Regression Models for Outlier Identification (Hurricanes and Typhoons) in Wave Hindcast Databases

R. MÍNGUEZ AND B. G. REGUERO

Environmental Hydraulics Institute "IH Cantabria," Universidad de Cantabria, Santander, Cantabria, Spain

A. LUCEÑO

Department of Applied Mathematics and Computational Sciences, Universidad de Cantabria, Santander, Cantabria, Spain

F. J. MÉNDEZ

Environmental Hydraulics Institute "IH Cantabria," Universidad de Cantabria, Santander, Cantabria, Spain

(Manuscript received 31 March 2011, in final form 4 July 2011)

ABSTRACT

The development of numerical wave prediction models for hindcast applications allows a detailed description of wave climate in locations where long-term instrumental records are not available. Wave hindcast databases (WHDBs) have become a powerful tool for the design of offshore and coastal structures, offering important advantages for the statistical characterization of wave climate all over the globe (continuous time series, wide spatial coverage, constant time span, homogeneous forcing, and more than 60-yr-long time series). However, WHDBs present several deficiencies reported in the literature. One of these deficiencies is related to typhoons and hurricanes, which are inappropriately reproduced by numerical models. The main reasons are (i) the difficulty of specifying accurate wind fields during these events and (ii) the insufficient spatiotemporal resolution used. These difficulties make the data related to these events appear as "outliers" when compared with instrumental records. These bad data distort results from calibration and/or correction techniques. In this paper, several methods for detecting the presence of typhoons and/or hurricane data are presented, and their automatic outlier identification capabilities are analyzed and compared. All the methods are applied to a global wave hindcast database and results are compared with existing hurricane and buoy databases in the Gulf of Mexico, Caribbean Sea, and North Atlantic Ocean.

1. Introduction

In the last decade, the traditional approach to climatology based on observations has evolved toward a state-of-the-art data assimilation system, which is used to reprocess all past environmental observations in combination with numerical models consistent with atmospheric equations. The improved methodology allows us to obtain the best estimate of the state and evolution of the atmosphere. It can also be considered as a reintegration of our knowledge about the atmosphere into an easily accessible global atmospheric reanalysis database. This source of information provides different climate variables, such as wind fields, in a regular grid.

These atmospheric reanalysis databases can be subsequently reprocessed using wind wave models, which allow the simulation of the wave generation and propagation processes all over the globe. As in the meteorological case, these models provide consistent datasets to define the wave climatology. However, since wave models do not incorporate wave instrumental observations, the resulting databases are called wave hindcast rather than reanalysis.

In the last years, the importance of wave hindcast databases for the design of offshore and coastal structures has increased considerably. The main reason is their ability to provide a detailed description of wave climate (i.e., long continuous time series records with wide spatial coverage) in locations where long-term instrumental

Corresponding author address: Roberto Minguez, Environmental Hydraulics Institute "IH Cantabria," Universidad de Cantabria, Avda de los Castros, Santander, Cantabria 39005, Spain.
E-mail: roberto.minguez@unican.es

DOI: 10.1175/JTECH-D-11-00059.1

records are not available. However, hindcast models use (i) several simplifying assumptions of reality and (ii) discrete forcing fields consisting of surface winds at different times, and for these reasons hindcast results present differences when compared with instrumental data (buoys and/or satellites; Caires and Sterl 2005; Cavaleri and Sclavo 2006). Besides, if the orography is complex, the hindcast inaccuracy becomes more evident (Cavaleri and Bertotti 2004) as a result of the inappropriate spatial and temporal resolution and inaccurate description of wind fields.

An additional problem related to wave hindcast databases is the bad performance during hurricanes and typhoons. These inconsistencies are produced because of the difficulty of specifying accurate wind fields and the scarcity of high-quality wave measurements during these events. Thus, to better catch up ocean surface behavior when hurricane and typhoons occur, models with higher spatial and temporal resolution must be used. These models take advantage of (i) the advances made in recent years in the analysis of the time and space evolution of surface wind fields, especially in North Atlantic basin hurricanes (Powell et al. 1998), and (ii) the high-quality wind datasets from remote sensing systems. However, these models are too time consuming and they should only be used when and where the global wave hindcast does not appropriately reproduce the wave climate (i.e., during those hurricanes and typhoons that produce important discrepancies between hindcast results and instrumental data).

Coastal management and design demand the appropriate definition of the wave climate. This requirement has resulted in an increased interest in collecting information through instrumental devices (i.e., buoys and satellites). For example, the National Oceanic and Atmospheric Administration (NOAA) National Data Buoy Center (NDBC) has a fairly dense rich array of moored data buoys around the United States. In addition, several satellite missions [Skylab, GEOS-3, *Seasat*, *Geosat*, the Ocean Topography Experiment (TOPEX)/Poseidon, the *European Remote Sensing Satellites-1* and *-2* (*ERS-1*) and (*ERS-2*), the *GEOSAT Follow-On* (*GFO*), *Jason-1*, the *Environmental Satellite* (*Envisat*), and *Jason-2*] incorporate altimetry sensors for the evaluation of different ocean climate variables with a high level of precision (i.e., ± 3 cm; Krogstad and Barstow 1999). These measurements are considerably more accurate than wave hindcast databases (WHDBs). However, there are also several shortcomings to be considered, such as disruptions on normal use due to failures, and temporal and spatial inhomogeneous records, which limit their use to certain regions, mostly related to developed countries. These reasons have motivated an increased interest in developing different wave generation models, such as the Wave Model (WAM) developed by the Wave Model Development and Implementation Group (WAMDI; WAMDI Group 1988) or Wave Watch (Tolman 1997, 1999, 2002). These models try to reproduce wave generation and propagation processes using wind fields as input data (Caires et al. 2004; Pilar et al. 2008; Dodet et al. 2010).

Since instrumental (buoys and/or satellites) and hindcast sources of information have advantages and drawbacks (Cavaleri and Sclavo 2006), several authors attempt to combine both types of information. Caires and Sterl (2005) establish a nonparametric correction based on analogs taken from a learning dataset. Cavaleri and Sclavo (2006) obtain calibrated decadal time series at a large number of points over the Mediterranean Sea. They use the overall information on models, buoys, and satellites. Tomás et al. (2008) include spatial correlation in the calibration process, proposing a spatial calibration procedure based on empirical orthogonal functions and a nonlinear transformation of the spatial–time modes. Mínguez et al. (2011) propose a calibration method based on a nonlinear regression problem in which the corresponding correction parameters vary smoothly along the possible wave directions by means of cubic splines. This procedure is based on a point-to-point basis including wave direction, but without considering the spatial correlation between neighboring nodes. However, none of these approaches provide a rational criterion to detect data associated with hurricanes and typhoons, which should be treated with care within the calibration process. Note that failing to exclude these outlying observations may provoke large distortion of calibration results. Besides, these data should be treated and analyzed separately for the results to be fully reliable. Efforts in this direction can be found in Cardone et al. (1976, 1996). This outlier detection task is of great importance if hindcast database information is used for maximum significant wave analysis, especially for the design of coastal protection and offshore structures, because it may underestimate maximum significant wave heights associated with given return periods, thus compromising safety and functionality.

Because of the difficulties of defining the wave climate, we are forced to work with mathematical and statistical models, as those proposed in this paper. Nevertheless, mathematical and statistical models are simplifications of reality and their results must be used with caution. For instance, it is known that in certain regions of the world, hurricane data may be present in instrumental records. Therefore, it is interesting to have statistical methods to automatically detect and/or remove outliers and other unduly influential observations. This would protect the results of the analysis from the influence of these rare events. Note that the techniques proposed in this paper would allow deciding "where" and "when" specific numerical models for hurricanes and typhoons should be used instead of wave hindcast databases.

There is a large amount of literature on outlier detection; see, for example, the books by Hawkins (1980), Belsley et al. (1980), Cook and Weisberg (1982), Atkinson (1985), Chatterjee and Hadi (1988), and Barnett and Lewis (1994), and the articles by Pregibon (1981), Gray and Ling (1984), Gray (1986), Cook (1986), Jones and Ling (1988), Weissfeld and Schneider (1990a,b), Schwarzmann (1991), Paul and Fung (1991), Simonoff (1991), Nyquist (1992), Hadi and Simonoff (1993), Atkinson (1984), Peña and Yohai (1995), Barrett and Gray (1997), Mayo and Gray (1997), Billor et al. (2001), and Winsnowski et al. (2001). As can be seen in these books and articles, the literature has focused mainly on the area of least squares linear regression. Other statistical models and estimation methods, such as reweighed techniques (Luceño 1997, 1998a,b), nonlinear methods (Castillo et al. 2004), heteroscedastic models (Cheng 2011), or some robust estimators (Rousseeuw and Leroy 1987; Rousseeuw and Van Driessen 1999) have received comparatively less attention.

The aim of this paper is twofold: first to present several outlier detection techniques for hurricanes and typhoons, and second to compare results from those techniques giving some recommendations.

The paper is organized as follows. In section 2, the dataset used for this study is described. section 3 presents four different methods for outlier detection. In section 4, the functioning of the different methods is illustrated through several examples using data from the Gulf of Mexico, the Caribbean Sea, and the North Atlantic Ocean. Finally, in section 5 relevant conclusions are drawn and some recommendations are given.

2. Data sources

For this study we have used the following database information:

(i) Significant wave height data from 43 buoys from NOAA/NDBC (http://www.ndbc.noaa.gov/) over the Gulf of Mexico, the Caribbean Sea, and the Atlantic Ocean. The main characteristics of the buoys used are given in Table 1, and their locations are shown in Fig. 1.

(ii) Atlantic Hurricane Database (HURDAT): This database consists of an ASCII (text) file containing the 6-hourly center locations (latitude and longitude in tenths of degrees) and intensities (maximum 1-min surface wind speeds in knots and minimum central pressures in millibars) for all tropical storms and hurricanes from 1851 to 2009 (Jarvinen et al. 1984; Landsea et al. 2004, 2008). Figure 1 shows the hurricane tracks from Atlantic HURDAT database and the tracks of some Atlantic storms.

(iii) Global Ocean Waves (GOW): This is a global wave hindcast from 1948 onward developed by the Environmental Hydraulics Institute "IH Cantabria." It uses the third-generation model Wave Watch III (Tolman 1997, 1999) forced by 6-hourly wind fields from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) atmosphere model. The GOW database has different spatial scales: (i) a global grid at $1.5° \times 1°$ (longitude–latitude) spatial resolution, (ii) an Atlantic coast grid at $0.5° \times 0.5°$ spatial resolution, and (iii) a Caribbean coast grid at $0.25° \times 0.25°$ spatial resolution.

To increase the confidence in wave hindcast databases, results must be postprocessed and validated with instrumental data (buoys and/or satellites). For this task, hindcast versus instrumental data pairs coincident in time and space must be selected. For this particular case, and because of the hindcast homogeneity both in time and space, database information is interpolated to the buoy positions and to the times where buoy data are recorded. These data pairs are used for validation and calibration. The aim of this paper is to propose methods for detecting data pairs associated with hurricanes and typhoons previously to validating and/or applying any calibration–correction technique.

An example of these data and the hurricane effect on hindcast validation is shown in Fig. 2, where the instrumental and hindcast significant wave records at buoy 42059 (eastern Caribbean) are plotted. Note in Fig. 2a that the hindcast time series captures appropriately the magnitude and temporal evolution of the instrumental significant wave height record; however, there exist clear discrepancies when hurricane events occur, especially during Dean 2007 and Omar 2008. This effect is also shown in the scatterplot (Fig. 2b), where instrumental and hindcast data occurring during these tropical storms present important discrepancies, which would affect the calibration process and detract the good performance of the hindcast if they were not accounted for appropriately. This paper does not try to detect and remove all data related to hurricanes, but only those data that differ substantially between hindcast and instrumental records. In Fig. 2b there are many data points recorded during the occurrence of different tropical storms where the hindcast performs appropriately. The reason for this behavior is that the hurricane wave generation is a local

TABLE 1. General characteristics of the 43 buoys from the NDBC used for the outlier detection analysis.

