Calibration of ocean wave measurements by the TOPEX, Jason-1, and Jason-2 satellites

R. D. RAY

NASA Goddard Space Flight Center
Greenbelt MD 20771, USA

B. D. BECKLEY

SGT, Inc.
Greenbelt MD 20771, USA

Corresponding author address:

Richard D. Ray
NASA/GSFC, Code 698
Greenbelt, MD 20771
USA

tel: (301) 301-6102
email: richard.ray@nasa.gov

May 6, 2012

The calibration and validation of ocean wave height measurements by the TOPEX, Jason-1, and Jason-2 satellite altimeters is addressed by comparing the measurements internally among themselves and against independent wave measurements at moored buoys. The two six-month verification campaigns, when two of the satellites made near-simultaneous measurements along the same ground track, are invaluable for such work and reveal subtle aspects that otherwise might go undetected. The two Jason satellites are remarkably consistent; Topex reports waves generally 1–2% larger. External calibration is complicated by some systematic errors in the buoy data. We confirm a recent report by Durrant et al. that Canadian buoys underestimate significant wave heights by about 10% relative to U.S. buoys. Wave heights from all three altimetric satellites require scaling upwards by 5–6% to be consistent with U.S. buoys.

Keywords ocean waves, satellite altimetry, significant wave height, Jason-2 validation.

1 Introduction

The subject of this work is the validation of significant wave heights H_s measured by the *Jason-2* satellite altimeter (Lambin et al., 2010), as well as its predecessor missions *TOPEX/Poseidon* and *Jason-1*, and the possible recalibration of these data to form a consistent 20-year time series of global wave data. Validation of the wave-height data from *Jason-2* has already been reported, at least in preliminary fashion, in several studies (e.g., Ablain et al. 2010; Abdalla et al., 2010; Quartly, 2010). Further and continuing analyses are warranted for a number of reasons. Foremost,

the use of altimetric time series to study long-term climatic trends (e.g., Young et al., 2011) requires that data be routinely monitored for quality and especially for drift. In addition, different approaches to the validation-calibration problem, such as satellite-satellite comparisons versus *in situ* buoy comparisons, can be mutually complementary and are therefore certainly warranted. In the present paper extensive use is made of the extremely valuable "verification campaign" data in which two successive satellites made near-simultaneous measurements along the same ground track. Such simultaneous measurements are unique opportunities to compare and calibrate overlapping missions, and many subtle aspects of altimetry were discovered during the first such verification campaign for *Jason-1*. To analyze satellite wave data collected outside these special campaigns, extensive use is made here of *in situ* wave data collected at a network of meteorological buoys. Buoy measurements, however, are themselves subject to certain systematic errors, not all of them understood (e.g., Durrant et al., 2009; Swail et al., 2009; Gemmrich et al., 2011). The inconsistencies noted by Durrant et al. (2009) especially call into question some previous altimeter calibration results.

In some cases different published calibration data for nominally the same satellite have been inconsistent because based on different released versions of the data. These differences may stem from errors in earlier releases, or changes in various correction tables, or especially from use of different waveform retracking algorithms. Some authors have analyzed only operational or real-time data because of the importance of these data for assimilation into operational wave forecasting systems. Other authors have concentrated on the highest quality (usually retracked) climate data records, for use in non-operational and climate studies. The present paper takes the second approach. For *Jason-2*, at least for the present time, this issue does not arise because the two types of records are identical (aside from minor variations related to interpolation along track—see discussion by

Lillibridge et al., 2011). The issue does arise for the earlier satellites.

This paper considers the entire *TOPEX*, *Jason-1*, *Jason-2* time series, with special emphasis given to the newer *Jason-2* data. Such a high-quality time series, now two decades in duration, with a fixed ground-track pattern covering nearly the whole global ocean, is ideal for wave-climate studies, and ensuring the consistency of these data and monitoring their quality over time should be an ongoing and essential task.

2 Data

The *Jason-2* Geophysical Data Records (GDRs) used here correspond to the so-called GDR-T product, usually referred to as simply the GDR (Lambin et al., 2010). The H_s data in these GDRs are also essentially equivalent to those available (as of this writing) on the Operational GDRs and Interim GDRs (Scharroo et al., 2009; Lillibridge et al., 2011). The waveform tracking algorithm used for all three products is the so-called MLE4 algorithm (Thibaut et al., 2010).

The *Jason-1* data used here are from the GDR-C release. Those H_s data are also based on the MLE4 tracking algorithm. Note that in our earlier work on *Jason-1* (Ray and Beckley, 2003), we used the IGDR-A, in which the wave data were based on MLE3 tracking. The different waveform algorithms do result in small differences in the H_s data, as well as in the normalized backscatter coefficient (Quartly, 2009), and our statistics regarding *Jason-1* differ somewhat from our earlier analysis. Note that GDR-B, which also employed MLE4 retracking, differs from GDR-C owing to some updated instrument correction tables (see AVISO PODAAC User Handbook, 2008); the largest adjustments were for wave heights around 6 m.

The *TOPEX* data (the much smaller dataset from the Poseidon altimeter will not be considered) are based on the MGDR-B product from the physical oceanography data center at the Jet Propul-

sion Laboratory. These data are still consistent with those used in the earlier Ray-Beckley analysis. A limited amount of *TOPEX* data has been retracked by P. Callahan (JPL), and this retracking does result in improved statistical characteristics of the wave data (Ray and Beckley, 2003), such as a more physically realistic 1-D probability density function. At this writing, however, the full dataset has not been retracked, so only the MGDR-B data are used here. For the buoy-altimeter statistical comparisons below, the *TOPEX* data for part of 1997–1999 (repeat cycles 160–236) are not used because those wave data are known to have drifted markedly (e.g., Queffeulou, 2004).

The buoy data used in Section 4 are extracted from two sources: (1) the National Data Buoy Center (NDBC), part of the U.S. National Oceanic and Atmospheric Administration (Hamilton, 1986; Teng, 2002), (2) the Department of Fisheries and Oceans, Canada. The buoy measurements from these two sources are not entirely consistent (Durrant et al., 2009; and below).

3 Analysis of Verification Campaign Data

3.1 Example profiles

This section is restricted to analyzing data collected during the verification campaign for *Jason-2* (July 2008 to January 2009), as well as a reanalysis of similar data for *Jason-1* (late January through mid-August 2002). During these periods, each lasting about six months, the two satellites—either the two Jason satellites or *Jason-1* with *TOPEX/Poseidon*—flew in tandem formation along the same ground track. (Here 'tandem' is in the sense of a tandem bicycle, one immediately behind the other; usage of the term within the altimeter community is inconsistent.)

The general characteristics of the H_s data can be seen in Figure 1 for a partial track about 6000 km long across the middle Pacific Ocean. There is considerable fine structure apparent in this profile, with a rich spectrum of features including a sharp front near time -500 s. A high-

frequency noise envelope of order 10 cm is apparent. The two profiles from the two altimeters are very consistent, and only very close examination reveals small differences, mostly at short scales of about 20 s or shorter. Differencing the two H_s profiles (lower panel) shows no evident broad-scale differences; the RMS difference is 18 cm.

The data shown have undergone fairly standard tests for data quality (see GDR Handbook), including tests of various instrument hardware flags, quality flags for the radiometer brightness temperatures, rain flags, and preset land-mask checks. Data editing is otherwise fairly loose: mean 1-Hz data values must be formed from at least four 20-Hz measurements; the standard deviation of the 20-Hz H_s data may not exceed 1.5 m; the backscatter coefficient σ^0 must fall in the range 6–27 dB. The data gap in Figure 1 near time −160 s is a result of the rain-flag check, without which there would be a small cluster of outliers present with errors reaching about one meter.

The data in Figure 1 were obtained at the primary Ku-band frequency. Figure 2 shows corresponding data at C-band. As is well known (e.g., Thibaut et al., 2004), the C-band altimetric wave data are considerably noisier; the RMS difference in this case between *Jason-1* and *Jason-2* is 44 cm. For the remainder of this paper, only the more precise Ku-band data will be considered.

Figure 3 shows another H_s profile from the *Jason-2* verification period, which is of interest because it highlights the relationship between outliers in H_s and "blooms" (the term was invented by the late George Hayne) of σ^0, the normalized backscatter coefficient. These blooms are regions of abnormally enhanced radar cross-section, and in most instances appear to be over regions of very light winds (Mitchum et al., 2004). Waveform tracking algorithms built on the standard rough-surface scattering model tend to break down when applied to these blooms, leading to errors in estimated sea-surface heights and wave heights. In Figure 3 the bloom near 26°S is seen to be associated with large H_s outliers in both *Jason-1* and *Jason-2*. A properly designed edit on

σ^0, built perhaps on detecting localized along-track peaks, might possibly handle many of these cases, but we have not attempted it. Most cases also escape detection by testing on the variance of the 20-Hz data, which tends to be only slightly inflated inside these blooms, and generally within our standard 1.5-m test. We have instead found that application of a median filter, of about 7 1-Hz samples full-width, removes many of these cases and also profitably reduces some of the very short-scale noise evident in the H_s data. Note that, in our experience, *TOPEX* H_s data are far less affected by σ^0 blooms—see Figure 3 of Ray and Beckley (2003) which shows an example from the *Jason-1* verification campaign, for the same ground-track as in Figure 3 and with a σ^0 bloom in very nearly the same location.

A somewhat analogous problem that typically arises over regions displaying moderate to large σ^0 concerns the multiple appearance of wave heights set precisely to zero. Figure 4a shows an example for a *Jason-2* pass over the Mediterranean Sea. For most of this pass σ^0 is in the range 12–16 dB, normal to moderately high, the latter presumably indicating reflection off a calm sea. According to the OSTM/Jason-2 User Handbook (2009), zero H_s values can be valid data, but in this case they appear to be invalid, and using these data to smooth along track would be inadvisable. Thibaut et al. (2004) state that such zeroes arise from waves estimated as negative but truncated to zero. Aside from deleting such data entirely, which would tend to limit severely the data returns in some regions such as the Mediterranean, it is not always clear how best to handle such problems. For the purposes at hand, our approach is simply to delete all data with $H_s = 0$ and leave intact the neighboring data; such a course at least avoids skewing any statistical distributions near zero wave height.

