77×10^{-35}	9.29×10^{-4}	6.02×10^{-181}
433	4.94×10^{-163}	1.83×10^{-149}	5.96×10^{-47}	1.36×10^{-5}	1.98×10^{-162}

They correspond to buoys 188, 202, and 238, with 202 the buoy that did result in a rejection when making use of the adjusted FDR p-value without the use of projections.

Let us analyze these results in detail. For buoy 188, the null hypothesis of Gaussianity is rejected when making use of the Epps test without projecting the data (Table 7) and when the (100,1) parameters are used for the projection (Table 8). Note that as this projection results in a time series similar to the original, this is expected to happen. Despite these rejections, as the other performed tests do not result in a rejection, none of the two corresponding adjusted p-values result in a rejection. For the case of buoy 238, it also occurs that none of the two corresponding adjusted p-values result in a rejection. However, for this buoy, it is due to none of the performed tests having resulted in a rejection, independently of whether projections were made or not. The case of buoy 202 differs as it does not result in a rejection (adjusted FDR p-value smaller than 0.05) when projections are made but it does when the data is not projected. This is in part due to the FDR for dependent data in a conservative methodology, which loses power when the amount of performed tests increases. The p-values smaller than 0.05 for this buoy correspond to the Lovato and Velasco test without a projection and projecting with the parameters (100,1). Note that the same behavior has been commented above for buoy 188 with the Epps test.

5. Conclusions

This manuscript is dedicated to empirically studying the Gaussianity of the sea waves heights, which is commonly assumed in the literature. The measurements provided by 62 buoys along the US coast are studied and obtained that over 96% of the studied time series were drawn from a non-Gaussian process. The analysis is novel in that it makes use of a test that is powerful against non-Gaussian processes with one-dimensional Gaussian marginals.

The Gaussian distribution is also known as normal distribution. The name *normal* comes from the fact that it is the most common distribution in nature in one-dimensional spaces. However, as it is shown in this study that it might not be the most common one in higher dimensional spaces. This goes in line with the fact that non-parametric statistics, where no distribution of the data is assumed, is nowadays the most common way to study multivariate and functional data.

The implication of this study is that the Gaussian assumption should not be used in simulating sea wave heights. This is due to the assumption that a false hypothesis results incurs an error. That not using the appropriate models results in incurring an error is not new. In fact [10] is dedicated to measuring and comparing the error caused by different wave height calculation models. As commented there, it is important to incur the lowest possible error, that is, to have an accurate model, in order to design appropriately offshore engineering structures, such as drilling ships and offshore platforms, under different marine scenarios.

A direct consequence of the study presented here is that it can be assumed, though, the one-dimensional marginal of the process is generally non-Gaussian. Further studies on the matter could include the study of the two-dimensional marginal distribution of the process, and the distribution of the full process. Once this distribution is known, the error in assuming Gaussianity could be easily quantified. The literature contains already studies on the two-dimensional marginal distribution of the process ([45–47] and the references therein). There, a model is assumed and the parameters of it obtained. Knowing the one-dimensional marginal distribution is not Gaussian will help in selecting the appropriate model.

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

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 4 October 2021).

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

Abbreviations

The following abbreviations are used in this manuscript:

FDR	False discovery rate
GMT	Greenwich mean time
UTC	Coordinated universal time

References

1. Forristall, G.Z. On the statistical distribution of wave heights in a storm. *J. Geophys. Res. Ocean.* **1978**, *83*, 2353–2358. [CrossRef]
2. Azaïs, J.M.; León, J.R.; Ortega, J. Geometrical characteristics of Gaussian sea waves. *J. Appl. Probab.* **2005**, *42*, 407–425. [CrossRef]
3. Karmpadakis, I.; Swan, C.; Christou, M. Assessment of wave height distributions using an extensive field database. *Coast. Eng.* **2020**, *157*, 103630. [CrossRef]
4. Tayfun, M.A. Distribution of Large Wave Heights. *J. Waterw. Port Coastal Ocean. Eng.* **1990**, *116*, 686–707. [CrossRef]
5. Mori, N.; Liu, P.C.; Yasuda, T. Analysis of freak wave measurements in the Sea of Japan. *Ocean. Eng.* **2002**, *29*, 1399–1414. [CrossRef]
6. Stansell, P. Distributions of freak wave heights measured in the North Sea. *Appl. Ocean. Res.* **2004**, *26*, 35–48. [CrossRef]
7. Stansell, P. Distributions of extreme wave, crest and trough heights measured in the North Sea. *Ocean. Eng.* **2005**, *32*, 1015–1036. [CrossRef]
8. Casas-Prat, M.; Holthuijsen, L. Short-term statistics of waves observed in deep water. *J. Geophys. Res. Ocean.* **2010**, *115*. [CrossRef]
9. Katalinić, M.; Parunov, J. Uncertainties of Estimating Extreme Significant Wave Height for Engineering Applications Depending on the Approach and Fitting Technique—Adriatic Sea Case Study. *J. Mar. Sci. Eng.* **2020**, *8*, 259. [CrossRef]
10. Chen, B.; Kou, Y.; Wu, F.; Wang, L.; Liu, G. Study on evaluation standard of uncertainty of design wave height calculation model. *J. Oceanol. Limnol.* **2021**, *39*, 1188–1197. [CrossRef]
11. Wang, L.; Li, J.; Liu, S.; Ducrozet, G. Statistics of long-crested extreme waves in single and mixed sea states. *Ocean. Dyn.* **2021**, *71*, 21–42. [CrossRef]
12. Ochi, M.K. *Ocean Waves: The Stochastic Approach*; Cambridge Ocean Technology Series; Cambridge University Press: Cambridge, UK, 1998. [CrossRef]