Region	Name	ID	Lat	Lon (0°–360°)	Depth (m)	T_0	T_f	Spectral?
Florida eastern Gulf Mexico	Grays Reef	41008	31.402°	−80.871°	18	1988	2008	From 1996
	—	41003	30.4°	−80.1°	—	1977	1982	No
	St. Augustine	41012	30.041°	−80.533°	37.2	2002	2008	Yes
	East Cape Canaveral	41009	28.519°	−80.166°	44.2	1988	2008	From 1996
	—	41006	29.3°	−77.4°	—	1982	1996	From 1996
	East Cape Canaveral	41010	28.906°	−78.471°	872.6	1988	2008	From 1996
	—	42025	24.9°	−80.4°	—	1991	1995	No
	East Southeast Pensacola	42039	28.791°	−86.008°	307	1995	2008	From 1996
	—	42009	29.3°	−87.5°	—	1980	1987	No
	South of Dauphin Island	42040	29.205°	−88.205°	274.3	1995	2008	From 1996
Northeast United States	Nantucket	44007	40.503°	−69.247°	59.1	1982	2008	From 1996
	Gulf of Maine	44005	43.189°	−69.14°	201.2	1978	2008	From 1996
	Boston	44013	42.346°	−70.651°	60	1984	2008	From 1996
	SE Cape Cod	44018	41.255°	−69.305°	63.7	2002	2008	Yes
	Georges Bank	44011	41.111°	−66.58°	88.4	1984	2008	From 1996
	Nantucket	44008	40.503°	−69.247°	59.1	1982	2008	From 1996
	—	44001	38.7°	−73.6°	—	1975 1979	1990 1991	No
	—	44012	38.8°	−74.6°	—	1984	1992	No
	Delaware Bay	44009	38.464°	−74.702°	28	1984	2008	From 1996
	Virginia Beach	44014	36.611°	−74.836°	47.5	1990	2008	From 1996
Southeast United States	—	44006	36.3°	−75.4°	—	1980 1988	19941996	No
	East Cape Hatteras	41001	34.704°	−72.734°	4425.7	1976	2008	From 1996
	Onslow Bay	41036	34.211°	−76.953°	30.8	2006	2008	Yes
	East of Charleston	41002	32.382°	−75.415°	3546	1973	2008	From 1996
	Southeast of Charleston	41004	32.501°	−79.099°	33.5	1978	2008	From 1996
	Bermuda	41048	30.978°	−69.649°	5261	2007	2008	Yes
Western Atlantic	Bahamas	41047	27.469°	−71.491°	5231	2007	2008	Yes
	Bahamas	41046	23.867°	−70.87°	5498.6	2007	2008	Yes
Western Gulf Mexico	—	10000	27.5°	−88°	—	1972	1976	No
	South of Southwest Pass	42001	25.9°	−89.667°	3246	1975	2008	From 1996
	South of Grand Isle	42041	27.504°	−90.462°	—	1999	2005	From 1999
	North mid–Gulf of Mexico	42038	27.421°	−92.555°	—	2004	2006	Yes
	East of Brownsville	42002	25.79°	−93.666°	3566.16	1973	2008	From 1996
	Freeport	42019	27.913°	−95.36°	83.2	1990	2008	From 1996
	Corpus Christi	42020	26.966°	−96.695°	88.1	1990	2008	From 1996
Caribbean	Middle Atlantic	41041	14.357°	−46.008°	3502	2005	2008	Yes
	West Atlantic	41040	14.477°	−53.008°	5267.2	2005	2008	Yes
	Eastern Caribbean	42059	15.006°	−67.496°	4900	2007	2008	Yes
	—	41018	15°	−75°	—	1994	1996	No
Western Caribbean	Bay of Campeche	42055	22.017°	−94.046°	3380.5	2005	2008	Yes
	Yucatan Basin	42056	19.874°	−85.059°	4446	2005	2008	Yes
	Western Caribbean	42057	16.834°	−81.501°	293	2005	2008	Yes
	Central Caribbean	42058	15.093°	−75.064°	4042	2005	2008	Yes

effect. As shown in Fig. 2c, there are four tropical storm tracks passing within 2° distance from the buoy location; however, there are only considerable discrepancies during two of these events:

(i) Dean 2007 evolved from east to west and went through the buoy location on 18 August. At that time, its hurricane category was H5. This is why discrepancies during this event are so high.
(ii) Noel 2007 was born close to the buoy location, being an extratropical storm at the time it passed close to the buoy on 25 October. The maximum category during this event was tropical or subtropical storm. For these reasons, discrepancies may be considered to be within tolerable limits.
(iii) Gustav 2008 was analogous to Noel 2007; its category was tropical or subtropical depression at the time it passed close to the buoy location on 25 August.
(iv) Omar 2008 reached category H1 on 15 October, when it passed close to the buoy location, increasing up to category H4 on 16 October, 500 km away from the buoy location, also producing remarkable discrepancies.

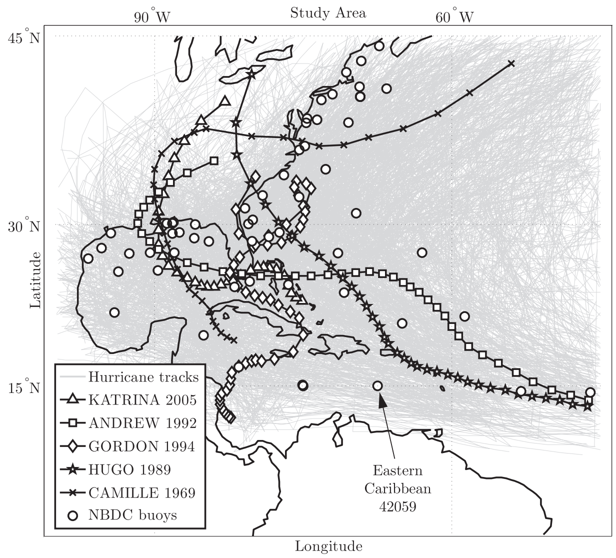

FIG. 1. Area of study showing NBDC buoys locations (open circles), tracks of tropical storms and hurricanes database, and the tracks of some Atlantic storms.

3. Outlier detection techniques

In this section, we start considering the weighted general linear regression model and continue showing different methods to deal with outliers.

a. Weighted least squares

Consider the standard linear regression model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\mathbf{Y} = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$ is a $n \times 1$ response variable vector, $\mathbf{X}$ is a $n \times k$ matrix of predictor variables often called the "design matrix," $\boldsymbol{\beta}$ is a $k \times 1$ vector of regression coefficients or parameters, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^{\mathrm{T}}$ is a $n \times 1$ vector of random errors assumed to be jointly normally distributed random variables $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{V})$, where $\sigma^2\mathbf{V}$ is a positive definite variance-covariance matrix.

Regression parameters $\boldsymbol{\beta}$ are usually estimated using the weighted least squares (WLS) method:

$$\underset{\boldsymbol{\beta}}{\text{Minimize}}\, \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{W}\boldsymbol{\varepsilon} = \underset{\boldsymbol{\beta}}{\text{Minimize}}\, (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{W}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}), \quad (2)$$

where $\mathbf{W} = \mathbf{V}^{-1}$. For Eq. (1), WLS coincides with the maximum likelihood (ML) estimation method. Note that, for homoscedastic models, $\mathbf{W}$ corresponds to the identity matrix (i.e., $w_{ii} = 1; i = 1, \ldots, n; w_{ij} = 0; i, j = 1, \ldots, n$ and $i \neq j$), and Eq. (2) becomes the traditional least squares (LS) method. However, we include matrix $\mathbf{W}$ in the formulation so that regression formulas remain valid for the reweighting approach presented in section 3b. Fitting results are (Draper and Smith 1981) the following:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{Y}), \quad (3)$$

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}, \quad (4)$$

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{P}\mathbf{W}\mathbf{Y}, \quad (5)$$

where the hat (ˆ) refers to estimates, and

$$\mathbf{P} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}, \quad (6)$$

$$\mathrm{Var}(\hat{\mathbf{Y}}) = \sigma^2\mathbf{P}, \quad (7)$$

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \hat{\mathbf{Y}} = (\mathbf{I} - \mathbf{P}\mathbf{W})\mathbf{Y}, \quad (8)$$

$$\mathrm{Var}(\hat{\boldsymbol{\varepsilon}}) = \sigma^2(\mathbf{V} - \mathbf{P}), \quad (9)$$

$$\mathrm{Var}(\hat{\varepsilon}_i) = \sigma^2(v_{ii} - p_{ii}); \quad i = 1, \ldots, n, \quad (10)$$

where v_{ii} and p_{ii} are the ith diagonal element of $\mathbf{V}$ and the projection matrix $\mathbf{P}$, respectively. The residual mean square estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{W}\boldsymbol{\varepsilon}}{n - k}. \quad (11)$$

1) DIFFERENCES BETWEEN INFLUENTIAL OBSERVATIONS AND OUTLIERS

Influential observations can be defined, according to Belsley et al. (1980), as those observations having a larger and excessive impact on the calculated values of some estimates. There are numerous influence measures in the literature, which according to Chatterjee and Hadi (1986) can be classified into five groups based on 1) residuals, 2) the prediction matrix, 3) volume of confidence ellipsoids, 4) influence functions, and 5) partial influence. In contrast, outliers are data that cannot be explained by the model, because they are produced under different dynamics than regular data. One can find outliers that are influential, as well as outliers that are not. Some outliers present large residuals and therefore are easy to detect. However, it is important to realize that some outliers may have small residuals because they have large influence on the parameter estimates; when outliers of this type appear in groups, they are often more difficult to detect even though they are very influential. Finally, there may be some outliers with small residuals that are not influential; these are also difficult to detect, but they are much less important.

Figures 2a,b show (i) the significant wave height evolution in time and (ii) the scatterplots corresponding to buoy 42059 (eastern Caribbean) and hindcast interpolated

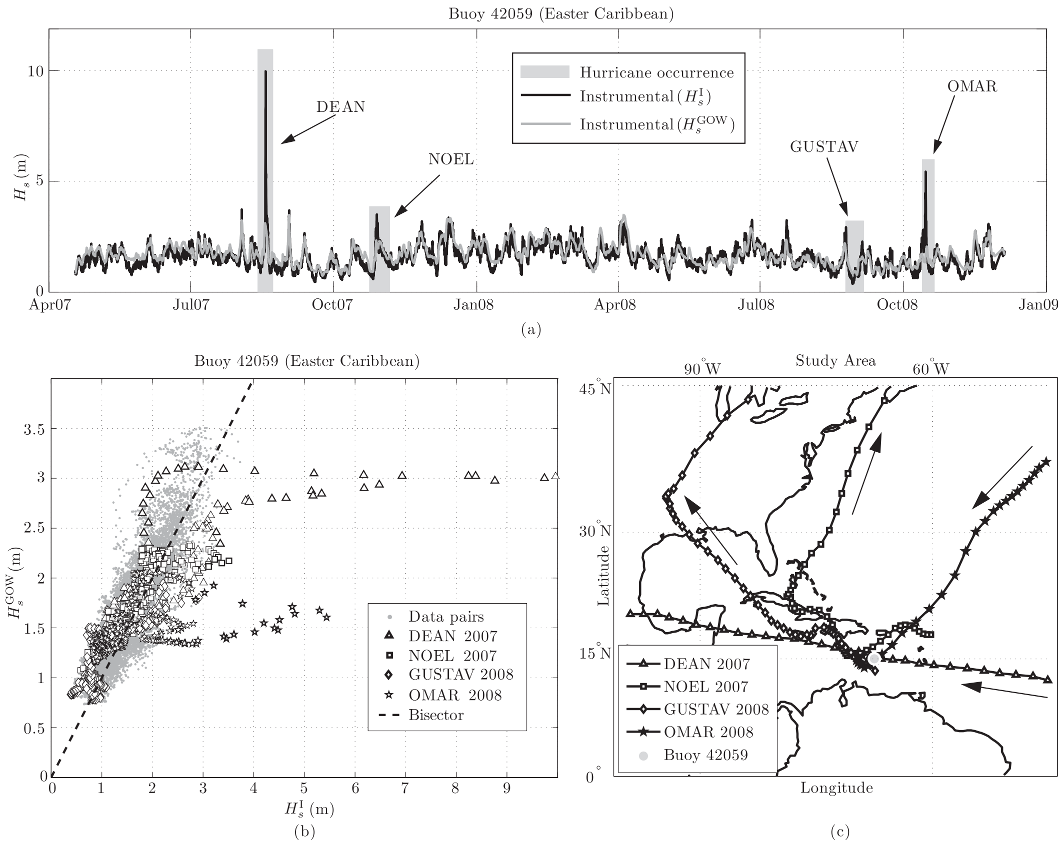

FIG. 2. Data associated with buoy 42059 (eastern Caribbean): (a) instrumental and hindcast significant wave height (m) time evolution, (b) scatterplot including bisector, and (c) tracks of hurricanes passing within a 2° distance from the buoy location.

data. According to these plots, many outliers related to hurricanes seem to have large residuals, but a moderate influence on the fitted regression model.