Thibaut et al. (2004) present H_s histograms for *Jason-1* with a large spike at zero, implying a fairly significant number of these zero values. For the period of the verification campaign we find

0.56% of the *Jason-2* data and 0.39% of the *Jason-1* data are exactly zero. There is a pronounced geographical bias, with most zero data occurring in regions tending to have fairly low wave height: shallow or enclosed seas as well as the Gulf of Mexico, the Caribbean, the South China Sea, and the Indonesian seas, extending somewhat into the Coral and Philippine Seas. By comparison, the *TOPEX* altimeter data appeared far less affected. Figure 4b,c shows H_s along the same Pass 44, for *Jason-1* cycle 16 and *TOPEX* cycle 359, both collected during the *Jason-1* verification phase. Although there are a few obvious outliers in *TOPEX* near the coast of Sicily, the data otherwise appear quite satisfactory, whereas the *Jason-1* profile is considerably corrupted, and in a manner similar to that shown in the top panel.

3.2 Global statistics

Table 1 lists wave height statistics derived from 1-Hz along-track measurements during both the *Jason-1* and *Jason-2* verification campaigns. The data are global, but somewhat over-weighted toward higher latitudes owing to the pattern of the satellite ground-tracks (with turning latitudes at 66°). The statistics for the two campaigns must also be expected to differ somewhat owing to different seasons of observation.

The mean wave heights for *Jason-1* and *Jason-2* are very close, differing only by 1.2 cm (in good agreement with that reported by Ablain et al., 2010). The higher moments of the H_s distributions are also close. The *Jason-1* and *TOPEX* wave heights are slightly less consistent.

These aspects are brought out more clearly by the observed joint probability densities shown in Figure 5. The Jason data are very consistent, with all density contours centered almost exactly on the unit slope. In contrast there is a small tendency for *TOPEX* wave heights to be slightly higher than *Jason-1*, which was also the case reported earlier based on using *Jason-1* IGDR data

with a different retracking algorithm (Ray and Beckley, 2003). The left panel also shows a subtle bowing of contours around 3 and 6 m, which arises from an imperfection in the *TOPEX* density function—a dearth of waves at 2.9 m and an apparent shift of 6-m waves toward higher elevations (Figure 4 of Ray and Beckley, 2003; also Quartly, 2000); both imperfections can be corrected by retracking.

Note that the data used to construct these figures, unlike the data used elsewhere in this section and below, have not been subjected to median smoothing. If median filtering is applied, the contours grow much tighter together, reflecting suppression of high-frequency noise. For the data of Figure 5 the RMS differences between satellites are (left panel) 22 cm and (right panel) 21 cm, while median filtering reduces these differences (as in Table 1) to 11.5 cm and 12.1 cm, respectively.

A linear fit to the joint Jason data, based on Figure 5, would be almost precisely unity, and an explicit regression calculation confirms this to 4 significant figures. However, the *TOPEX* and *Jason-1* density clearly peaks off the diagonal; the crest appears fairly linear except very near the origin.

To determine a linear fit to the joint *TOPEX* and *Jason-1* density function, one must account for errors in both variables (e.g., Madansky, 1959), which classical regression methods do not. For the case at hand, a method treating each variable symmetrically is appropriate, and we use orthogonal regression (e.g., Isobe et al., 1990), which minimizes the sum of squares of the orthogonal distances from the data points to the fitted line. Note, however, that the H_s differences described in Table 1 show an exceedingly large kurtosis, or a very long tail presumably reflecting the presence of some large outliers. So before computing the linear regression we have employed a 1% trimming (Everitt, 2002) of the differences, which reduces the kurtosis to near zero. The resulting estimated

linear relationship is: slope $\beta_1 = 1.017$, intercept $\beta_0 = 2.7$ cm, with standard errors smaller than the quoted precision. (The value of β_1 given in the figure is without median filtering.) This slope is significantly smaller than the 1.030 quoted by Ray and Beckley (2003), but there are several differences in the input data: the use of *Jason-1* GDR-C rather than IGDR, the existence of roughly 10% more data than before, and somewhat different outlier detection schemes.

The main point of the present section is that the two 6-month verification campaigns show (1) that the two Jason wave-height datasets are very consistent and can safely be used to form a continuous time series across the two missions, and (2) that the *TOPEX* wave heights are approximately 2% larger than Jason heights, which should be accounted for before the datasets are combined. These results, of course, strictly apply only to the two 6-month periods of data. The next section uses buoy data to extend the analyses over the whole period of available altimeter data.

4 Buoy Comparisons

Comparison of buoy measurements of significant wave height to altimeter measurements from nearby ground tracks has long been a standard approach to altimeter calibration (Lefèvre and Cotton, 2001, and references therein). Our previous efforts with *Jason-1* data (Ray and Beckley, 2003) were based on only six months of altimetry, and subsequently published calibrations are presumably more reliable since based on considerably more data. In the following Section 4.1, data from all three subject altimeters are reanalyzed. The results are bulk calibration coefficients for each entire mission. Section 4.3 takes a second approach in order to monitor possible changes in the measured H_s data over time by employing a technique devised by Mitchum (2000) to monitor drift in altimetric sea-level data.

The use of buoys for this task necessarily has certain limitations. With most buoys located fairly near the North American continent, geographical limitations imply that the statistical distribution of observed waves is not completely representative of the global environment. This implies that regression results, which can be sensitive to the statistical distribution of the data (especially if the ratio of height measurement error to height standard deviation is moderately large), may be similarly limited and not strictly applicable to the global ocean. A related aspect is the lack of large H_s data in the buoy data, such as those more routinely observed in the Southern Ocean. Finally, the use of buoy data can be impacted by their own systematic errors (Durrant et al., 2009). All these limitations should be kept in mind in the following discussion.

All deep-ocean buoys from the NDBC and Canadian on-line archives were used in the following analyses except for buoys deemed too close to land and therefore potentially subject to either land contamination in the altimeter beam pattern or to large H_s spatial gradients from rapid bathymetric gradients. A total of 44 buoys were used, although the number returning data at any given time was always fewer. Eight of the buoys are from the Canadian network. Buoy locations are shown in Figure 6.

4.1 Buoy-altimeter statistics

As in many previous studies (e.g., Dobson et al., 1987; Lefèvre and Cotton, 2001; Zieger et al., 2009; Durrant et al., 2009), we have constructed a dataset of joint altimeter and buoy near-simultaneous measurements of H_s. Compiling such near-collocated data requires acceptance criteria for temporal and spatial separation. Time is of little concern, since hourly buoy data ensures separation times no longer than 30 minutes and interpolation in time further reduces error. The criteria for spatial separation is often based on wave-correlation statistics derived by Monaldo (1988) and generally

restricted to distances less than 50 km, or sometimes 100 km.

We experimented with relaxing the maximum distance between buoy and altimeter track to 150 km. For each buoy and for each of its neighboring tracks falling within 150 km, the RMS difference for all valid passes of the *Jason-1* altimeter was computed. These RMS values are plotted in Figure 7 as a function of distance between the track and buoy. Although many buoy-track pairs have RMS less than 30 cm even at large distances, the figure does show a general inflation of larger RMS with distance. Based on this figure we gauged 70 km as the maximum distance we would tolerate. Note that this does eliminate some buoys completely from our analysis because they are more than 70 km from all groundtracks (e.g., buoys 41002, 41048, 44011, 46006 are eliminated).

It may be of interest (and mild amusement) to examine the largest discrepancy (124 cm) shown in Figure 7. This arose from buoy 46072, whose location is shown along with its four closest altimeter tracks in Figure 8. The anomalous discrepancy is with Track 126, and one notices the closest point along that track falls on the opposite side of the Aleutian Island chain! It is thus not surprising to see large disagreement between this track and the buoy. (Our criterion of 70 km does eliminate this pair.)

Figure 9 shows a scatter plot of 3452 pairs of altimeter-buoy H_s measurements for *Jason-2*, based on 29 NDBC buoys, with RMS difference 22.9 cm and Pearson correlation coefficient 0.983. The agreement must be judged very good, but one can detect even by eye a slight systematic difference in the highest wave heights, with the altimeter yielding slightly smaller heights. Altimeters *TOPEX* and *Jason-1* yield similar diagrams.

Table 2 summarizes statistics for all three altimeters and the corresponding coefficients of a linear fit to the data. If one wishes to adjust the altimeter data to agree more closely with the

buoy data, then the adjusted data would be given by

$$H_s^{\text{corrected}} = \beta_1 H_s + \beta_0, \tag{1}$$

using coefficients from this table.

It is by now widely appreciated within the community that these kinds of linear fits to the buoy-altimeter data should not be done with classical linear regression methods (i.e., ordinary least squares or some robust variation), which assume no measurement error in the independent variable and generally yield slopes biased small (Madansky, 1959). As in Section 3 orthogonal regression has here been used to determine the β_1, β_0 coefficients. In fact, the appropriate fitting method depends on the relative magnitudes of the measurement errors, or at least their ratio (Cheng and Van Ness, 1999), and orthogonal regression is warranted only if this ratio is near unity, thus treating both variables symmetrically—an approach that appears more justified for the altimeter-altimeter fits of Section 3. Nonetheless, several studies employing triple-collocation analyses (Caires and Sterl, 2003; Abdalla et al., 2011) have concluded that the buoy and altimeter measurement errors are roughly comparable. Based on this evidence orthogonal regression is a reasonable approach. Zieger et al. (2009) employ a technique sometimes called reduced major axis (RMA) regression, which minimizes the triangular areas formed by vertical and horizontal projections from each point to the line. Isobe et al. (1990) usefully tabulates formulae for these various kinds of regressions. For the *Jason-2* data RMA yields a slope $\beta_1 = 1.0519$ versus 1.0528 for orthogonal regression, a difference of little consequence and less than the standard error (Table 2).

The coefficients for *TOPEX* in Table 2 are consistent with our earlier estimates (Ray and Beckley, 2003) within their error limits (which were previously ± 0.014 on β_1, not 0.0014 as stated

by Queffeulou, 2004). Our new results for *Jason-1* are not consistent with our previous results, but, as noted above, these were for IGDR data (and for a very small sample size of 368).