13. Said, S.E.; Dickey, D.A. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika* **1984**, *71*, 599–607. [CrossRef]
14. 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]
15. Longuet-Higgins, M.S.; Deacon, G.E.R. The statistical analysis of a random, moving surface. *Philos. Trans. R. Soc. London. Ser. A Math. Phys. Sci.* **1957**, *249*, 321–387.
16. 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]
17. Jishad, M.; Yadhunath, E.; Seelam, J.K. Wave height distribution in unsaturated surf zones. *Reg. Stud. Mar. Sci.* **2021**, *44*, 101708. [CrossRef]
18. 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]
19. Naess, A. On the distribution of crest to trough wave heights. *Ocean. Eng.* **1985**, *12*, 221–234. [CrossRef]
20. Boccotti, P. On Mechanics of Irregular Gravity Waves. *Atti Della Accad. Naz. Dei Lincei Mem.* **1989**, *19*, 110–170.
21. Klopman, G. Extreme wave heights in shallow water. In *WL | Delft Hydraulics*; Report H2486; Deltares: Delft, The Netherlands, 1996.
22. van Vledder, G.P. *Modification of the Glukhovskiy Distribution*; Report H1203; Technical Report, Delft Hydraulics; Deltares: Delft, The Netherlands, 1991.
23. Battjes, J.A.; Groenendijk, H.W. Wave height distributions on shallow foreshores. *Coast. Eng.* **2000**, *40*, 161–182. [CrossRef]
24. Mendez, F.J.; Losada, I.J.; Medina, R. Transformation model of wave height distribution on planar beaches. *Coast. Eng.* **2004**, *50*, 97–115. [CrossRef]
25. 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]
26. Haver, S. On the joint distribution of heights and periods of sea waves. *Ocean. Eng.* **1987**, *14*, 359–376. [CrossRef]
27. Mendes, S.; Scotti, A. The Rayleigh-Haring-Tayfun distribution of wave heights in deep water. *Appl. Ocean. Res.* **2021**, *113*, 102739. [CrossRef]
28. Xiaolong, L.; Zhiwen, C.; Ze, S.; Wenwei, C.; Jun, Y.; Jun, D.; Yonglin, Y. Study on long-term distribution and short-term characteristics of the waves near islands and reefs in the SCS based on observation. *Ocean. Eng.* **2020**, *218*, 108171. [CrossRef]
29. Butler, R.W.; Machado, U.; Rychlik, I. Distribution of wave crests in a non-Gaussian sea. *Appl. Ocean. Res.* **2003**, *31*, 57–64. [CrossRef]
30. Petrova, P.; Cherneva, Z.; Guedes Soares, C. Distribution of crest heights in sea states with abnormal waves. *Appl. Ocean. Res.* **2006**, *28*, 235–245. [CrossRef]
31. 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]
32. Tayfun, M.; Alkhalidi, M. Distribution of sea-surface elevations in intermediate and shallow water depths. *Coast. Eng.* **2020**, *157*, 103651. [CrossRef]
33. Epps, T.W. Testing That a Stationary Time Series is Gaussian. *Ann. Stat.* **1987**, *15*, 1683–1698. [CrossRef]
34. Lobato, I.; Velasco, C. A simple Test of Normality for Time Series. *Econom. Theory* **2004**, *20*, 671–689. [CrossRef]
35. Psaradakis, Z. *Normality Tests for Dependent Data*; Working and Discussion Papers WP 12/2017; Research Department, National Bank of Slovakia: Bratislava, Slovakia, 2017.
36. Bontemps, C.; Meddahi, N. Testing Normality: A GMM approach. *J. Econom.* **2005**, *124*, 149–186. [CrossRef]
37. 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]
38. 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]
39. 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]
40. Jarque, C.M.; Bera, A.K. Efficient tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. *Econ. Lett.* **1980**, *6*, 255–259. [CrossRef]
41. López, I.; Carballo, R.; Iglesias, G. Intra-annual variability in the performance of an oscillating water column wave energy converter. *Energy Convers. Manag.* **2020**, *207*, 112536. [CrossRef]
42. Ciappi, L.; Cheli, L.; Simonetti, I.; Bianchini, A.; Talluri, L.; Cappietti, L.; Manfrida, G. Wave-to-wire models of wells and impulse turbines for oscillating water column wave energy converters operating in the Mediterranean Sea. *Energy* **2022**, *238*, 121585. [CrossRef]
43. Benjamini, Y.; Yekutieli, D. The Control of the False Discovery Rate in Multiple Testing under Dependency. *Ann. Stat.* **2001**, *29*, 1165–1188. [CrossRef]
44. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300. [CrossRef]

45. Lin, Y.; Dong, S.; Tao, S. Modelling long-term joint distribution of significant wave height and mean zero-crossing wave period using a copula mixture. *Ocean. Eng.* **2020**, *197*, 106856. [CrossRef]
46. Huang, W.; Dong, S. Joint distribution of individual wave heights and periods in mixed sea states using finite mixture models. *Coast. Eng.* **2020**, *161*, 103773. [CrossRef]
47. Huang, W.; Dong, S. Joint distribution of significant wave height and zero-up-crossing wave period using mixture copula method. *Ocean. Eng.* **2021**, *219*, 108305. [CrossRef]