2) INFLUENCE MEASURES

To assess the effect of outliers associated with hurricanes on the estimators, we use different influence measures, some of them based on the deletion approach (i.e., the influence of the ith observation on a given estimator is calculated comparing results using all data versus results obtained removing the ith observation from the dataset). We have considered the following statistics, which are valid only for $\mathbf{W} = \mathbf{V}^{-1}$ diagonal matrix so that $w_{ii} = v_{ii}{}^{-1}$:

(i) The ith diagonal element of the projection matrix $\mathbf{P}$:

$$p_{ii} = \mathbf{x}_i(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{x}_i^{\mathrm{T}}; \quad i = 1, \ldots, n, \qquad (12)$$

where $\mathbf{x}_i$ is the ith row of the design matrix, which represents the amount of leverage of the response value y_i on the corresponding response estimate $\hat{y}_i$. Note that $\mathrm{Var}(\hat{y}_i) = \sigma^2 p_{ii}$. High leverage points in regression (i.e., points that are outlying in the x space) should be further examined (Hoaglin and Welsch 1978).

(ii) Internally studentized residuals, which are a scaled version of residuals, that is,

$$z_i = \frac{\sqrt{w_{ii}}\varepsilon_i}{\hat{\sigma}\sqrt{1 - w_{ii}p_{ii}}}; \quad i = 1, \ldots, n. \qquad (13)$$

For "outlier" identification purposes, an internally studentized residual corresponds to suspected "bad" data with a $1 - \alpha$ confidence level (e.g., 0.99) if $|z_i| > \Phi^{-1}(1 - \alpha/2)$.

(iii) Externally studentized residuals, a second version of studentized residuals in (13) where $\hat{\sigma}$ is replaced by $\hat{\sigma}_{(i)}$ and $\hat{\sigma}^2_{(i)}$ is the estimator of σ^2 when the ith observation is omitted:

$$\hat{\sigma}^2_{(i)} = \frac{(n-k)\hat{\sigma}^2}{(n-k-1)} - \frac{w_{ii}\varepsilon_i^2}{(n-k-1)(1-w_{ii}p_{ii})}; \quad i = 1, \ldots, n. \qquad (14)$$

Large values of the two studentized residuals are related to outliers in the response-factor space and represent points not well fitted by the model.

(iv) Ratio between estimation variance in (7) and residual variance in (9):

$$\mathrm{RATIO}_i = \frac{w_{ii}p_{ii}}{1 - w_{ii}p_{ii}}; \quad i = 1, \ldots, n. \tag{15}$$

This statistic serves the same purpose as (12), but it is often more sensitive to detect leverage points.

(v) The standardized squared modulus of the difference between the vector estimate $\hat{\boldsymbol{\beta}}$ for the whole set of data and the same vector when the ith observation is omitted $\hat{\boldsymbol{\beta}}_{(i)}$:

$$\frac{1}{\hat{\sigma}^2}\left[\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(i)}\right]^{\mathrm{T}}\left[\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(i)}\right] = \frac{p^*_{ii}}{\hat{\sigma}^2}\left(\frac{w_{ii}\varepsilon_i}{1 - w_{ii}p_{ii}}\right)^2; \quad i = 1, \ldots, n, \tag{16}$$

where $p^*_{ii} = \mathbf{x}_i(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-2}\mathbf{x}_i^{\mathrm{T}}$. This measure is based on the sensitivity curve (Chatterjee and Hadi 1986).

(vi) The increase in the trace of the matrix $(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}$ after removing the ith observation:

$$\mathrm{trace}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}_{(i)} - \mathrm{trace}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1} = \frac{w_{ii}p^*_{ii}}{1 - w_{ii}p_{ii}}; \quad i = 1, \ldots, n. \tag{17}$$

Note that (16) is the product of (17) by z_i^2 given by (13).

(vii) The weighted squared standardized distance (WSSD; Daniel and Wood 1980) of the ith observation in the x space:

$$\mathrm{WSSD}_i = \frac{1}{s_y^2}\sum_{j=1}^{k}\hat{\beta}_j\left[\sqrt{w_i}x_{ij} - \overline{x}_j^{(w)}\right]^2; \quad i = 1, \ldots, n, \tag{18}$$

where s_y^2 is an estimate of σ^2 and $\overline{x}_j^{(w)} = \left(1/\sum_{i=1}^{n} w_{ii}\right)\sum_{i=1}^{n} w_{ii}x_{ij}$.

3) HETEROSCEDASTIC TRANSFORMATIONS

When the homoscedastic assumption (constant variance) does not hold, it is often possible to transform the response variable to stabilize the variance by using the transformation:

$$Z = g(Y) = \begin{cases} KY^{1-\gamma} & \text{if} \quad \gamma \neq 1 \\ K\log(Y) & \text{if} \quad \gamma = 1, \end{cases} \tag{19}$$

for some appropriate value of γ. This value of γ can be estimated using two different methods:

(i) Including the transformation in (19) within a nonlinear LS model. Thus, the estimated value $\hat{\gamma}$ is obtained jointly with the regression parameters.

(ii) Using repeated observations of the response variable Y at approximately the same point in the x space. The estimated parameter $\hat{\gamma}$ is obtained from fitting the model:

$$\log(\hat{\sigma}_{Y_i}) = \delta + \gamma\log(\hat{\mu}_{Y_i}) + \varepsilon_{Y_i}, \tag{20}$$

where $(\hat{\mu}_{Y_i}, \hat{\sigma}_{Y_i})$ are the estimated mean and standard deviation of Y for each set of repeated observations.

The second alternative is preferable, if one can find sets of repeated observations, because it allows using solutions given in section 3a. Consequently, heteroscedastic data can be analyzed using WLS, an appropriate transformation of the response variable, or a combination of both. We also show next that weights can be recalculated iteratively to match them with the observed standardized residuals.

b. *Reweighted least squares*

The aim of many outlier detection methods is to determine whether an observation should be considered as an outlier or not, without allowing for intermediate situations. In contrast, the reweighted least squares (RWLS) method aims at empirically determining a weight $0 \leq w_{ii} \leq 1$ for every observation ranging continuously from 0, for observations that are completely unreliable, and up to 1, for observations that are completely reliable. This can be attained by applying the following recursive procedure:

- Step 0: Set $w_{ii} = 1$; $i = 1, \ldots, n$.
- Step 1: Compute weighted least squares regression solving Eq. (2).
- Step 2: Compute new weights from the residuals of the last fit.

Steps 1 and 2 are repeated till convergence.

A key issue for the successful application of this algorithm is the new weight computation in step 2. From different formulas proposed in the literature (Huber 1981; Chatterjee and Mächler 1997; Luceño 1998b), we choose Tuckey's biweight:

$$w_{ii} = \begin{cases} \left[1 - \left(\dfrac{u_i}{6}\right)^2\right]^2 & \text{if} \quad |u_i| \leq 6, \\ 0 & \text{if} \quad |u_i| > 6 \end{cases} \tag{21}$$

where $u_i = (\varepsilon_i/\sigma^*)$ is a standardized residual based on the scaled median absolute deviation estimator $\sigma^* = (\mathrm{med}_i|\varepsilon_i|/c^*)$ of σ, with $c^* = 0.6745$ (for consistency of σ^*).

Within the RWLS scheme, outliers related to hurricanes and typhoons are characterized with low w_{ii} weights. Note that in addition to its multiple outlier detection capabilities, reweighting also provides a better performance on model estimation, because the influence of potential outliers is removed from the final estimates.

c. Nonlinear weighted least squares

Regression models presented previously allow the treatment of nonlinear and/or heteroscedastic problems using adequate transformations and/or weighting.

Tomás et al. (2008) and Mínguez et al. (2011) show that potential nonlinear relationships of the type $y_i = ax_i^b + \varepsilon_i$ and heteroscedastic variance $\mathrm{Var}(\varepsilon_i) = cx_i^d$ provide very good calibration results. For this reason, an outlier detection method based on a nonlinear heteroscedastic regression model is presented.

An intrinsically (nonlinearizable) nonlinear regression model can be written as

$$y_i = f_\mu(x_i; \boldsymbol{\beta}) + \varepsilon_i; \quad i = 1, 2, \ldots, n, \qquad (22)$$

where the function f_μ is known and nonlinear in the parameter vector $\boldsymbol{\beta}$, and ε_i; $i = 1, \ldots, n$ are jointly normally distributed $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{V})$ errors as in Eq. (1).

As in (2), the standard nonlinear weighted least squares (NWLS) method, for $\mathbf{W} = \mathbf{V}^{-1}$ diagonal, can be stated as

$$\underset{\boldsymbol{\beta}}{\mathrm{Minimize}}\, \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{W}\boldsymbol{\varepsilon} = \underset{\boldsymbol{\beta}}{\mathrm{Minimize}} \sum_{i=1}^{n} w_{ii}\left[y_i - f_\mu(x_i; \boldsymbol{\beta})\right]^2, \qquad (23)$$

where n is the number of observations. Note that analogously to the linear case, nonlinear regression models can also be used including weights in the formulation.

For wave hindcast data, a simple scatterplot of hindcast versus instrumental data allows observing how the variance of the regression model changes over the regression function. Consequently, we consider a nonlinear heteroscedastic regression model in which the standard deviation σ_i of the ith error is a function of the predictor variable (x_i):

$$\sigma_i = f_\sigma(x_i; \boldsymbol{\theta}) = w_{ii}^{-1/2}, \qquad (24)$$

where $\boldsymbol{\theta}$ is a new $s \times 1$ vector of coefficients or parameters. If the parameter vector $\boldsymbol{\theta}$ were known, estimation of the parameter vector $\boldsymbol{\beta}$ could be based on the NWLS method in (23). However, the values of $\boldsymbol{\theta}$ are usually unknown, and can be estimated using maximum likelihood methods. Thus, assuming that random errors are uncorrelated and normally distributed random variables each with mean zero and standard deviation given by (24), the whole set of model parameters ($\boldsymbol{\beta}$ and $\boldsymbol{\theta}$) can be jointly estimated maximizing the log-likelihood function:

$$\ell(\boldsymbol{\beta}, \boldsymbol{\theta}) = -\sum_{i=1}^{n} \log\left[f_\sigma(x_i; \boldsymbol{\theta})\right] - \frac{1}{2}\sum_{i=1}^{n}\left[\frac{y_i - f_\mu(x_i; \boldsymbol{\beta})}{f_\sigma(x_i; \boldsymbol{\theta})}\right]^2. \qquad (25)$$

The estimates $\hat{\boldsymbol{\beta}}$ that maximize the log-likelihood function in (25), and solve (23), allow calculating the residual vector $\hat{\boldsymbol{\varepsilon}}$, which is defined as

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - f_\mu(\mathbf{x}; \hat{\boldsymbol{\beta}}). \qquad (26)$$

Observe that the maximization of the log-likelihood function can be done using any of the available solvers for nonlinear programming, possibly subject to bounds on variables. One such solver is MINOS (Murtagh and Saunders 1998) under General Algebraic Modeling System (GAMS) (Brooke et al. 1998) which allows for upper and lower bounds on parameters to be estimated, and uses a reduced-gradient algorithm (Wolfe 1963) combined with the quasi-Newton algorithm described in Murtagh and Saunders (1978), or the Trust Region Reflective Algorithm under Matlab, also capable of dealing with upper and lower bounds through the function fmincon. For details about the method see Coleman and Li (1994) and Coleman and Li (1996). To improve convergence properties both the gradient and Hessian of the objective function are calculated analytically (see the appendix for details).

Following the analogy between WLS and NWLS, it is also possible to apply reweighting strategies within nonlinear regression models, which will enhance the quality of parameter estimates reducing the effect of possible existing outliers. This will lead to an increase in the computational time, or a somewhat more difficult to fit nonlinear regression model.