The effects of linear adjustment of the altimeter data, as well as insight into the extremes which are important in climatological studies, may be better assessed with quantile (or Q-Q) plots (e.g., Wilk and Gnanadesikan, 1968), as shown in Figure 10. These diagrams are based on the same collocated dataset used for the regressions. The agreement between buoys and altimetry at all percentiles is excellent after adjustment. Close examination even reveals better agreement at the first percentile for the very smallest waves.

4.2 Buoy inconsistencies

So far the discussion has ignored the last row of Table 2, which represents buoy-altimeter statistics derived from using 8 buoys of the Canadian network (3 in the Pacific, 5 in the Atlantic). Unlike the NDBC buoys, the Canadian buoys tend to observe wave heights smaller, rather than larger, than the altimetry, and the resulting regression yields a slope less than unity. Systematic differences between national buoy networks were first reported, to our knowledge, by Challenor and Cotton (2003) who also noted systematic differences with British and Japanese buoy networks. Durrant et al. (2009) recently concluded that Canadian buoys were underestimating wave heights relative to the NDBC buoys by about 10%, and our slope estimates shown in Table 2 for *Jason-1* confirm their result.

Figure 11 attempts to shed further light on these systematic differences by showing estimates of the orthogonal regression coefficient β_1 derived from each individual buoy over the *Jason-1* mission (2002 through 2008). All Canadian-based coefficients fall below all NDBC-based ones. Within the NDBC network, however, there is considerable scatter, but the figure suggests no convincing

dependence on either the mean wave height at a given buoy or its standard deviation. Nor have we discovered any dependence on average wave period (reported by the buoys) or on platform size (to the extent we have been able to find this information); see also the discussion by Durrant et al. (2009). Note that Gemmrich et al. (2011) found several discontinuities at decimeter size in northeast Pacific buoy time series, including two jumps exceeding 40 cm in the Canadian 46004 and 46184 buoys, but all the Canadian points in Figure 11 are nonetheless tightly clustered together.

These inconsistency problems among different *in situ* measurements are becoming widely appreciated, and a recent report has identified the critical need to assess and improve the quality of observations from the present networks of moored buoys (Swail et al., 2010). The recent work of Gemmrich et al. (2011) should further emphasize this need.

4.3 Buoy-altimeter differences in time

This section analyzes the buoy-altimeter wave-height discrepancies as a function of time in order to assess system stability and to detect drift errors or other temporally dependent errors. The approach is to apply the methodology devised by Mitchum (1998, 2000) to use tide-gauge data to monitor drift and other errors in altimetric sea-level measurements. The application to significant wave height is somewhat simpler, because unlike the sea-level case we can here use absolute heights as opposed to relative heights.

For details of the technique readers should consult Mitchum's original papers. The compilation of the altimeter-buoy collocation data differs somewhat from that used above. Briefly, at each buoy, ΔH_s time series are formed with altimeter measurements from nearby satellite tracks. Which point along any particular track is adopted for the comparisons is determined by minimizing the difference variance or maximizing altimeter-buoy correlations. Mitchum also allows the gauge (or buoy) data

to be shifted in time to allow for propagation effects, again by minimizing variances, and even though our software has this capability it seems less relevant to H_2 data than to sea level data so we did not allow any temporal shifting. Each buoy then yields several difference time series, one for each nearby track. All altimeter-buoy differences are then averaged for each satellite repeat cycle, yielding a final, single difference time series. The time series, so formed, are shown in Figure 12.

Several aspects of Figure 12 (top panel) merit discussion. None of the difference time series for the three satellites is centered about zero. Except for a brief two-year span of *TOPEX* all series imply that the mean altimeter wave heights are smaller than the mean buoy heights, in agreement with the previous section. The diagram for *TOPEX* clearly shows the period beginning in 1997 where H_s began drifting markedly, increasing by more than 30 cm in two years (e.g., Queffeulou, 2004). The sharp discontinuity in February 1999 was the result of switching to a second redundant altimeter aboard *TOPEX*. Neither *Jason-1* nor *Jason-2* shows signs of such anomalous drifting.

All of the series show evidence of an annual cycle. This is to be expected when, for example, the buoy and altimeter wave data differ by a scaling parameter, as our previous discussion suggested. Because the buoy data are limited to the northern hemisphere, an annual cycle should be apparent in the mean, so any scaling error with altimetry would generate a similar signal in the mean differences.

Figure 12 (bottom) shows the same data after adjusting the altimeter data by the β_1, β_0 linear corrections of Table 2. (A specially developed adjustment was obviously used for the anomalous 1997–1999 *TOPEX* data. The adjustment was taken as a simple linear function of time, which appears adequate to the task.) All three series are now seen to be more closely centered about zero with reduced evidence of annual oscillations. There is, however, a suggestion of a small change over time in the *Jason-1* data. In fact, if some of the fine detail of Figure 12 is reliable, *Jason-*

1 actually experienced a small jump around the beginning of 2005. It is not clear at this stage whether such fine detail is reliable—there are other small jumps at other times, and some of these must undoubtedly arise from changes (dropouts or additions) in the buoy network. In fact, five of our adopted buoys first came online in 2005, so to test sensitivity to this we removed these five buoys and regenerated Figure 12. The small jump in early 2005 remained. Of the various subtle offsets in Figure 12, the *Jason-1* offset in 2005 does appear the most exceptional.

Fitting a straight line to all adjusted time series in Figure 12 gives the following drift estimates:

TOPEX:	0.16 ± 0.10 cm yr^{-1}
Jason-1:	0.52 ± 0.17
Jason-2:	0.39 ± 0.62
Combined	0.06 ± 0.04

The drift in *Jason-2* is not statistically significant. (If the *TOPEX* time series is split into its separate Sides A and B—pre 1997 and post 1999—the trends are 0.90 ± 0.31 and -0.66 ± 0.50 cm yr^{-1}, respectively. The Side-A drift is clearly evident in Figure 12; the Side-B drift is hardly significant.) The drift in the combined 20-year time series is very small, less than 1 mm per year.

Notwithstanding these small residual drifts or offsets, and the many possible errors in the buoy data against which the calibration is being done, it seems safe to conclude that *Jason-2* data are of the same high quality as *Jason-1*, and that no large anomalous drifts of the sort that *TOPEX* experienced in 1997 has been affecting either the *Jason-1* or *Jason-2* data.

Application of Mitchum's technique to the *TOPEX* and *Jason-1* data, after the satellites had each been moved into their final interleaved orbits, results in the difference time series shown in Figure 13. After adjustment (bottom panel) the data appear comparable in quality to the results of Figure 12.

5 Discussion

Are the various calibration analyses presented above consistent? Surely the most robust results presented here are from the two verification campaigns, since these are based on more than 8 million measurements each and must be more reliable than either cross-over comparisons with other satellites or buoy comparisons. These analyses suggest that *TOPEX* significant wave heights are 1.7% larger on average than *Jason-1* wave heights, whereas *Jason-1* and *Jason-2* wave heights are on average essentially equivalent. The only qualification to these statements is that they strictly describe only the two six-month verification periods.

Comparisons against the NDBC buoy data (Table 2) suggest that, within quoted error bars, all three satellites report wave heights roughly comparable: between 5 and 6% smaller than the buoys. The β_1 scaling factors specifically suggest *TOPEX* and *Jason-1* heights should be in the ratio 1.063/1.056, or only 0.6% difference. However, these buoy results apply to the entire missions. If the small jump seen in Figure 12 at the start of 2005 is real, then we cannot expect exact agreement with the verification-campaign ratio. If the *Jason-1* data are separated into two periods before and after December 2004, orthogonal regression against the buoy data gives a slope $\beta_1 = 1.073$ for the early data, and 1.059 for the later data. Then the proper ratio with *TOPEX* is $1.073/1.056 = 1.016$, in near-perfect agreement with the direct ratio of 1.017.

Similarly, the regression slope for the late data (1.059) is very close, within 1%, for the *Jason-2* value of 1.053, and certainly within their uncertainties (Table 2).

Thus, these several independent strands of the calibration problem—buoy comparisons, 6-month satellite overlaps, and the Figure 12 time series—lead to a fairly consistent picture of the satellite-derived wave data.

There are, however, several topics left hanging. One of the most important is whether our

results can lend support to climate studies such as the recent work of Young et al. (2011) who used altimeter data to determine trends in global extreme wave heights (specifically in the 90th and 99th percentile waves). To some extent our results do support such studies, and diagrams such as Figure 10 confirm that the altimetry can properly map the 99th percentile at the available buoys. However, the 99th percentile in this collocation dataset, consisting of wave samples only at the infrequent times of satellite overpasses and limited severely in geographic extent, is unlikely to be representative of true 99th percentile global waves. Moreover, our drift results (Figure 12) apply to a great extent to the observed mean wave heights. It is possible that the tails of the wave-height distributions behave differently. We have attempted to address this issue by partitioning the collocation data of Section 4.1 into individual years, computing annual 99th percentile wave heights for both altimeter and buoy data, and plotting their differences in time. The result is Figure 14. (The advantage of using yearly intervals for this is that is avoids complications from the seasonal cycle.) Unfortunately, the error bars in Figure 14 are too large to constrain trends with any useful precision. More work is needed to confirm altimeter measurements of the largest waves.

A second topic of concern more directly addresses the main topic at hand: what are the possible effects on our analysis of changes or malfunctions in the buoy network, in the buoy instrumentation, and/or in the associated data processing algorithms? Errors—primarily jumps in H_s—of the sort recently documented by Gemmrich et al. (2011), must at some level lead to corresponding errors in our calibration analyses. Some presumably account for the scatter in individual buoy β_1 estimates (Figure 11). Also, some of the small jumps evident in Figure 12 may originate with the buoy data. Gemmrich et al. suggest a possible approach to addressing and correcting some of these problems, but we are hardly here in a position to do so, since a complete history of the network seems a required starting point. Accepting such errors as they now exist surely does not invalidate our

results. Even the current buoy data are sufficiently robust to maintain the noise envelope in Figure 12 to generally within ±10 cm and to detect easily any anomalies such as the 1997–1998 *TOPEX* drift.