1) Residual covariance matrix and studentized residuals

Using a first-order Taylor series expansion of the function in (26) around the optimal estimated parameter vector $\hat{\boldsymbol{\beta}}$, the estimated differential residual vector is obtained as

$$d\hat{\boldsymbol{\varepsilon}} = d\mathbf{y} - \left.\frac{\partial f_\mu(\mathbf{x}; \boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} d\boldsymbol{\beta} = d\mathbf{y} - \mathbf{H}\,d\boldsymbol{\beta}, \qquad (27)$$

where $\mathbf{H}$ is the $n \times k$ Jacobian matrix evaluated at $\hat{\boldsymbol{\beta}}$. It readily follows that

$$\frac{\partial \hat{\boldsymbol{\varepsilon}}}{\partial \mathbf{y}} = \mathbf{I} - \mathbf{H}\left.\frac{\partial \boldsymbol{\beta}}{\partial \mathbf{y}}\right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = \mathbf{I} - \mathbf{H}\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}} = \mathbf{S}, \qquad (28)$$

where the $k \times n$ matrix $\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}}$ contains the derivatives of vector $\boldsymbol{\beta}$ with respect to $\mathbf{y}$ evaluated at $\hat{\boldsymbol{\beta}}$, matrix $\mathbf{I}$ is the n-dimensional identity matrix, and matrix $\mathbf{S}$ is the so-called residual *sensitivity* matrix.

Integration of (28) allows obtaining the first-order linear approximation to the (nonlinear in $\hat{\boldsymbol{\beta}}$) transformation (26) from $\mathbf{y}$ to $\hat{\boldsymbol{\varepsilon}}$:

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{S}\mathbf{y} + \mathbf{k}, \tag{29}$$

where $\mathbf{k}$ is the integration constant vector.

The corresponding estimated residual covariance matrix $\boldsymbol{\Omega} = \mathrm{Var}(\hat{\boldsymbol{\varepsilon}})$ is

$$\boldsymbol{\Omega} = \mathbf{S}\mathbf{C}_y\mathbf{S}^{\mathrm{T}}, \tag{30}$$

where matrix $\mathbf{C}_y$ is the error covariance matrix provided by (24):

$$\mathbf{C}_y = \begin{bmatrix} f_\sigma(x_1, \hat{\boldsymbol{\theta}})^2 & 0 & \cdots & 0 \\ 0 & f_\sigma(x_2, \hat{\boldsymbol{\theta}})^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & f_\sigma(x_n, \hat{\boldsymbol{\theta}})^2 \end{bmatrix}. \tag{31}$$

Therefore, considering (28), the general expression for matrix $\boldsymbol{\Omega}$ is

$$\boldsymbol{\Omega} = (\mathbf{I} - \mathbf{H}\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}})\mathbf{C}_y(\mathbf{I} - \mathbf{H}\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}})^{\mathrm{T}}, \tag{32}$$

where matrices $\mathbf{H}$ and $\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}}$ depend on the selected $f_\mu(\mathbf{x}; \boldsymbol{\beta})$ and $f_\sigma(\mathbf{x}; \boldsymbol{\theta})$ functions. Note that (32) is a nonlinear equivalent to (9).

Finally, from (26) and (32), studentized residuals are computed as

$$z_i = \frac{\hat{\varepsilon}_i}{\sqrt{\Omega_{i,i}}} = \frac{y_i - f_\mu(x_i; \hat{\boldsymbol{\beta}})}{\sqrt{\Omega_{i,i}}} \quad i = 1, \ldots, n, \tag{33}$$

where $\Omega_{i,i}$ is the ith diagonal element of $\boldsymbol{\Omega}$.

Vector $\boldsymbol{z}$ provides the studentized residuals, and hence can be used straightforwardly for outlier identification as in the linear case.

2) SENSITIVITY MATRIX FROM SENSITIVITY ANALYSIS

Section 3c(1) shows that the sensitivity matrix $\mathbf{S}$, which allows calculating the estimated residual covariance matrix $\boldsymbol{\Omega}$, which depends on matrix $\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}}$. This matrix is obtained below, based on sensitivity analysis results reported in Castillo et al. (2006).

For the maximum likelihood estimation problem, which is an unconstrained nonlinear optimization problem, the Karush–Kuhn–Tucker (KKT) first-order optimality conditions at its optimal solution $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}, \ell)$ (Bazaraa et al. 1993; Luenberger 1984) reduce to

$$\nabla_{\boldsymbol{\eta}}\ell(\hat{\boldsymbol{\eta}}, \mathbf{y}) = \mathbf{0}, \tag{34}$$

where $\hat{\boldsymbol{\eta}} = (\hat{\boldsymbol{\beta}}; \hat{\boldsymbol{\theta}})$, and $\nabla_{\boldsymbol{\eta}}$ stands for the vector of partial derivatives (of ℓ) with respect to $\boldsymbol{\eta}$.

This condition establishes that the gradient of the objective function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ at the optimal solution $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ must be zero.

To obtain sensitivity equations, we perturb or modify $\boldsymbol{y}$ so that $\hat{\boldsymbol{\eta}}$ is modified accordingly to continue satisfying the KKT conditions (Castillo et al. 2006). After manipulating the resulting expressions, the required sensitivity equation reduces to the following linear system of equations:

$$(-\mathbf{H}_{\boldsymbol{\eta}\boldsymbol{\eta}})\frac{\partial \boldsymbol{\eta}}{\partial \mathbf{y}}_{[(k+s)\times n]} = \mathbf{H}_{\boldsymbol{\eta}\mathbf{y}}, \tag{35}$$

where the vectors and submatrices in (35) are defined below (dimensions in parentheses):

$$\mathbf{H}_{\boldsymbol{\eta}\boldsymbol{\eta}[(k+s)\times(k+s)]} = \nabla_{\boldsymbol{\eta}\boldsymbol{\eta}}\ell(\boldsymbol{\eta}, \mathbf{y}), \tag{36}$$

$$\mathbf{H}_{\boldsymbol{\eta}\mathbf{y}[(k+s)\times n]} = \nabla_{\boldsymbol{\eta}\mathbf{y}}\ell(\boldsymbol{\eta}, \mathbf{y}), \tag{37}$$

which constitute Hessians with respect to parameters and data. Note that (36) is a nonlinear equivalent to (3) with unit weights.

Expression (35) allows deriving sensitivities of the parameter estimates with respect to the data. Under mild regularity conditions that are often satisfied (Coles 2001; Castillo et al. 2005) $-\mathbf{H}_{\eta\eta}$ (the Fisher information matrix) is invertible, and (35) has a unique solution. Matrix $\partial\boldsymbol{\eta}/\partial\mathbf{y}_{[(k+s)\times n]}$ can be partitioned in two different blocks associated with mean and standard deviation parameter functions, respectively:

$$\frac{\partial \boldsymbol{\eta}}{\partial \mathbf{y}}_{[(k+s)\times n]} = \begin{bmatrix} \dfrac{\partial \boldsymbol{\beta}}{\partial \mathbf{y}} \\ \dfrac{\partial \boldsymbol{\theta}}{\partial \mathbf{y}} \end{bmatrix}. \tag{38}$$

The first block corresponds to matrix $\mathbf{M}_{\boldsymbol{\beta}\mathbf{y}}$, which allows obtaining the sensitivity matrix $\mathbf{S}$ using Eq. (28).

From a computational point of view, inversion of the Hessian matrix $\mathbf{H}_{\boldsymbol{\eta}\boldsymbol{\eta}}$ is not needed because it can be easily factorized using LU algorithms. Sensitivities $(\partial\boldsymbol{\eta}/\partial\mathbf{y})$ are thus obtained using forward and backward elimination methods. Note that for the calculation of all sensitivities, second-order derivatives of the log-likelihood function with respect to parameters and data are needed. They can be obtained numerically by finite differences or analytically. For the analytical case, a detail derivation of Jacobians and Hessians with respect to parameters and data is given in the appendix. Although analytical results seem to be complex, we rather like this approach because it is easy to implement using any programming language, and it avoids possible numerical problems deriving from finite differences. In addition, to calculate studentized residuals, only the computation of the diagonal elements of the $\mathbf{\Omega}$ matrix is required, which considerably reduces the computational time.

d. Minimum covariance determinant estimator

A different method capable of detecting outliers is the minimum covariance determinant estimator (Rousseeuw and Van Driessen 1999), which is used in this paper for comparison purposes. The minimum covariance determinant estimator (MCD) method looks for the h observations out of n whose classical covariance matrix has the lowest possible determinant. This method allows us to calculate a robust distance:

$$\mathrm{RD}_i = \sqrt{(\mathbf{x}_i - \overline{\boldsymbol{\mu}}_{\mathrm{MCD}})\overline{\mathbf{\Sigma}}_{\mathrm{MCD}}^{-1}(\mathbf{x}_i - \overline{\boldsymbol{\mu}}_{\mathrm{MCD}})^{\mathrm{T}}}, \quad (39)$$

where $\overline{\boldsymbol{\mu}}_{\mathrm{MCD}}$ and $\overline{\mathbf{\Sigma}}_{\mathrm{MCD}}^{-1}$ are robust MCD location and scatter estimates, respectively, so as to determine whether the associated observation i is an outlier or not. Under the normal assumption, the outliers correspond to those values whose robust distances are larger than a given cutoff value usually defined as $\sqrt{\chi^2_{p,1-\alpha/2}}$ for some small $0 < \alpha < 1$. The robust distance in (39) is a robustification of the Mahalanobis distance.

4. Case study

In this section we illustrate the performance of the methods presented in section 3. We have applied them to the 43 buoys from the NDBC given in Table 1 and shown in Fig. 1. In this application we only deal with two variables: y_i corresponds to the ith value of the response variable (buoy data), and x_i is the predictor variable (interpolated hindcast data) corresponding to the ith observation. However, methods presented in the paper are valid for multivariate analysis. We could, for example, use more than one function of X in the regression Eqs. (1) or (22). Consequently, we have investigated some of these more complex models, but we will only show results for those models we have found to work best.

Before performing the analysis, the particular regression models we have chosen are presented:

- For the WLS method (section 3a), the response variable is transformed using Eq. (19) and the estimate $\hat{\gamma}$ is calculated based on Eq. (20). Because the relationship between X and Y is approximately linear, we apply the same power transformation $1 - \gamma$ to the covariate X and response Y, which leads to the following regression model:

$$Y^{1-\gamma} = \beta_0 + \beta_1 X^{1-\gamma} + \varepsilon. \quad (40)$$

  This model is linear with respect to β_0 and β_1 and nonlinear with respect to γ. However, because the estimate of γ is obtained previously rather than using a nonlinear iteration, model (40) can be considered linear for practical purposes.
- The RWLS method (section 3b) is applied using Eq. (40).
- For the NWLS model (section 3c), the following parameterization is used for the mean and dispersion functions:

$$f_\mu(x_i, \boldsymbol{\beta}) = \beta_0 x_i^{\beta_1}, \quad (41)$$

$$f_\sigma(x_i, \boldsymbol{\theta}) = \theta_0 x_i^{\theta_1}. \quad (42)$$

- Transformed data $Y^{1-\gamma}$ and $X^{1-\gamma}$ are also used within the MCD framework (section 3d).

Note that previous to deciding the particular regression model for each case, alternative expressions have been considered particularly to check whether other transformations of X and Y could be useful. We only provide those giving a better performance.

a. Detailed results for eastern Caribbean buoy 42059

We first analyze some detailed results for buoy 42059 (eastern Caribbean) shown in Fig. 2. We have applied the WLS (section 3a), RWLS (section 3b), NWLS (section 3c), and MCD (section 3d) methods. For the WLS and NWLS, outliers are identified using the internally studentized residuals z_i given in (13) and (33), respectively. In both cases, a case is identified as an outlier if $|z_i| > \Phi^{-1}(1 - \alpha/2)$. Results for different significance levels $\alpha = \{0.1, 0.05, 0.01, 0.001, 0.0001\}$ are shown in Figs. 3a,b, where data removed for each significance level are highlighted by using different dot marker specifiers. Table 2 also provides the number of data points detected as outliers for each significance level, and the computational time in seconds. Note that models have been run on a portable computer with one processor clocking at 2.39 GHz and 3.25 GB of RAM. From all these results the following observations are pertinent:

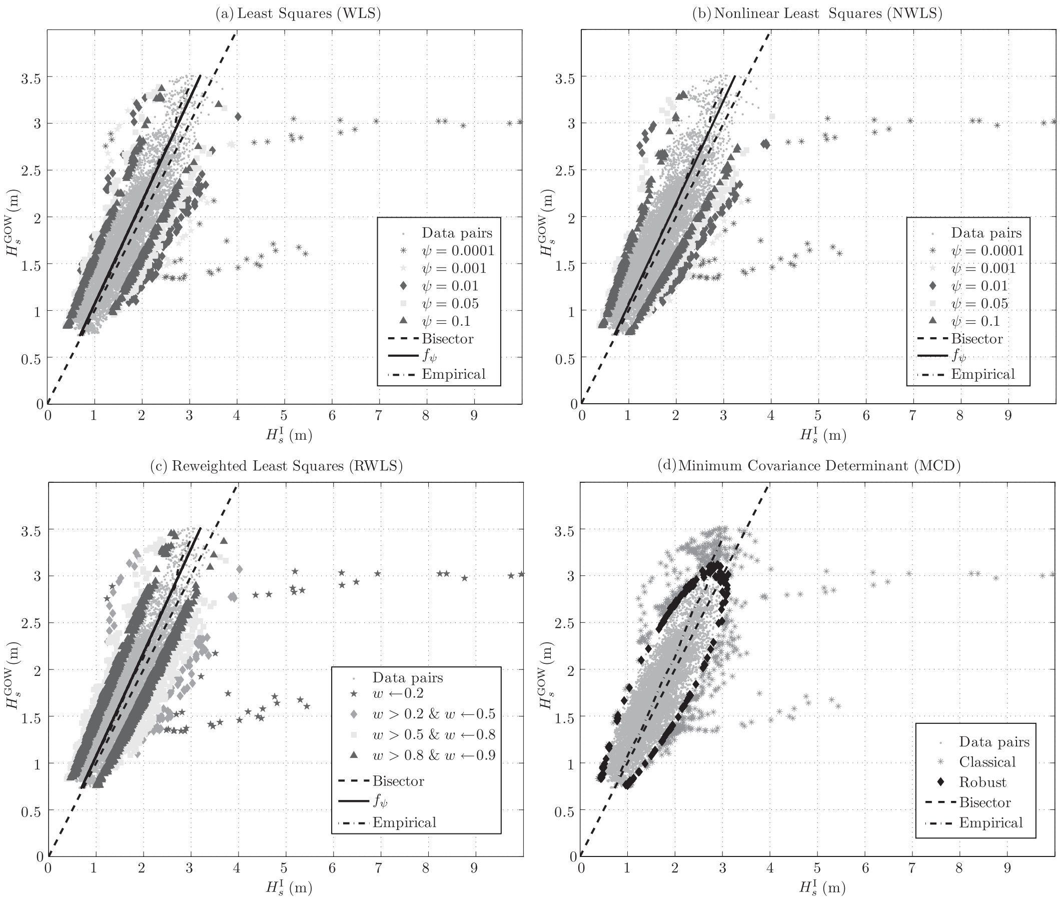

FIG. 3. Outlier detection performance at buoy 42059 (eastern Caribbean): (a) WLS, (b) NWLS, (c) RWLS, and (d) MCD.

(i) Both WLS and NWLS provide similar results. The numbers of outliers detected by the two methods for each significance level are almost the same.

(ii) The WLS method requires the evaluation of the optimal γ value in transformation in (19) for the homoscedastic assumption to hold, which for this particular buoy corresponds to $\hat{\gamma} \approx 0.41$. However, once this value is calculated, the problem is easily solvable using (3)–(11), which requires little computational time. On the other hand, the nonlinear version requires solving an optimization problem, which takes longer to solve although it is easily solvable using standard nonlinear mathematical programming techniques.

(iii) Since the RWLS method iteratively updates the weights associated with each case, the detection criterion is established as a function of the final weights w_{ii}. Outliers relevant for calibration purposes are those whose weights are lower than about 0.2 (note that $0 \leq w_{ii} \leq 1; i = 1, \ldots, n$). RWLS also appropriately detects the most relevant outliers (see Fig. 3c). The computational time increases slightly with respect to WLS, but decreases considerably with respect to NWLS. The iterative process

TABLE 2. Number of detected outliers from applying different outlier detection techniques on buoy 42059 (eastern Caribbean).

Method	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.001$	$\alpha = 10^{-4}$	CPU time (s)
WLS	1048	551	182	70	42	≈0.15
NWLS	965	523	181	70	45	≈0.5
	$0.8 < w \leq 0.9$	$0.5 < w \leq 0.8$	$0.2 < w \leq 0.5$	$w \leq 0.2$		CPU time (s)
RWLS	1645	819	79	41	—	≈0.19
	Classical			Robust		CPU time (s)
MCD	569			741		≈1

usually converges in a few iterations (e.g., for this particular buoy, it requires six iterations).

(iv) It is also important to realize that RWLS avoids the dichotomy outlier versus "not outlier" for each particular case in the sample. In contrast, the fitted regression line is estimated giving to each case a weight ($0 \leq w_{ii} \leq 1$) ranging from 0 to 1 according to our empirically determined degree of credibility on the goodness of each case in the sample.

(v) Hurricane data related to Dean (2007) and Omar (2008) (see Fig. 2), where the discrepancies are remarkable, are correctly detected with both WLS and NWLS methods using a significance level $\alpha = 0.0001$, as well as with the RWLS using a weight threshold of $w = 0.2$.

For comparison purposes, we have also applied the MCD approach (section 3d). Results are also given in Fig. 3d and Table 2. The MCD method is applied using the function mcdcov from MATLAB toolbox LIBRA (Verboven and Hubert 2005), which is an implementation of the fast-MCD algorithm proposed by Rousseeuw and Van Driessen (1999). Note that Fig. 3d shows the data detected using the classical approach based on Mahalanobis distance along with those for the robust approach. From these results, we can conclude that both methods (classical and robust) related to the MCD approach provide unsatisfactory results, since besides detecting data associated with outliers, they also eliminate extreme values of hindcast and instrumental distributions that are close to the regression line. These points are appropriately reproduced by the hindcast, and extremely important from the engineering design point of view. In addition, computational cost is much higher with respect to the other methods.

b. Results for the remainder buoys

Table 3 provides the following information related to the performance of the WLS and NWLS methods on the 43 buoys from the NDBC: the number of cases at each buoy location (n), the number of detected outliers for significance levels $\alpha_1 = 0.001$ and $\alpha_2 = 0.0001$ ($n_{\alpha_1}, n_{\alpha_2}$), the mean and standard deviation of the studentized residual absolute value ($|\overline{z}|$, $\sigma_{|z|}$), the maximum and minimum studentized residual absolute value ($|z|_{\max}$, $|z|_{\min}$), and the CPU time in seconds. Note that $|\overline{z}|$, $\sigma_{|z|}$, $|z|_{\max}$ and $|z|_{\min}$ are based on data removed using $\alpha_2 = 0.0001$.

From results given in Table 3 the following observations are pertinent:

(i) Both WLS and NWLS approaches provide satisfactory results on outlier identification in most cases, with the computational time required for WLS being lower.

(ii) In all buoys, and for the same significance level, the number of data points detected using the nonlinear approach is higher, and the maximum studentized residual absolute value $|z|_{\max}$ is also higher, resulting in a more conservative approach, which may produce better postcalibration results.

For comparison purposes, Fig. 4 shows the performance of both WLS and NWLS methods on three different buoys: 41040, 41046, and 41047. For buoy 41040 both methods perform appropriately, detecting the most relevant outliers. However, whereas the mean, standard deviation, and minimum studentized residual absolute value are similar (see the corresponding row in Table 3), the maximum studentized residual absolute values are 11.8317 and 19.2277, respectively, the NWLS $|z|_{\max}$ value being considerably higher. In this particular case, using $\alpha = 0.0001$, the performance can be considered equivalent from the postcalibration process perspective. This effect is also observed in buoys 41047 and 41046. In buoy 41047, the maximum studentized residual absolute values are 4.4024 and 5.6516, relatively close, but in 41046 the maximum studentized residual absolute values are 5.2096 and 9.3842, where differences increase considerably with respect to buoy 41047. On the other hand, the minimum studentized residual absolute values are very similar in both locations. Thus, NWLS method provides more conservative detection results at the $\alpha = 0.0001$ level, as shown in Figs. 4c,d, because it includes as outliers those points associated with H_s^{GOW} between 3 and 4 m, and H_s^{I} around 6 m. Note that both methods are also capable of detecting points associated with negative studentized residuals, as shown in Figs. 4a,b (left side of the regression lines), which may be related to points taken during disruption of normal use of the instrumental device.

An important contribution of our analysis is the assessment of the effect on outlier detection of using the different diagnostic statistics given in (12)–(18). We have confirmed that the most appropriate statistic for this particular application is the internally studentized residual, since the other statistics detect high leverage points that are not usually related to hurricanes. Regarding the differences between internally and externally studentized residuals, differences are negligible for the buoys considered.

Table 4 provides the following information related to the performance of the RWLS and NWLS methods on the 43 buoys from the NDBC: the number of cases at each buoy location (n), the number of detected outliers for weights holding $0.2 \leq w_1 \leq 0.5$ and $w_2 \leq 0.2$ (n_{w_1}, n_{w_2}), the number of detected outliers for significance levels $\alpha_1 = 0.001$ and $\alpha_2 = 0.0001$ ($n_{\alpha_1}, n_{\alpha_2}$), the

TABLE 3. Summarizing results from applying WLS and NWLS outlier detection techniques on the 43 buoys from the NDBC.

		WLS		NWLS		WLS	NWLS	WLS	NWLS	WLS	NWLS	WLS	NWLS	WLS	NWLS
ID	No.	n_{α_1}	n_{α_2}	n_{α_1}	n_{α_2}	$\lvert\overline{z}\rvert$	$\lvert\overline{z}\rvert$	$\sigma_{\lvert z\rvert}$	$\sigma_{\lvert z\rvert}$	$\lvert z\rvert_{\max}$	$\lvert z\rvert_{\max}$	$\lvert z\rvert_{\min}$	$\lvert z\rvert_{\min}$	time (s)	time (s)
41008	137 122	442	111	774	362	4.2757	4.7235	0.3421	0.7477	5.4274	7.3709	3.8945	3.8932	0.3438	10.2188
41003	15 037	79	21	113	58	4.2153	4.6897	0.2933	0.7558	4.8156	6.6724	3.9014	3.8907	0.1250	1.0781
41012	48 665	109	17	379	169	4.5718	4.5959	0.7011	0.9658	6.1662	10.9116	3.8985	3.8911	0.2188	3.5313
41009	273 795	1058	341	2057	1054	4.3736	4.7649	0.4484	0.8143	6.2315	8.5761	3.8907	3.8912	0.5938	18.6719
41006	98 052	389	99	693	316	4.3094	4.6585	0.3565	0.7724	5.2834	7.9569	3.8987	3.8944	0.2813	8.4688
41010	290 837	1048	292	2558	1372	4.4027	4.8687	0.4321	1.0359	6.0154	10.9560	3.8915	3.8909	0.7344	23.5938
42025	24 107	47	0	168	69	0	4.6992	0	0.7854	0	7.6880	0	3.9206	0.0938	1.7500
42039	109 759	539	209	829	474	4.3279	5.0119	0.3526	0.9298	5.4317	8.7355	3.8927	3.8922	0.2500	11.2656
42009	16 509	88	17	160	101	4.3553	4.9739	0.4329	0.9035	5.3190	7.8597	3.8908	3.8925	0.1250	2.0625
42040	108 092	598	235	902	523	4.6673	5.1866	0.8600	1.3959	8.6480	12.5488	3.8937	3.8912	0.3438	11.3594
44007	219 280	737	289	2194	1106	4.4072	4.8247	0.3573	0.9785	5.5247	12.5047	3.8906	3.8965	0.5625	17.9688
44005	203 184	309	60	1184	455	4.3134	4.5796	0.3847	0.7489	6.1976	9.7382	3.9137	3.8910	0.4375	20.6563
44013	191 121	363	131	1961	882	4.1446	4.9813	0.1878	1.3401	4.8251	13.4606	3.8971	3.8909	0.4375	17
44018	50 955	104	12	239	77	4.0787	4.3682	0.1351	0.5781	4.3345	7.6285	3.8945	3.9062	0.2031	4.4531
44011	182 806	355	72	789	344	4.1847	4.7747	0.5386	0.9145	8.2330	9.5811	3.8997	3.8995	0.4375	18.0781
44008	205 335	547	142	977	424	4.8552	4.8312	2.1657	1.0678	12.8360	12.3123	3.8929	3.8949	0.5781	16.6250
44001	9015	21	8	50	22	4.7839	4.7903	0.7567	1.4153	5.6723	10.3077	3.9115	3.9076	0.0313	1.0313
44012	35 014	179	51	337	204	4.2983	4.7619	0.2815	0.7615	5.4204	6.9618	3.9074	3.8920	0.1563	2.8125
44009	180 367	693	136	1606	824	4.2184	4.7491	0.2922	0.8653	5.4091	11.6203	3.8916	3.8911	0.4531	14.7344
44014	141 588	768	283	948	425	4.8632	4.7238	0.9744	0.9211	9.6952	12.3205	3.8922	3.8921	0.4219	10.8438
44006	9198	40	18	42	29	4.5780	5.1854	0.4624	1.0739	5.4003	7.6501	3.9254	3.9403	0.0625	1.0625
41001	187 253	564	199	1264	614	4.7033	4.8276	1.1646	1.1173	9.9324	16.7261	3.8906	3.8911	0.5000	17.8438
41036	45 367	108	33	317	141	4.2977	4.9978	0.3128	1.2102	5.0835	9.1222	3.9295	3.9023	0.1250	3.2813
41002	193 022	882	310	1507	765	5.4547	5.0393	1.9334	1.3083	10.7551	12.6517	3.8913	3.8913	0.5156	16.1406
41004	136 731	465	138	1083	570	4.7424	5.1960	0.9099	1.9037	7.9380	18.3539	3.8936	3.8916	0.3438	11.2031
41048	13 264	44	12	90	53	4.2602	5.0406	0.3263	0.9646	4.9236	7.8573	3.9011	3.9001	0.1250	1.8125
41047	9250	46	13	92	43	4.0779	4.4495	0.1490	0.4962	4.4024	5.6516	3.8907	3.8952	0.1250	1.1250
41046	9928	64	20	124	76	4.5710	5.3070	0.3987	1.4910	5.2096	9.3842	4.0068	3.8919	0.0313	1.2813
10000	955	10	5	16	10	4.4737	4.9588	0.2860	0.7059	4.9356	6.4495	4.2241	4.0774	0	0.2813
42001	236 281	975	339	1867	1002	4.6371	5.0152	0.8528	1.2876	9.0892	14.6672	3.8908	3.8924	0.5781	15.4531
42041	33 562	153	48	246	136	4.5476	5.0405	0.7737	1.3249	7.2164	10.6177	3.9025	3.8978	0.1094	2.7656
42038	16 537	105	19	92	54	4.3114	4.8445	0.3693	0.6728	5.3217	6.7321	3.8965	3.8995	0.1094	1.5625
42002	236 760	886	255	1500	777	4.5308	5.0328	0.6117	1.3841	6.9903	11.9223	3.8913	3.8936	0.5313	15.4688
42019	139 808	626	224	1038	540	4.5715	4.9334	0.7347	1.4301	7.7769	14.4643	3.8924	3.8914	0.3125	12.2188
42020	136 294	597	244	915	493	4.8074	5.2244	1.1559	1.8425	8.7640	15.3183	3.8913	3.8967	0.3906	11.9375
41041	31 298	146	68	192	90	5.5995	6.9746	1.9807	4.3726	10.4220	19.9469	3.9235	3.8970	0.0781	4.1875
41040	24 191	135	74	181	121	5.4958	5.4673	2.0014	2.4022	11.8317	19.2277	3.8994	3.8924	0.0938	3.0313
42059	14 135	70	42	70	45	6.9869	7.4736	2.8397	3.2642	14.3411	15.0454	3.9078	3.9323	0.0625	1.2344
41018	8669	11	1	17	1	3.9272	4.0830	0	0	3.9272	4.0830	3.9272	4.0830	0.0625	0.7813
42055	22 964	95	40	122	69	4.7890	5.1393	0.8298	1.1693	7.0481	8.8778	3.9155	3.9128	0.1250	2.2500
42056	30 195	172	120	315	202	6.9233	5.9229	2.3239	2.0479	11.7877	12.4083	3.9208	3.8909	0.1250	2.7188
42057	9602	131	101	148	116	4.9631	6.1350	0.5244	1.2292	6.4778	9.6274	3.9382	3.8964	0.0625	0.7813
42058	16 216	58	13	40	16	5.1281	5.6985	1.5274	2.1645	8.5397	11.0783	3.8929	3.9295	0.1250	1.9219