Finally, it is clear that our calibration results are dependent on adoption of the NDBC network as the standard. Calibration against, for example, the Canadian DFO network, yields very different results. According to Challenor and Cotton (2003), other national (or international) networks would also yield disparate calibrations. Again we are in no position to solve such issues, except to lend our support to the views of Swail et al. (2010) that better consistency and accuracy of *in situ* ocean wave measurements is essential, not only for these sorts of calibration exercises, but for climate science in general.

Acknowledgments. This work was funded by the U. S. National Aeronautics and Space Administration under the Ocean Surface Topography program and the MEaSUREs program. Buoy data were obtained from the National Data Buoy Center, a part of the National Oceanic and Atmospheric Administration, and from Fisheries and Oceans Canada.

References

Abdalla, S., P. A. E. M. Janssen, and J.-R. Bidot. 2010. Jason-2 OGDR wind and wave products: monitoring, validation and assimilation. *Marine Geodesy,* 33: 239–255.

Abdalla, S., P. A. E. M. Janssen, and J.-R. Bidot. 2011. Altimeter near real time wind and wave products: random error estimation. *Marine Geodesy,* 34: 393–406.

Ablain, M., S. Philipps, N. Picot, and E. Bronner. 2010. Jason-2 global statistical assessment and cross-calibration with Jason-1. *Marine Geodesy,* 33: 162–185.

AVISO and PODAAC User Handbook. 2008. IGDR and GDR Jason products. AVISO SMM-MU-M5-OP-13184-CN and JPL D-21352, edition 4.1, 130 pp.

Caires, A. and A. Sterl. 2003. Validation of ocean wind and wave data using triple collocation. *Journal of Geophysical Research*, 108(C3): 3098.

Cheng, C.-L., and J. W. Van Ness. 1999. *Statistical Regression with Measurement Error.* Arnold, London.

Challenor, P. G., and P. D. Cotton. 2003. The joint calibration of altimeter and in situ wave heights. *Advances in the Application of Marine Climatology,* WMO Tech. Doc. 1081, JCOMM Tech. Rep. 13, pp. 139–148.

Dobson, E., F. Monaldo, J. Goldhirsh, and J. Wilkerson. 1987. Validation of Geosat altimeter-derived wind speeds and significant wave heights using buoy data. *Journal of Geophysical Research,* 92: 10719–10731.

Durrant, T. H., D. J. M. Greenslade, and I. Simmonds. 2009. Validation of *Jason-1* and *Envisat* remotely sensed wave heights. *Journal of Atmospheric and Oceanic Technology,* 26: 123–134.

Everitt, B. S. 2002. *The Cambridge Dictionary of Statistics*, 2nd ed. Cambridge Univ. Press, Cambridge.

Gemmrich, J., B. Thomas, and R. Bouchard. 2011. Observational changes and trends in northeast Pacific wave records. *Geophysical Research Letters,* 33: L22601.

Hamilton, G. D. 1986. National Data Buoy Center programs. *Bulletin of the American Meteorological Society,* 73: 985–993.

Isobe, T., E. D. Feigelson, M. G. Akritas, and G. J. Babu. 1990. Linear regression in astronomy. *Astrophysical Journal,* 364: 104–113.

Lambin, J., R. Morrow, L.-L. Fu, J. K. Willis, H. Bonekamp, J. Lillibridge, J. Perbos, G. Zaouche, P. Vaze, W. Bannoura, F. Parisot, E. Thouvenot, S. Coutin-Faye, E. Lindstrom, and M. Mignogno. 2010. The OSTM/Jason-2 mission. *Marine Geodesy,* 33: 4–25.

Lefèvre, J.-M., and P. D. Cotton. 2001. Ocean surface waves. In *Satellite Altimetry and Earth Sciences,* (Eds.: L.-L. Fu and A. Cazenave), Chapter 7, pp. 305–328. Academic Press, San Diego.

Lillibridge, J., R. Scharroo, G. Jacobs, L. Russell, and V. Tabor. 2011. Quality assessment of the Jason-2 operational and interim geophysical data records. *Marine Geodesy,* 34: 191–213.

Madansky, A. 1959. The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Association,* 54: 173–205.

Mitchum, G. T. 1998. Monitoring the stability of satellite altimeters with tide gauges. *Journal of Atmospheric and Oceanic Technology,* 15: 721–730.

Mitchum, G. T. 2000. A improved calibration of satellite altimetric heights using tide gauge sea levels with adjustment for land motion. *Marine Geodesy,* 23: 145–166.

Mitchum, G. T., D. W. Hancock, G. S. Hayne, and D. C. Vandemark. 2004. Blooms of σ^0 in the TOPEX radar altimeter data. *Journal of Atmospheric and Oceanic Technology,* 21: 1232–1245.

Monaldo, F. 1988. Expected differences between buoy and radar altimeter estimates of wind speed and significant wave height and their implications on buoy-altimeter comparisons. *Journal of Geophysical Research,* 93: 2285–2302.

OSTM/Jason-2 Products Handbook. 2009. CNES: SALP-MU-M-OP-15815-CN, EUMETSAT: EUM/OPS-JAS/MAN/08/0041, NASA/JPL: OSTM-29-1237, NOAA/NESDIS: Polar Series/OSTM J400, Issue 1 Rev. 4, 67 pp.

Quartly, G. D. 2000. The gate dependence of geophysical retrievals from the TOPEX altimeter, *Journal of*

Atmospheric and Oceanic Technology, 17: 1247–1251.

Quartly, G. D. 2009. Optimizing σ^0 information from the Jason-2 altimeter, *IEEE Geoscience and Remote Sensing Letters,* 6: 398–402.

Quartly, G. D. 2010. Jason-1/Jason-2 metocean comparisons and monitoring. *Marine Geodesy,* 33: 256–271.

Queffeulou, P. 2004. Long-term validation of wave height measurements from altimeters. *Marine Geodesy,* 27: 495–510.

Ray, R. D., and B. D. Beckley. 2003. Simultaneous ocean wave measurements by the Jason and Topex satellites, with buoy and model comparisons. *Marine Geodesy,* 26: 367–382.

Scharroo, R., J. Lillibridge, and E. Leuliette. 2009. Good, better, best: A comparison of Jason-2 O/I/GDR products. OSTST meeting, Seattle.

Swail, V., et al. 2010. Wave measurements, need and development for the next decade. *Proceedings OceanObs'09*, Vol. 2, Venice, September 2009, (Eds.: J. Hall, D. E. Harrison, and D. Stammer), ESA Publ. WPP-306, doi:10.5270/OceanObs09.cwp.87

Teng, C.-C. 2002. Wave measurements from NDBC buoys and C-MAN stations. *Proc. OCEANS '02 MTS/IEEE*, Biloxi, pp. 517–524, doi:10.1109/OCEANS.2002.1193322

Thibaut, P., L. Amarouche, O. Z. Zanife, N. Steunou, P. Vincent, and P. Raizonville. 2004. Jason-1 altimeter ground processing look-up correction tables. *Marine Geodesy,* 27: 409–431.

Thibaut, P., J. C. Poisson, E. Bronner, and N. Picot. 2010. Relative performance of the MLE3 and MLE4 retracking algorithms on Jason-2 altimeter waveforms. *Marine Geodesy,* 33: 317–335.

Wilk, M. B., and R. Gnanadesikan. 1968. Probability plotting methods for the analysis of data. *Biometrika,* 55: 1–17.

Young, I. R., S. Zieger, and A. V. Babanin. 2011. Global trends in wind speed and wave height. *Science,* 332: 451–455.

Zieger, S., J. Vinoth, and I. R. Young. 2009. Joint calibration of multiplatform altimeter measurements of

wind speed and wave height over the past 20 years. *Journal of Atmospheric and Oceanic Technology,* 26: 2549–2563.

Table 1: Global H_s statistics during verification periods

	Mean (cm)	Median (cm)	Mode (cm)	Std. dev. (cm)	Skewness	Kurtosis
Jason-1	265	234	182	133	1.3	2.14
Jason-2	266	236	184	133	1.3	2.15
J1 minus J2	−1	−1	−1	12	0.7	> 200
Topex	273	237	190	135	1.2	2.03
Jason-1	271	243	190	132	1.2	1.87
Topex minus J1	2	2	2	11	3.4	> 200

From over 8 million 1-Hz measurements along satellite ground-track during two verification campaigns, after along-track median filtering.

Table 2: Altimeter–buoy H_s differences

	N	Mean (cm)	RMS (cm)	Linear relationship β_0	Linear relationship β_1	ρ
Topex	5652	−3.8	24	-7.8 ± 0.8	1.056 ± 0.005	0.983
Jason-1	7462	−5.4	24	-7.7 ± 0.8	1.063 ± 0.005	0.982
Jason-2	3452	−2.2	23	-8.7 ± 1.1	1.053 ± 0.007	0.983
Jason-1 *	2543	13.7	29	-3.2 ± 1.6	0.960 ± 0.008	0.975

* Final line employs Canadian buoys; all others are NDBC.

The mean refers to (altimeter minus buoy). The linear structural relationship corresponds to Buoy = $\beta_1 \times$ Altimeter + β_0 (in cm), assuming equal error variances. N is number of collocated altimeter-buoy measurement pairs. Uncertainties representing one standard error follow from large-sample formulae given by Isobe et al. (1990).