mean and standard deviation of the weights ($\overline{w}, \sigma_w$), the maximum and minimum weights ($w_{\max}$, $w_{\min}$), and the CPU time in seconds. Note that $\overline{w}, \sigma_w, w_{\max}, w_{\min}$ are for data removed using the $w_2 \leq 0.2$ criterion. This table also shows that the number of iterations required for convergence of the RWLS method is between 5 and 7, and is thus computationally faster than NWLS.

The number of outliers detected using RWLS (i.e., n_{w_1}) is very similar to the number detected using NWLS (i.e., n_{α_2}) with both methods capable of detecting all the relevant outliers. Differences are due to certain outliers detected by RWLS, which are related to lower values of the instrumental dataset. Figure 5 shows the performance of the RWLS method on buoys 41040, 41046, and 41047. Comparing these results with those in Figs. 4b,d,f, it can be observed that the outliers detected with NWLS using $\alpha = 0.0001$ and RWLS using $w < 0.5$, associated with hurricanes (higher values of the instrumental record), are almost the same. However, RWLS also includes data records related to the medium and lower part of the instrumental distribution, which are not considered as outliers by NWLS.

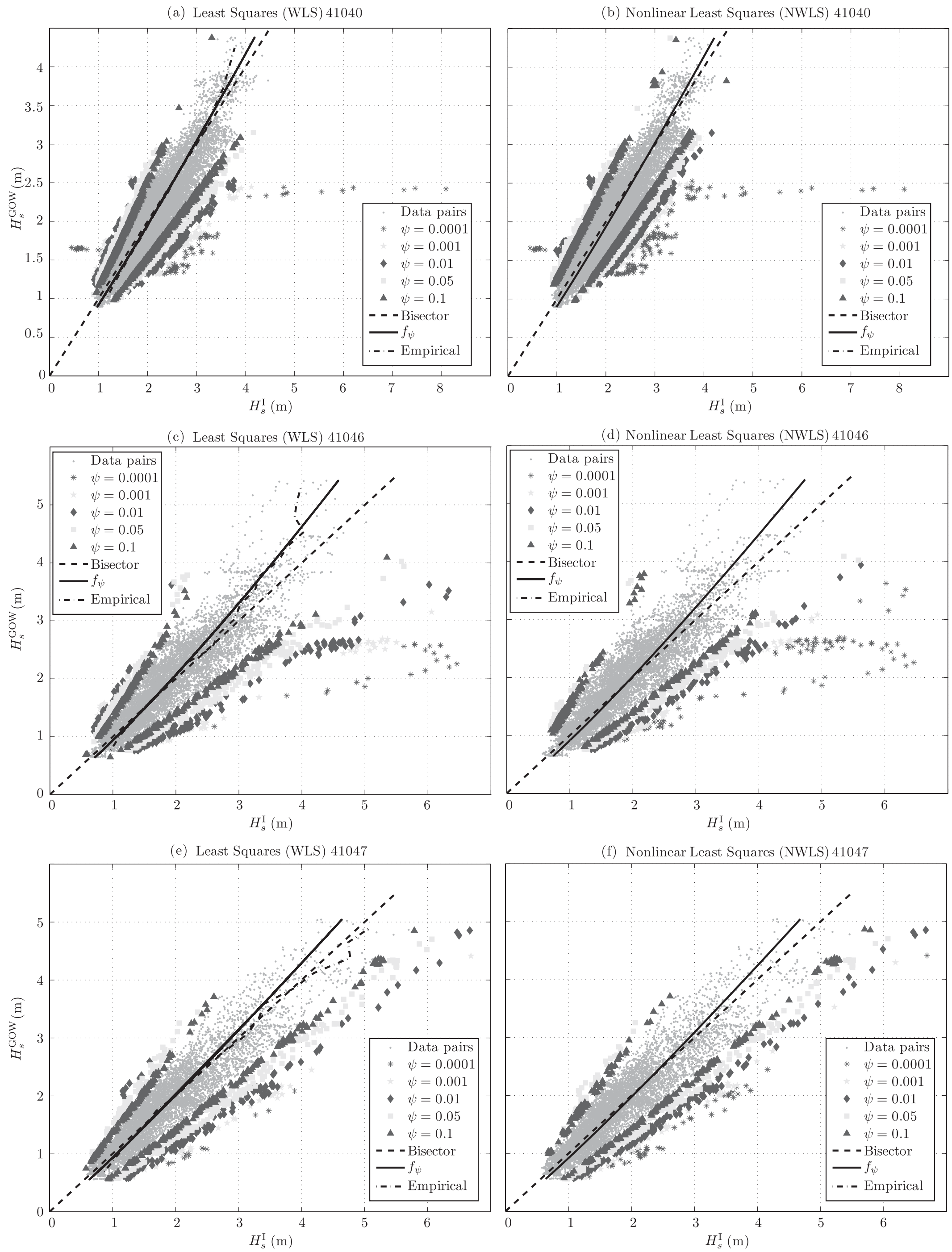

FIG. 4. Outlier detection performance at buoys 41040, 41046, and 41047 using WLS and NWLS methods.

Note that although computational time increases slightly with respect to WLS method, the RWLS detecting capabilities should be regarded as an insurance policy to obtain (i) better protection against outliers that are more difficult to detect, and (ii) better estimates for the model parameters, because suspected outliers are given small or null weights (see columns w_{max} and w_{min} in Table 4) depending on our belief in their true outlying nature.

5. Conclusions

Several methods for automatic ''outlier'' identification, when comparing wave hindcast versus instrumental

TABLE 4. Comparative results from applying RWLS and NWLS outlier detection techniques on the 43 buoys from the NDBC.

		RWLS		NWLS		RWLS						NWLS
ID	No.	n_{w_1}	n_{w_2}	n_{α_1}	n_{α_2}	$\overline{w}$	σ_w	$w_{\max}$	$w_{\min}$	Iter	Time (s)	time (s)
41008	137 122	514	29	774	362	0.1238	0.0502	0.1900	0.0207	5	2.2344	10.5469
41003	15 037	147	12	113	58	0.1151	0.0499	0.1886	0.0420	6	0.2969	1.2500
41012	48 665	154	10	379	169	0.1048	0.0794	0.1991	0	5	0.7656	3.4063
41009	273 795	1272	157	2057	1054	0.1166	0.0629	0.2000	0	6	3.5781	18.7188
41006	98 052	704	62	693	316	0.1166	0.0558	0.1991	0.0058	6	1.7188	8.4531
41010	290 837	1344	148	2558	1372	0.1148	0.0582	0.2000	0	6	3.9688	23.7656
42025	24 107	71	0	168	69	0	0	0	0	5	0.3594	1.6094
42039	109 759	565	91	829	474	0.1286	0.0495	0.1993	0.0134	6	1.5781	11.6719
42009	16 509	136	12	160	101	0.1150	0.0686	0.1936	0.0029	5	0.2031	2
42040	108 092	772	166	902	523	0.0933	0.0653	0.1993	0	5	1.1719	11.1563
44007	219 280	992	199	2194	1106	0.1160	0.0526	0.1973	0.0007	6	3.7656	18.2344
44005	203 184	304	15	1184	455	0.1445	0.0530	0.1985	0	5	2.8438	20.1719
44013	191 121	673	50	1961	882	0.1632	0.0296	0.1986	0.0677	7	3.2656	18.3125
44018	50 955	175	2	239	77	0.1889	0.0022	0.1905	0.1873	6	0.7969	4.2031
44011	182 806	418	7	789	344	0.1200	0.0717	0.1997	0	6	2.7969	18.0625
44008	205 335	755	53	977	424	0.1083	0.0751	0.2000	0	6	3.0469	17.1406
44001	9015	37	5	50	22	0.0346	0.0610	0.1429	0	5	0.1406	0.9219
44012	35 014	318	42	337	204	0.1293	0.0460	0.1962	0	6	0.6719	3.0625
44009	180 367	1061	56	1606	824	0.1449	0.0486	0.1998	0.0095	6	3.0938	14.4063
44014	141 588	1007	221	948	425	0.0664	0.0667	0.1986	0	6	2.1719	10.3125
44006	9198	31	9	42	29	0.1136	0.0387	0.1768	0.0483	5	0.1250	0.7969
41001	187 253	669	94	1264	614	0.0891	0.0711	0.1982	0	5	2.1719	17.7031
41036	45 367	104	12	317	141	0.1587	0.0359	0.1982	0.0773	5	0.6563	3.2344
41002	193 022	1309	230	1507	765	0.0708	0.0737	0.1995	0	6	3.2031	16.3438
41004	136 731	562	84	1083	570	0.0887	0.0726	0.1990	0	6	2.3750	10.9219
41048	13 264	39	2	90	53	0.1477	0.0280	0.1675	0.1279	6	0.2344	1.5781
41047	9250	78	5	92	43	0.1722	0.0240	0.1958	0.1329	5	0.1875	1.0938
41046	9928	157	26	124	76	0.0769	0.0719	0.1967	0	6	0.1875	1.2813
10000	955	20	11	16	10	0.0630	0.0745	0.1954	0	7	0.0625	0.1406
42001	236 281	1406	227	1867	1002	0.0870	0.0711	0.1974	0	6	3.0938	15.3281
42041	33 562	185	27	246	136	0.0985	0.0701	0.1974	0	5	0.4688	2.4688
42038	16 537	165	15	92	54	0.1092	0.0568	0.1950	0	5	0.2500	1.3438
42002	236 760	1083	136	1500	777	0.0966	0.0690	0.1983	0	6	3.8125	14.6094
42019	139 808	965	158	1038	540	0.0920	0.0669	0.1981	0	6	2.5000	12.0781
42020	136 294	737	163	915	493	0.0901	0.0714	0.1999	0	5	1.6875	12.2188
41041	31 298	202	59	192	90	0.0681	0.0711	0.1992	0	6	0.5000	4.0781
41040	24 191	160	70	181	121	0.0711	0.0661	0.1986	0	6	0.4063	3.3750
42059	14 135	79	41	70	45	0.0351	0.0605	0.1799	0	6	0.1875	1.6094
41018	8669	9	0	17	1	0	0	0	0	4	0.1250	1.0938
42055	22 964	109	33	122	69	0.0782	0.0704	0.1995	0	5	0.3906	2.2344
42056	30 195	206	137	315	202	0.0331	0.0570	0.1967	0	6	0.5625	2.6719
42057	9602	149	144	148	116	0.0366	0.0607	0.1975	0	7	0.1875	1.0469
42058	16 216	94	7	40	16	0.0545	0.0908	0.1904	0	5	0.2188	1.6563

time series, are analyzed and compared in this paper. We prove that these outlying data are mostly related to the presence of typhoons and/or hurricanes, which must be removed to avoid distorting postcalibration results. The main conclusions of the study are as follows:

(i) The best diagnostic statistic for outlier identification purposes in the WLS and NWLS methods is the internally studentized residual.

(ii) Both WLS and NWLS models perform appropriately in most cases. The WLS method is computationally faster; however, NWLS provides better postcalibration results because it is more conservative for the same significance level, which may be convenient if computational time is not relevant.

(iii) The RWLS method is also recommended for this specific application since it provides analogous results to NWLS. This method increases its relevance if there is a special interest on the final regression model parameters beyond outlier detection.

(vi) RWLS and NWLS provide systematic procedures to (i) detect outliers and (ii) remove outliers for

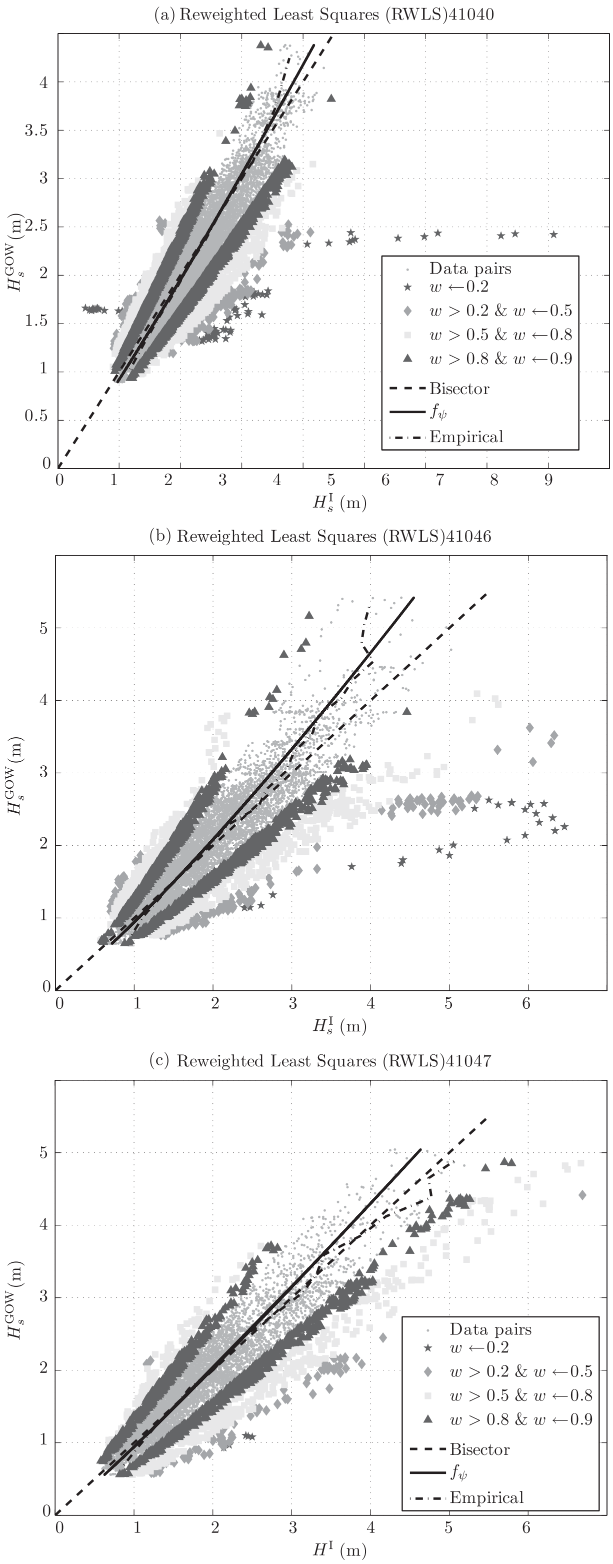

FIG. 5. Outlier detection performance at buoys 41040, 41046, and 41047 using RWLS method.

calibration purposes. In addition, NWLS allows us to identify those areas where the presence of hurricanes and typhoons is more relevant, which are related to high values of the maximum studentized residual. This is especially important if wave hindcast time series are intended to be used for engineering purposes.

(v) Methods based on the minimum covariance determinant (MCD) produce inappropriate results for this particular application. The main reason is the assumption of an underlying multivariate normal pattern that wave data do not follow, even after transforming the variables.

Note that our automatic hurricane–typhoon identification procedures allow detecting those areas and periods of time in which it is necessary to carry out a more accurate analysis by increasing the spatial and temporal resolution of winds during these events.

An open question is to assess the importance of using the proposed outlier detection techniques in new calibration studies. However, this effort is beyond the scope of the present paper.

Acknowledgments. This work was partly funded by projects ''GRACCIE'' (CSD2007-00067, Programa Consolider-Ingenio 2010) and ''AMVAR'' (CTM2010-15009) from the Spanish Ministry MICINN, by project C3E (200800050084091) from the Spanish Ministry MAMRM, and by project MARUCA (E17/08) from the Spanish Ministry MF. R. Mínguez is also indebted to the Spanish Ministry MICINN for the funding provided within the ''Ramon y Cajal'' program. Alberto Luceño acknowledges the support of the Spanish Grant MTM2008-00759. We also thank the editor and reviewers for their very helpful comments and suggestions, which have led to an improved manuscript.

APPENDIX

Derivatives for Sensitivity Matrix Calculations

The analytical derivation for all required matrices for the calculation of the sensitivity matrix is provided below. For this task, first- and second-order derivatives of the log-likelihood function with respect to parameters $\boldsymbol{\eta}$ at the optimum must be obtained. Note that all derivations are based on the chain rule.

a. First-order derivatives of the log-likelihood function

First-order derivatives of the log-likelihood function with respect to mean (μ) and standard deviation (σ) parameters are

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n}\left[\frac{y_i - f_\mu(x_i;\boldsymbol{\beta})}{f_\sigma^2(x_i,\boldsymbol{\theta})}\right]\frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_j}; \quad \forall j, \tag{A1}$$

$$\frac{\partial \ell}{\partial \theta_j} = -\sum_{i=1}^{n}\frac{1}{f_\sigma(x_i;\boldsymbol{\theta})}\frac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j} + \sum_{i=1}^{n}\frac{[y_i - f_\mu(x_i;\boldsymbol{\beta})]^2}{f_\sigma^3(x_i,\boldsymbol{\theta})}\frac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j}; \quad \forall j, \tag{A2}$$

where the derivatives of the functions f_μ and f_σ proposed in (41)–(42), and used in expressions (A1)–(A2) are

$$\begin{aligned}\frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_0} &= x_i^{\beta_1}; \quad \frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_1} = \beta_0 x_i^{\beta_1}\log(x_i); \quad \forall i\\ \frac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_0} &= x_i^{\theta_1}; \quad \frac{\partial f_{\sigma_1}(x_i;\boldsymbol{\theta})}{\partial \theta_1} = \theta_0 x_i^{\theta_1}\log(x_i); \quad \forall i.\end{aligned} \tag{A3}$$

b. Second-order derivatives of the log-likelihood function

Second-order derivatives of the log-likelihood function with respect to mean (μ) and standard deviation (σ) parameters are

$$\frac{\partial^2 \ell}{\partial^2 \beta_j} = \sum_{i=1}^{n}\frac{1}{f_\sigma^2(x_i,\boldsymbol{\theta})}\left\{[y_i - f_\mu(x_i;\boldsymbol{\beta})]\frac{\partial^2 f_\mu(x_i;\boldsymbol{\beta})}{\partial^2 \beta_j} - \left[\frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_j}\right]^2\right\}; \quad \forall j, \tag{A4}$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial \beta_l} = \sum_{i=1}^{n}\frac{1}{f_\sigma^2(x_i,\boldsymbol{\theta})}\left\{[y_i - f_\mu(x_i;\boldsymbol{\beta})]\frac{\partial^2 f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_j\,\partial \beta_l} - \frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_j}\frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_l}\right\}_{;\forall(j,l)}, \tag{A5}$$

$$\frac{\partial^2 l}{\partial^2 \theta_j} = -\sum_{i=1}^{n}\frac{f_\sigma(x_i;\theta)\dfrac{\partial^2 f_\sigma(x_i;\theta)}{\partial^2 \theta_j} - \left[\dfrac{\partial f_\sigma(x_i;\theta)}{\partial \theta_j}\right]^2}{f_\sigma^2(x_i;\theta)} + \sum_{i=1}^{n}\frac{[y_i - f_\mu(x_i;\beta)]^2}{f_\sigma^3(x_i;\theta)}\left\{\frac{\partial^2 f_\sigma(x_i;\theta)}{\partial^2 \theta_j} - \frac{3}{f_\sigma(x_i;\theta)}\left[\frac{\partial f_\sigma(x_i;\theta)}{\partial \theta_j}\right]^2\right\}; \quad \forall j, \tag{A6}$$

$$\frac{\partial^2 \ell}{\partial \theta_j \partial \theta_l} = -\sum_{i=1}^{n}\frac{f_\sigma(x_i;\boldsymbol{\theta})\dfrac{\partial^2 f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j \partial \theta_l} - \dfrac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j}\dfrac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_l}}{f_\sigma^2(x_i,\boldsymbol{\theta})} + \sum_{i=1}^{n}\frac{[y_i - f_\mu(x_i;\boldsymbol{\beta})]^2}{f_\sigma^3(x_i,\boldsymbol{\theta})}\left[\frac{\partial^2 f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j \partial \theta_l} - \frac{3}{f_\sigma(x_i;\boldsymbol{\theta})}\frac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j}\frac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j}\right]_{;\forall(j,l)}, \tag{A7}$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial \theta_l} = -2\sum_{i=1}^{n}\frac{[y_i - f_\mu(x_i;\,\boldsymbol{\beta})]}{f_\sigma^3(x_i,\boldsymbol{\theta})}\frac{\partial f_\mu(x_i;\,\boldsymbol{\beta})}{\partial \beta_j}\frac{\partial f_\sigma(x_i;\,\boldsymbol{\theta})}{\partial \theta_l}; \quad \forall(j,l), \tag{A8}$$

where the second derivatives of the functions f_μ and f_σ proposed in (41)–(42) are

$$\begin{aligned}&\frac{\partial^2 f_\mu(x_i;\boldsymbol{\beta})}{\partial^2 \beta_0} = 0; \quad \frac{\partial^2 f_\mu(x_i;\boldsymbol{\beta})}{\partial^2 \beta_1} = \beta_0 x_i^{\beta_1}\log^2(x_i); \quad \frac{\partial^2 f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} = x_i^{\beta_1}\log(x_i); \quad i = 1\\ &\frac{\partial^2 f_\sigma(x_i;\boldsymbol{\theta})}{\partial^2 \theta_0} = 0; \quad \frac{\partial^2 f_\sigma(x_i;\boldsymbol{\theta})}{\partial^2 \theta_1} = \theta_0 x_i^{\theta_1}\log^2(x_i); \quad \frac{\partial^2 f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_0 \partial \theta_1} = x_i^{\theta_1}\log(x_i); \quad i.\end{aligned} \tag{A9}$$

In addition, the evaluation of the second-order derivatives of the log-likelihood function with respect to parameters to be estimated and data $\mathbf{H}_{\boldsymbol{\eta}\mathbf{y}}$ is required. These are as follows:

$$\frac{\partial^2 \ell}{\partial \beta_j \partial y_i} = \frac{1}{f_\sigma^2(x_i,\boldsymbol{\theta})}\frac{\partial f_\mu(x_i;\boldsymbol{\beta})}{\partial \beta_j}; \quad j = 1, \ldots, k; \quad i = 1, \ldots, n, \tag{A10}$$

$$\frac{\partial^2 \ell}{\partial \theta_j \partial y_i} = \frac{2[y_i - f_\mu(x_i;\boldsymbol{\beta})]}{f_\sigma^3(x_i,\boldsymbol{\theta})}\frac{\partial f_\sigma(x_i;\boldsymbol{\theta})}{\partial \theta_j}; \quad j = 1, \ldots, s; \quad i = 1, \ldots, n. \tag{A11}$$

REFERENCES

Atkinson, A. C., 1984: Fast very robust methods for the detection of multiple outliers. *J. Amer. Stat. Assoc.*, **89,** 1329–1339.