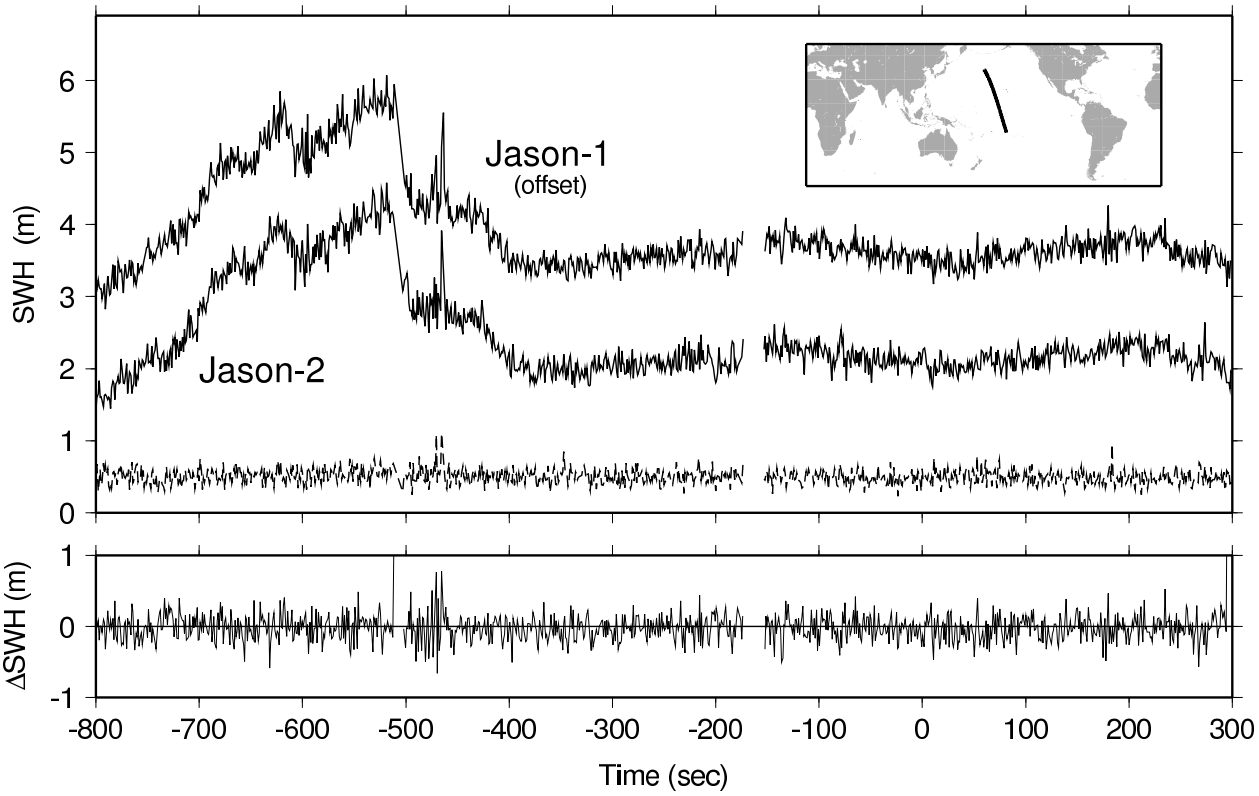

Figure 1: Significant wave height measurements by *Jason-1* and *Jason-2* Ku-band satellite altimeters (with *Jason-1* data offset by 1.5 m), for a partial track across the Pacific Ocean (see map inset), collected during a time when both altimeters were flying in tandem formation separated by only 55 s. Time on horizontal axis is relative to each satellite's equatorial crossing time, which occurred at 27 October 2008 09:21:25 and 09:22:21 UT, respectively. Data are from pass 210, *Jason-1* cycle 250 and *Jason-2* cycle 11. Lighter dashed line shows the standard deviation of the 20-Hz wave heights for *Jason-2*, which is nearly a constant 50 cm. Bottom panel shows the H_s differences between the two arcs, with a mean of -2 cm and RMS of 18 cm.

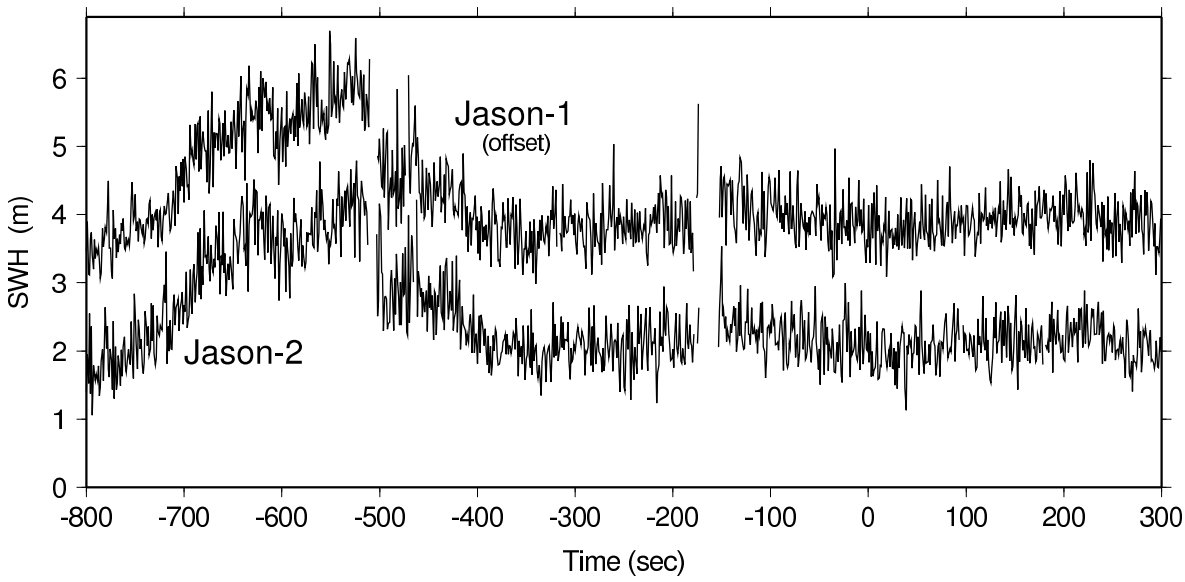

Figure 2: As in Figure 1, but for the C-band wave heights. The noise level is considerably greater. The RMS difference between the two arcs is 44 cm.

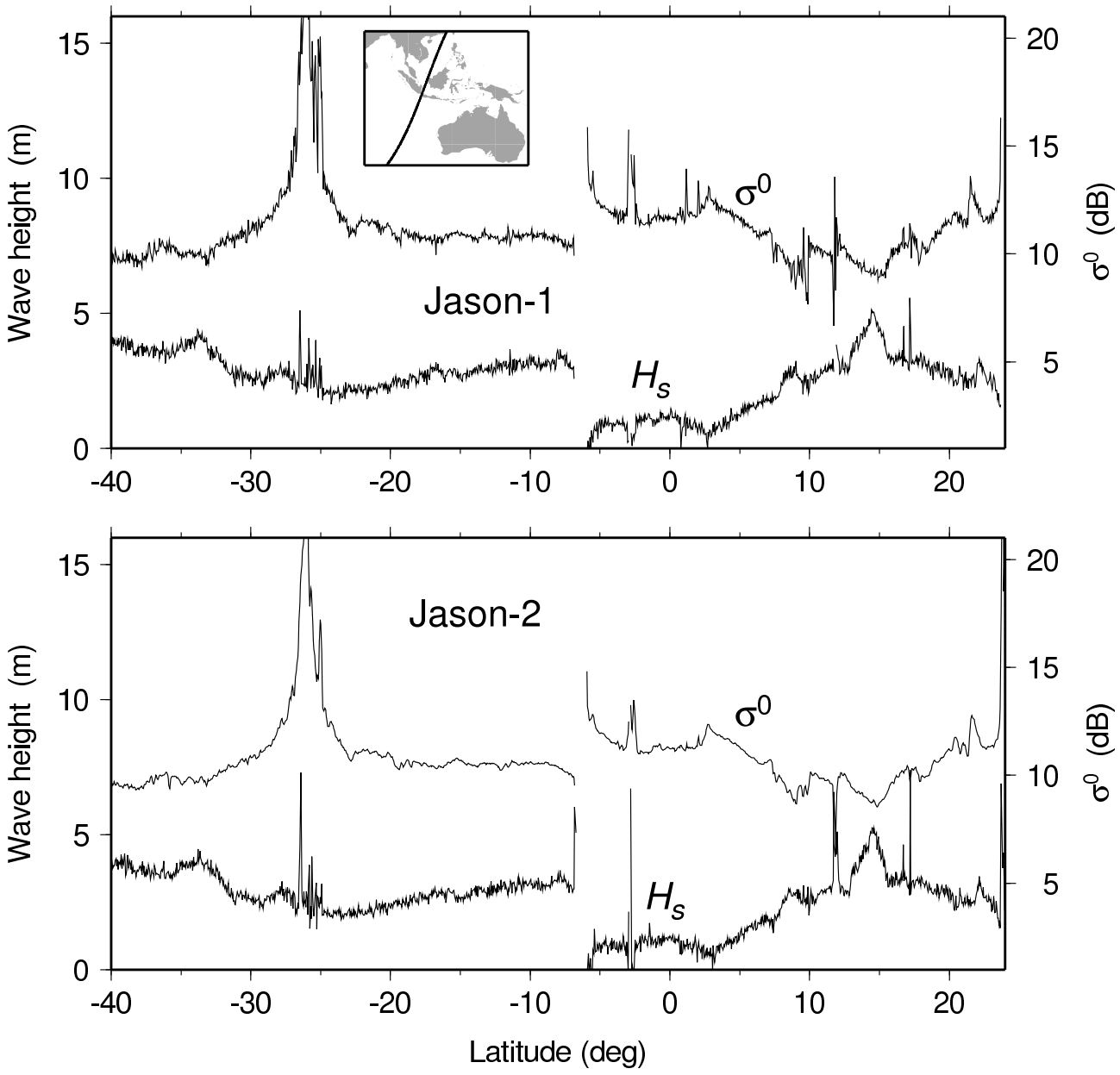

Figure 3: Comparison of near-simultaneous measurements by the *Jason-1* and *Jason-2* altimeters of significant wave height H_s and normalized backscatter coefficient σ^0, along a track crossing the eastern Indian Ocean and the South China Sea (see map inset). Data are from pass 229, *Jason-2* cycle 8 and *Jason-1* cycle 247, collected on 28 September 2008.

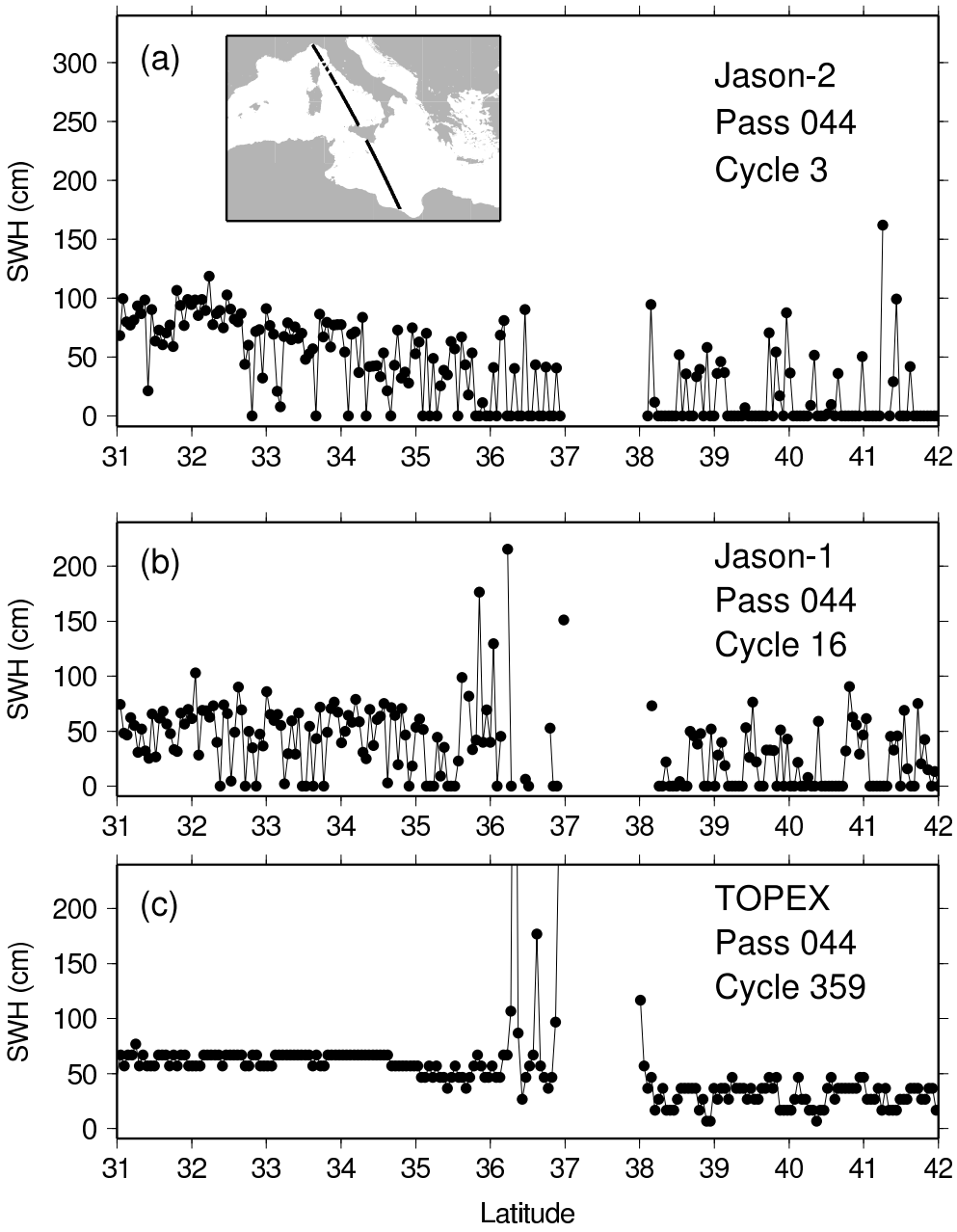

Figure 4: (a) *Jason-2* H_s data crossing the Mediterranean Sea and displaying anomalous wave heights of precisely zero. Such problems appear to occur when the backscatter coefficient σ^0 is moderate to large, presumably over very calm seas. (b) and (c) show similar (simultaneous) passes for *Jason-1* and *TOPEX*. It appears *TOPEX* wave data are less sensitive to high backscatter. The coarser digitization level (10 cm) used for *TOPEX* is evident.

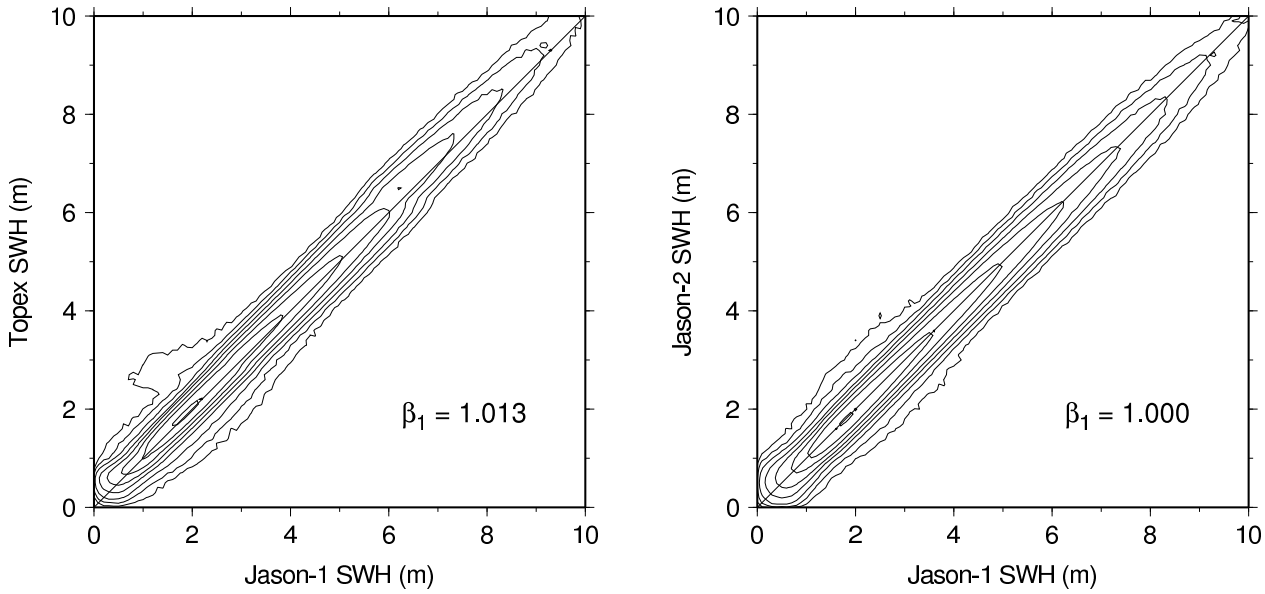

Figure 5: Observed joint density function of H_s measurements during (left) the *Jason-1* verification campaign of 2002 and (right) the *Jason-2* verification campaign of 2008. Left diagram is based on *Jason-1* repeat cycles 3–18 and *TOPEX* cycles 346–363 (excluding 361); right diagram is based on *Jason-2* cycles 1–20 and *Jason-1* cycles 240–259. Contours are roughly logarithmic: 0.0003, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1.0 m^{-2}. These data have *not* been subjected to median filtering, which tends to tighten the contours considerably. β_1 is estimated slope from an orthogonal regression. There is a small tendency for *TOPEX* waves to be slightly larger than *Jason-1* waves, while *Jason-1* and *Jason-2* are very consistent.

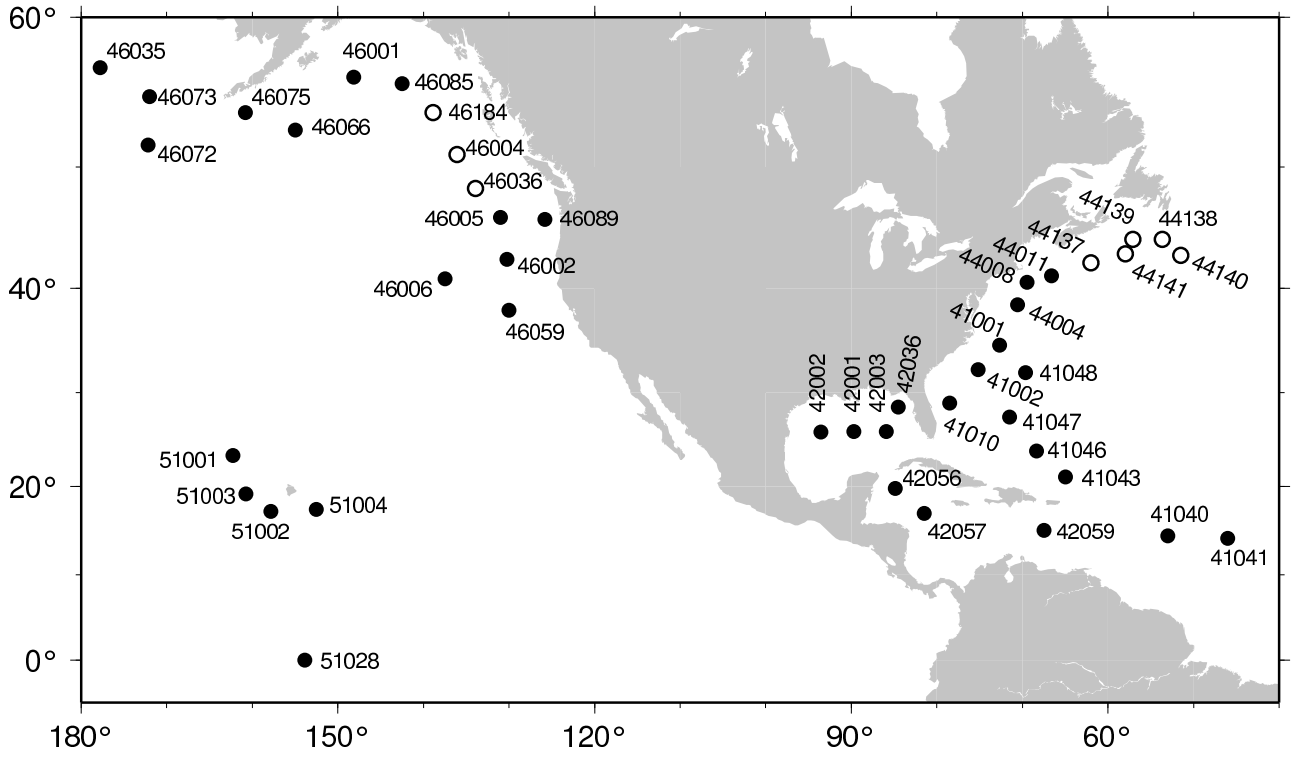

Figure 6: Locations of buoy wave measurements. Canadian buoys are denoted by open circles; U. S. buoys by filled circles. Buoy labels follow the World Meteorological Organization's standard name conventions. All buoys are in deep (> 700 m) water, save 42036 which sits on the Florida Shelf in about 50 m of water. Some of the buoys (e.g., 41040–41048 and 42056–42059) were established only recently, since 2005 or later.