——, 1985: *Plots, Transformations, and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Clarendon Press, 282 pp.

Barnett, V., and T. Lewis, 1994: *Outliers in Statistical Data*. 3rd ed. John Wiley & Sons, 604 pp.

Barrett, B. E., and J. B. Gray, 1997: Leverage, residual, and interaction diagnostics for subsets of cases in least squares regression. *Comput. Stat. Data Anal.,* **26,** 39–52.

Bazaraa, M. S., H. D. Sherali, and C. M. Shetty, 1993: *Nonlinear Programming: Theory and Algorithms*. 2nd ed. John Wiley & Sons, 656 pp.

Belsley, D. A., E. Kuh, and R. E. Welsch, 1980: *Regression Diagnostics: Identifying Influential Data and Sources of Multicollinearity*. John Wiley & Sons, 320 pp.

Billor, N., S. Chatterjee, and A. S. Hadi, 2001: Iteratively reweighted least squares method for outlier detection in linear regression. *Bull. Int. Stat. Inst.,* **1,** 470–472.

Brooke, A., D. Kendrick, A. Meeraus, and R. Raman, 1998: GAMS: A user's guide. GAMS Development Corporation, Washington, DC, 305 pp.

Caires, S., and A. Sterl, 2005: A new non-parametric method to correct model data: Application to significant wave height from the ERA-40 Re-Analysis. *J. Atmos. Oceanic Technol.,* **22,** 443–459.

——, ——, J. Bidlot, N. Graham, and V. Swail, 2004: Intercomparison of different wind-wave reanalyses. *J. Climate,* **17,** 1893–1913.

Cardone, V. J., E. G. Ward, and W. J. Pierson, 1976: Hindcasting the directional spectra of hurricane-generated waves. *J. Pet. Technol.,* **28** (4), 385–394.

——, R. E. Jensen, D. T. Resio, V. R. Swail, and A. T. Cox, 1996: Evaluation of contemporary ocean wave models in rare extreme events: The ''Halloween storm'' of October 1991 and the ''Storm of the Century'' of March 1993. *J. Atmos. Oceanic Technol.,* **13,** 198–230.

Castillo, E., A. S. Hadi, A. J. Conejo, and A. Fernández-Canteli, 2004: A general method for local sensitivity analysis with application to regression models and other optimization problems. *Technometrics,* **46** (4), 430–445.

——, ——, N. Balakrishnan, and J. M. Sarabia, 2005: *Extreme Value and Related Models with Applications in Engineering and Science*. John Wiley & Sons, 362 pp.

——, A. J. Conejo, C. Castillo, R. Mínguez, and D. Ortigosa, 2006: Perturbation approach to sensitivity analysis in nonlinear programming. *J. Optim. Theory Appl.,* **128,** 49–74.

Cavaleri, L., and L. Bertotti, 2004: Accuracy of the modelled wind and wave fields in enclosed seas. *Tellus,* **56A,** 167–175.

——, and M. Sclavo, 2006: The calibration of wind and wave model data in the Mediterranean Sea. *Coast. Eng.,* **53,** 613–627.

Chatterjee, S., and A. S. Hadi, 1986: Influential observations, high leverage points, and outliers in linear regression. *Stat. Sci.,* **1** (3), 379–393.

——, and ——, 1988: *Sensitivity Analysis in Linear Regression*. John Wiley & Sons, 315 pp.

——, and M. Mächler, 1997: Robust regression: A weighted least squares approach. *Comm. Stat. Theory Methods,* **26** (6), 1381–1394.

Cheng, T.-C., 2011: Robust diagnostics for the heteroscedastic regression model. *Comput. Stat. Data Anal.,* **55,** 1845–1866, doi:10.1016/j.csda.2010.11.024.

Coleman, T. F., and Y. Li, 1994: On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds. *Math. Programm.,* **67** (2), 189–224.

——, and ——, 1996: An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optim.,* **6,** 418–445.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer, 224 pp.

Cook, R. D., 1986: Assessment of local influence (with discussion). *J. Roy. Stat. Soc.,* **48B,** 133–169.

——, and S. Weisberg, 1982: *Residuals and Influence in Regression*. Springer, 240 pp.

Daniel, C., and F. S. Wood, 1980: *Fitting Equations to Data: Computer Analysis of Multifactor Data*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 458 pp.

Dodet, G., X. Bertin, and R. Taborda, 2010: Wave climate variability in the North–East Atlantic Ocean over the last six decades. *Ocean Modell.,* **31,** 120–131.

Draper, N. R., and H. Smith, 1981: *Applied Regression Analysis*. 2nd ed. Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., 709 pp.

Gray, J. B., 1986: A simple graphic for assessing influence in regression. *J. Stat. Comput. Simulat.,* **24,** 121–134.

——, and R. F. Ling, 1984: K-clustering as a detection tool for influential subsets in regression (with discussion). *Technometrics,* **26,** 305–330.

Hadi, A. S., and J. S. Simonoff, 1993: Procedures for the identification of multiple outliers in linear models. *J. Amer. Stat. Assoc.,* **88,** 1264–1272.

Hawkins, D. M., 1980: *Identification of Outliers*. Springer, 188 pp.

Hoaglin, D. C., and R. E. Welsch, 1978: The hat matrix in regression and ANOVA. *Amer. Stat.,* **32,** 17–22.

Huber, P. J., 1981: *Robust Statistics*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 320 pp.

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886-1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC 22, National Hurricane Center, 24 pp. [Available online at http://www.nhc.noaa.gov/pdf/NWS-NHC-1988-22.pdf.]

Jones, W. D., and R. F. Ling, 1988: A new unifying class of influence measures for regression diagnostics. *Proc. Stat. Comput. Sect.,* American Statistical Association, 305–310.

Krogstad, H. E., and S. F. Barstow, 1999: Satellite wave measurements for coastal engineering applications. *Coastal Eng.,* **37,** 283–307.

Landsea, C. W., and Coauthors, 2004: The Atlantic hurricane database reanalysis project: Documentation for the 1851–1910 alterations and additions to the HURDAT database. *Hurricanes and Typhoons: Past, Present, and Future,* R. J. Murnane and K. Liu, Eds., Columbia University Press, 177–221.

——, and Coauthors, 2008: A reanalysis of the 1911–20 Atlantic hurricane database. *J. Climate,* **21,** 2138–2168.

Luceño, A., 1997: Estimation of missing values in possibly partially nonstationary vector time series. *Biometrika,* **84,** 495–499, doi:10.1093/biomet/84.2.495.

——, 1998a: Detecting possibly non-consecutive outliers in industrial time series. *J. Roy. Stat. Soc.,* **60B,** 295–310.

——, 1998b: Multiple outliers detection through reweighted least deviances. *Comput. Stat. Data Anal.,* **26,** 313–326, doi:10.1016/S0167-9473(97)00036-4.

Luenberger, D. G., 1984: *Linear and Nonlinear Programming*. 2nd ed. Addison-Wesley, 491 pp.

Mayo, M. S., and J. B. Gray, 1997: Elemental subsets: The building blocks of regression. *J. Amer. Stat. Assoc.,* **51,** 122–129.

Mínguez, R., A. Espejo, A. Tomás, F. J. Méndez, and I. J. Losada, 2011: Directional calibration of wave reanalysis databases using instrumental data. *J. Atmos. Oceanic Technol.,* **28,** 1466–1485.

Murtagh, B. A., and M. A. Saunders, 1978: Large-scale linearly constrained optimization. *Math. Programm.,* **14,** 41–72.

——, and ——, 1998: MINOS 5.5 users' guide. Rep. SOL 83-20R, Department of Operations Research, Stanford University, Stanford, CA, 150 pp.

Nyquist, H., 1992: Sensitivity analysis in empirical studies. *J. Off. Stat.,* **8,** 167–182.

Paul, S. R., and K. Y. Fung, 1991: A generalized extreme studentized residual multiple-outlier-detection procedure in linear regression. *Technometrics,* **33,** 339–348.

Peña, D., and V. Yohai, 1995: The detection of influential subsets in linear regression by using an influence matrix. *J. Roy. Stat. Soc.,* **57B,** 339–348.

Pilar, P., S. Guedes Soares, and J. Carretero, 2008: 44-year wave hindcast for the North East Atlantic European coast. *Coast. Eng.,* **55** (11), 861–871.

Powell, M. D., S. H. Houston, L. R. Amat, and N. Morisseau-Leroy, 1998: The HRD real-time hurricane wind analysis system. *J. Wind Eng. Ind. Aerodyn.,* **77–78,** 53–64, doi:10.1016/S0167-6105(98)00131-7.

Pregibon, D., 1981: Logistic regression diagnostics. *Ann. Stat.,* **9,** 705–724.

Rousseeuw, P. J., and A. M. Leroy, 1987: *Robust Regression and Outlier Detection*. Wiley Series in Probability and Statistics, John Wiley & Sons, 352 pp.

——, and K. Van Driessen, 1999: A fast algorithm for the minimum covariance determinant estimator. *Technometrics,* **41,** 212–223.

Schwarzmann, B., 1991: A connection between local-influence analysis and residual diagnostics. *Technometrics,* **33,** 103–104.

Simonoff, J. S., 1991: Directions in robust statistics and diagnostics: Part II. *General Approaches to Stepwise Identification of Unusual Values in Data Analysis,* W. Stahel and S. Weisberg, Eds., Springer-Verlag, 223–242.

Tolman, H., 1997: User manual and system documentation of WAVEWATCH-III version 1.15. Tech. Note 151, NOAA/NWS/NCEP/OMB, 97 pp.

——, 1999: User manual and system documentation of WAVEWATCH-III version 1.18. Tech. Note 166, NOAA/NWS/NCEP/OMB, 110 pp.

——, 2002: User manual and system documentation of WAVEWATCH-III version 2.22. Tech. Note 222, NOAA/NWS/NCEP, 133 pp.

Tomás, A., F. J. Méndez, and I. J. Losada, 2008: A method for spatial calibration of wave hindcast data bases. *Cont. Shelf Res.,* **28,** 391–398.

Verboven, S., and M. Hubert, 2005: Libra: A matlab library for robust analysis. *Chemometr. Intell. Lab. Syst.,* **75** (2), 127–136, doi:10.1016/j.chemolab.2004.06.003.

WAMDI Group, 1988: The WAM model—A third generation ocean wave prediction model. *J. Phys. Oceanogr.,* **18,** 1775–1810.

Weissfeld, I., and H. Schneider, 1990a: Influence diagnostics for the normal linear model with censored data. *Aust. J. Stat.,* **32,** 11–20.

——, and ——, 1990b: Influence diagnostics for the weibull model fit to censored data. *Stat. Probab. Lett.,* **9,** 67–73.

Winsnowski, W. J., D. C. Montgomery, and R. S. James, 2001: A comparative analysis of multiple outlier detection procedures in the linear regression model. *Comput. Stat. Data Anal.,* **36,** 351–382.

Wolfe, P., 1963: Methods of nonlinear programming. *Recent Advances in Mathematical Programming,* R. L. Graves and P. Wolfe, Eds., McGraw-Hill, 76–77.