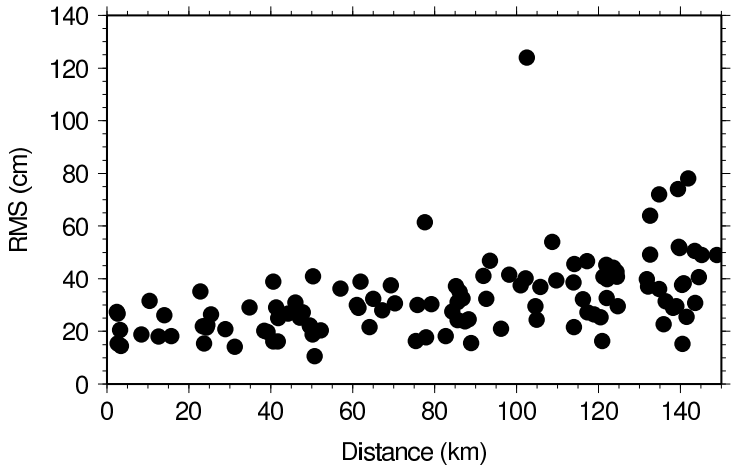

Figure 7: RMS wave-height differences (cm) between pairs of buoys and their closest points along particular altimeter tracks, based on 2002–2008 *Jason-1* data. Horizontal axis denotes the distance between the buoy and the altimeter track. Each plotted point represents the RMS difference between from 40 to 240 pairs of H_s measurements. As should be expected, a general inflation in RMS occurs as the distance between buoy and track grows.

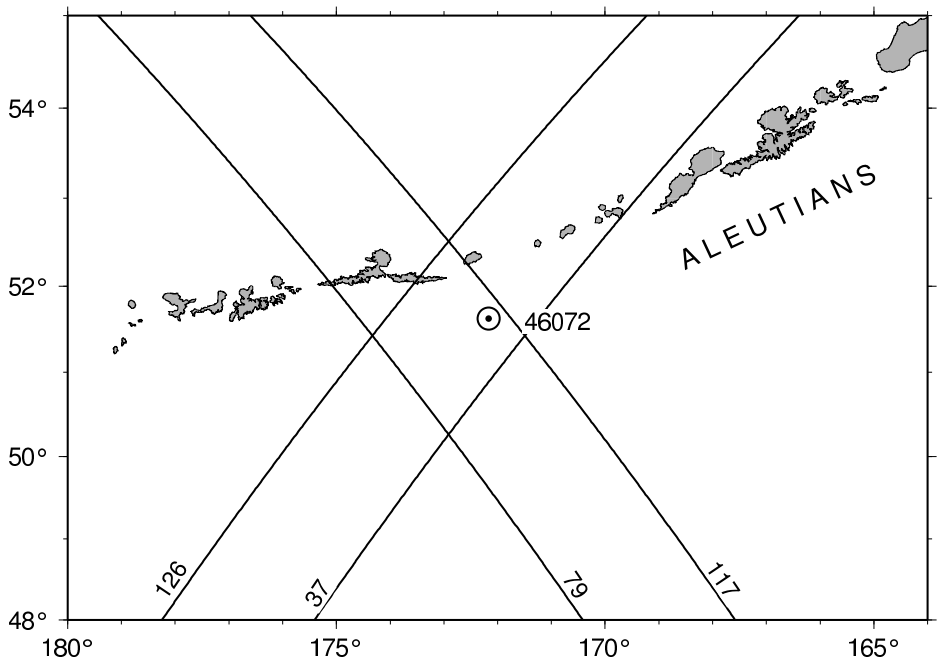

Figure 8: Location of NDBC buoy 46072 in the North Pacific and its four closest altimeter tracks. The largest outlier in Figure 7 corresponds to the RMS difference in H_s between this buoy and track 126, for which the closest point is on the opposite side of the Aleutians.

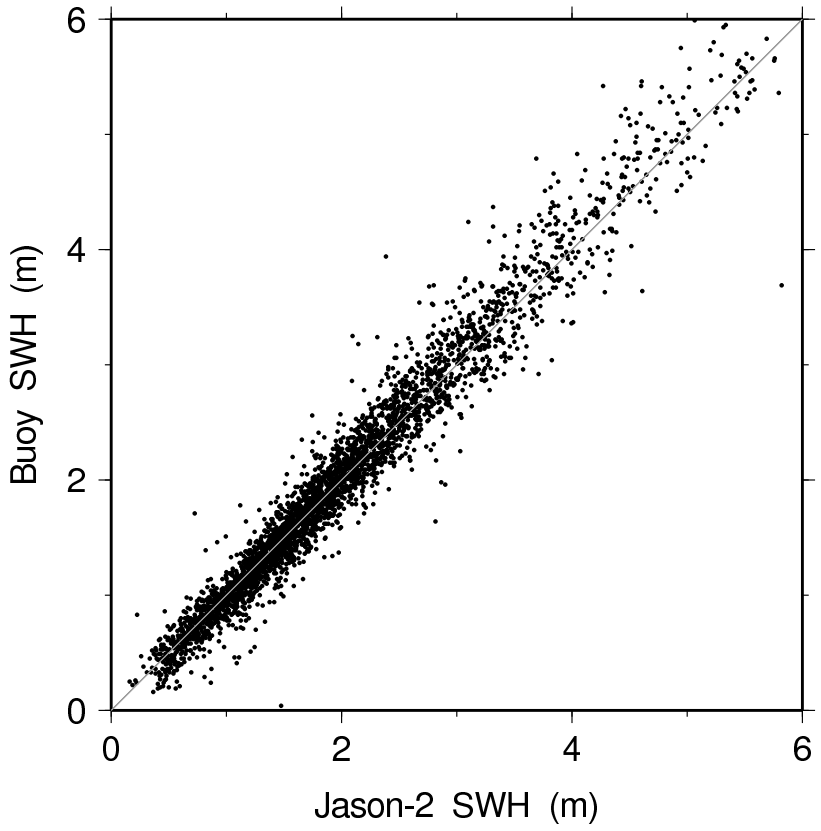

Figure 9: Scatter plot of 3452 pairs of collocated buoy and altimeter H_s measurements for the *Jason-2* satellite, from July 2008 through June 2011. All buoy data are from the NDBC archives.

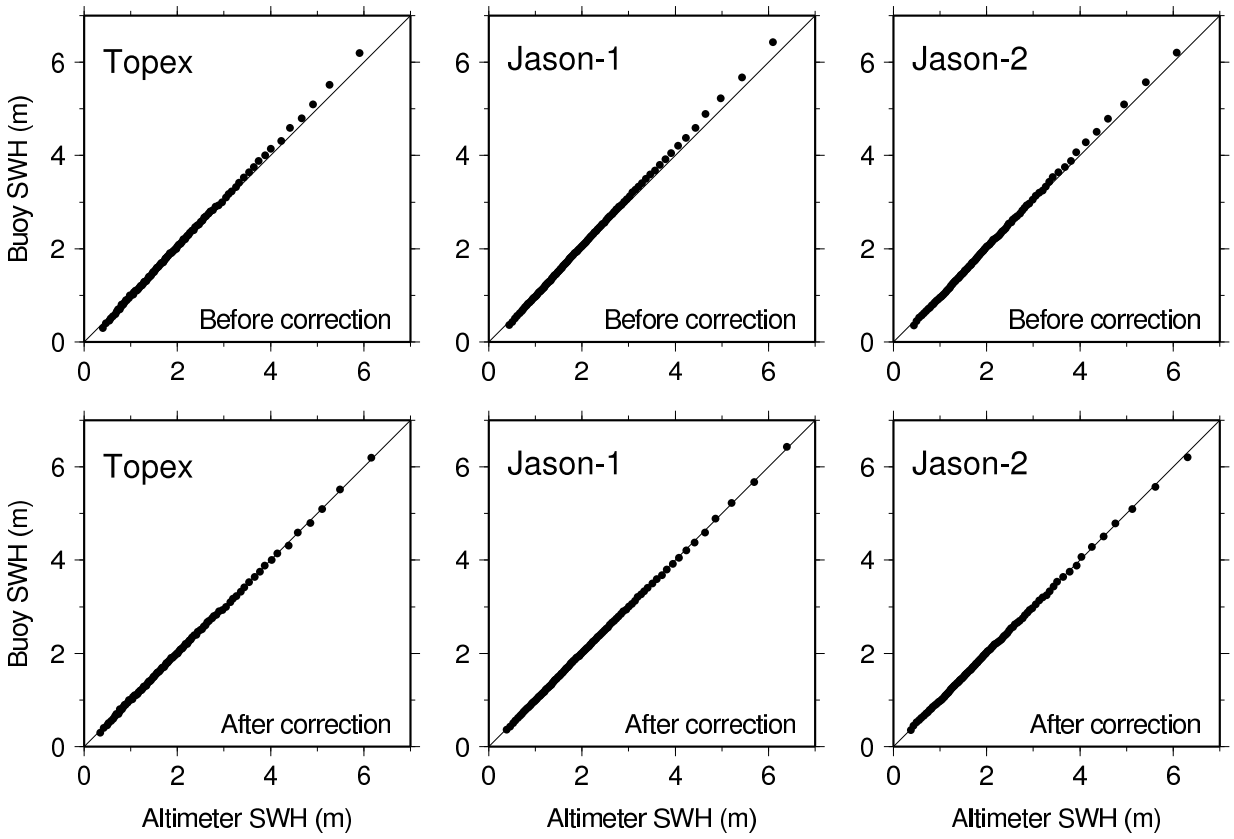

Figure 10: Quantile (or Q-Q) plots showing 99 H_s joint percentiles, from the first to 99th, for collocated buoy and altimeter observations. Top panels show original data, bottom show after adjustment using the NDBC-based linear coefficients of Table 2.

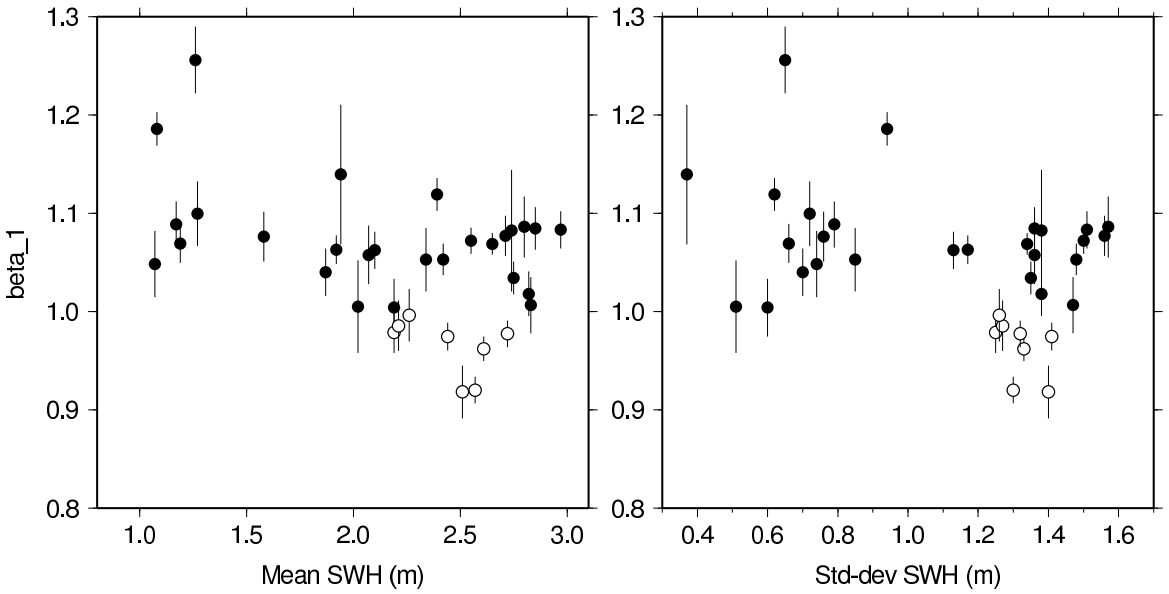

Figure 11: Estimates of the regression slope coefficient β_1 derived at individual buoys by comparing against *Jason-1* altimetric H_s measurements. Filled circles are NDBC buoys; open circles are Canadian DFO buoys. No obvious dependence of β_1 is observed as a function of either mean wave height (left panel) or wave standard deviation (right panel). The slope coefficient for every DFO buoy is less than the coefficient for every NDBC buoy.

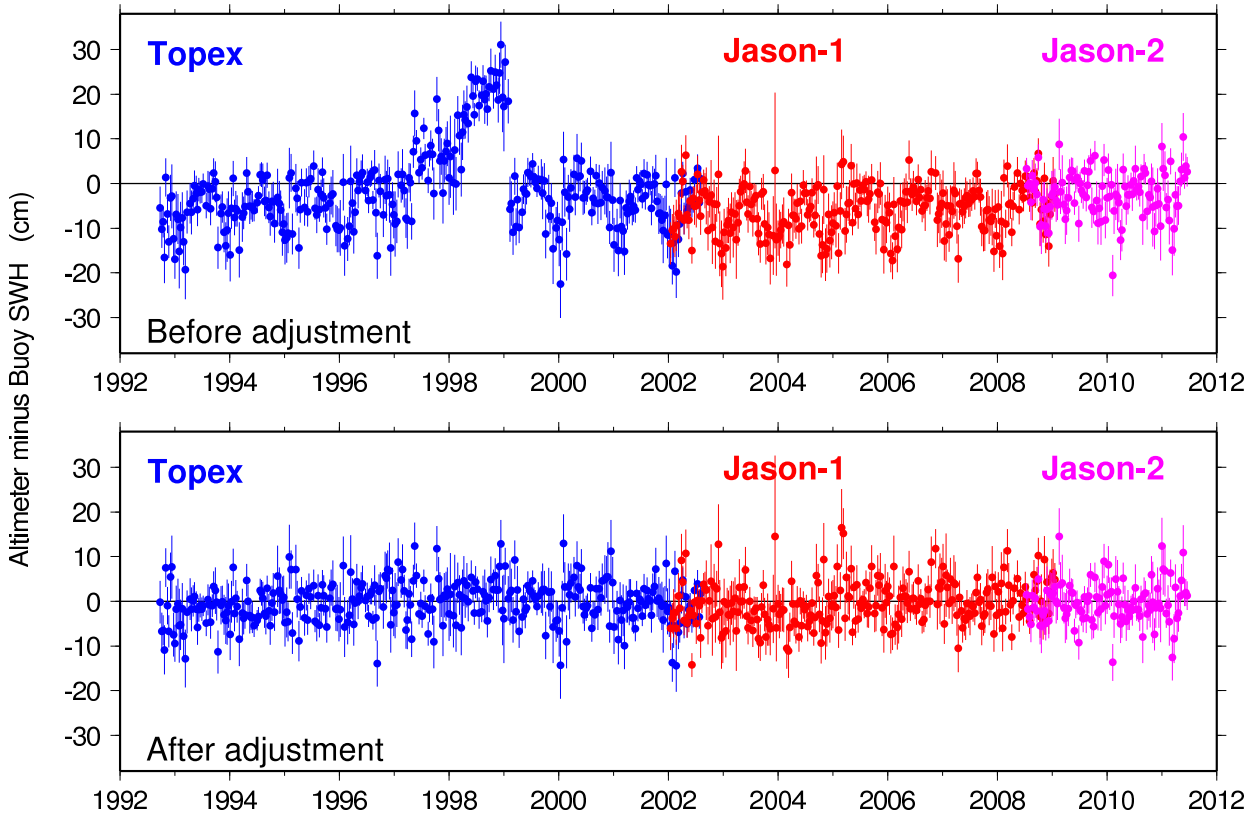

Figure 12: Mean altimeter minus buoy significant wave height differences, for each repeat cycle (i.e. every 9.9 days) for the *TOPEX* (blue), *Jason-1* (red) and *Jason-2*(magenta) satellites. Top panel based on original GDR data (see text for which GDR versions); bottom panel shows mean differences after each altimeter height has been adjusted according to the linear coefficients of Table 2 (the anomalous *TOPEX* data of 1997–1999 were handled specially). Adjustment noticeably reduces the bias and annual oscillations seen in top panel.

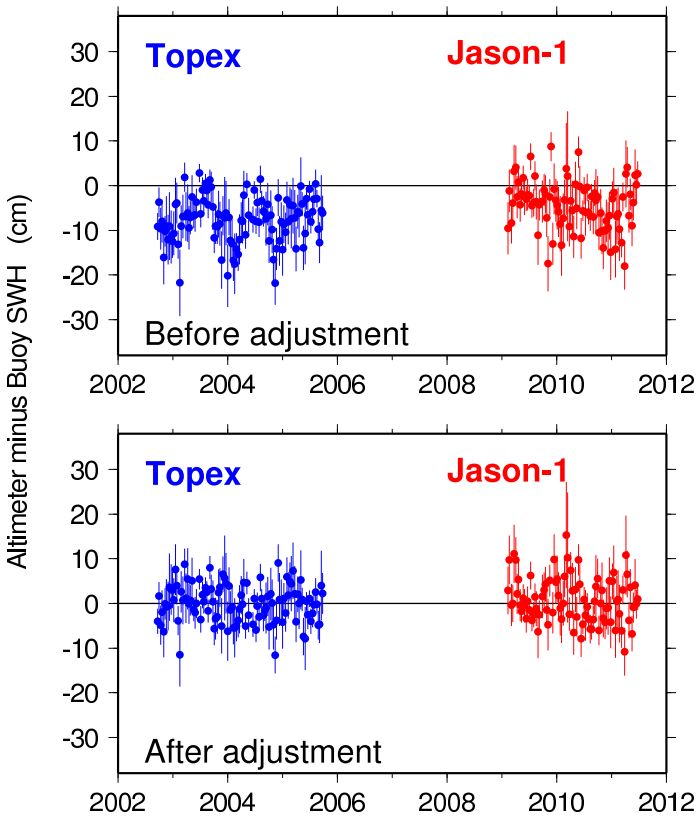

Figure 13: As in Figure 12, but after each satellite had been moved into its final interleaved orbit.

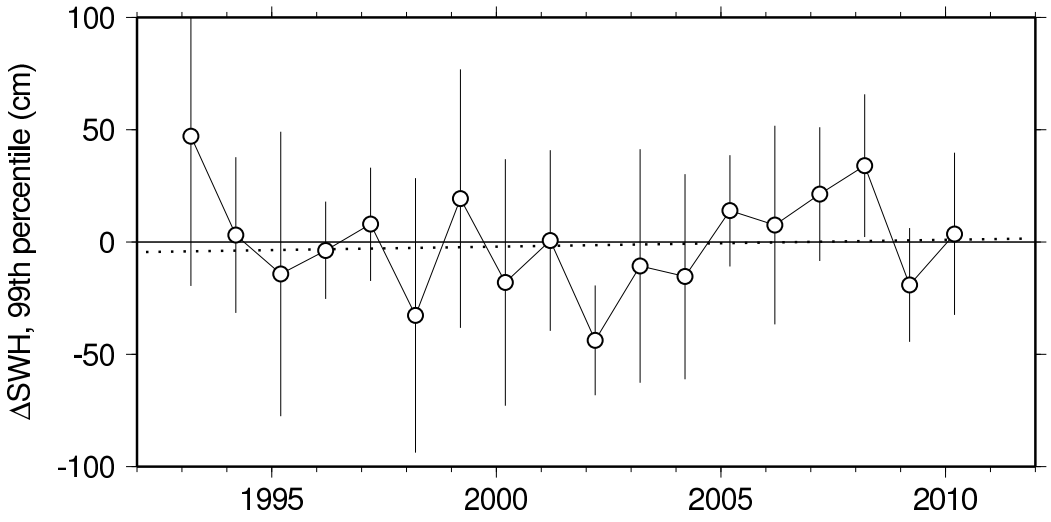

Figure 14: Annual differences (in cm) between the 99th percentile altimeter and buoy H_s data, based on the collocation datasets used in Section 4.1. Error bars are computed by a bootstrap resampling algorithm from each year's data. The error bars are too large to yield useful constraints on possible drift in altimeter measurements of large waves.