

Statistical Comparison of Global Significant Wave Heights from Topex and ERS-1 Altimeter and from Operational Wave Model WAM

C. Staabs¹ and E. Bauer²

¹Institut für Meereskunde, Universität Hamburg, Germany

²Potsdam-Institut für Klimafolgenforschung, Potsdam, Germany

Received 25 April 1997; accepted 14 August 1997

Abstract. Global data sets of significant wave height (H_s) from radar altimeters aboard the satellites ERS-1 and Topex and from the wave model WAM are compared for 1993. The bivariate distributions from the collocation of the data are displayed in scatter diagrams. The statistical properties of bivariate distributions, where both data sets contain random errors, are best derived from principal component analysis (PCA). From this, a good agreement between H_s of the ERS-1 altimeter and the WAM model is inferred. The assimilation of H_s from ERS-1 altimeter into WAM is seen to have two effects. Firstly, the rms deviation between these data sets is reduced significantly. Hence the bivariate distribution narrows considerably. Secondly, as the high sea states from ERS-1 altimeter are seen to be low-biased compared to WAM before assimilation, the high sea states of WAM are reduced. This results in a rotation of the principal axes toward the right direction. The positive impact from the assimilation is proven by the significant reduction of the rms deviation between H_s from WAM and the independent H_s from Topex. The rotation of the distribution, however, appears to overshoot the mark. © 1998 Elsevier Science Ltd. All rights reserved

1 Introduction

The measurement of significant wave heights (H_s) with buoys is limited to a few locations. This in-situ reference suffers from the relatively small number of data sets and the incomplete coverage of the natural range of variations of H_s . Radar altimeter from space-borne platforms give researchers an excellent tool to create global H_s data sets of high precision every few days, because H_s can be inferred directly from the shape of the returned radar pulse without any further a priori information [Barrick (1968)]. For instance, Topex pro-

vides a global set of H_s measurements along its ground track every ten days. Other sources for global H_s are wave models such as the spectral third-generation Wave Model (WAM) [WAMDI Group (1988)] operated routinely at the European Centre for Medium-range Weather Forecasts (ECMWF). In the present study we compare global H_s obtained from the altimeters of the ERS-1 (H_s^E) and Topex satellites (H_s^T) with those obtained from the wave model WAM (H_s^W).

In various studies, data from altimeters were validated against in-situ observations from buoys [e.g. Queffelec and Lefevre (1992), Cotton and Carter (1994), and Gower (1996)]. Uncertainties still remain about which of the H_s data sets agrees more closely with the "true" sea state. On average, the accuracy of H_s from altimeters was repeatedly confirmed to be below the commonly used error boundaries of 0.5 m or 10% (whichever is larger) in the range 1 to 20 m.

Global model calculations of the sea state are limited to reproducing observations accurately on relatively coarse spatial grids and with parameterized physical processes. Errors in wave modelling using WAM are caused mainly by incorrect wind forcing and less by insufficient resolutions [Bauer et al. (1992), Cardone et al. (1995)]. Data assimilation techniques are used to compensate for many of these limitations, and since August 16, 1993, H_s from the ERS altimeters has been assimilated into the WAM model at ECMWF [Lionello et al. (1995)].

Section 2 describes the used data sets and their preparation. Remarks on the statistical method used are given in Sect. 3. The bivariate distributions of H_s from collocated altimeter and model data are characterized by principal component analysis in Sect. 4. Conclusions on the assessment of the quality of the analyzed data sets follow in Sect. 5.

Correspondence to: C. Staabs

2 Data sets of H_s

Measured H_s of 1993 from the altimeters aboard the satellites ERS-1 and Topex are collocated to H_s modeled by WAM. The altimeter data are preprocessed using a quality control and along track averaging [see Bauer et al. (1992)]. The altimeter data were averaged over 30-point intervals giving a spatial resolution of about 200 km which corresponds closely to the resolution of the model. Spikes which occurred mainly at the edges of data gaps were rejected. The spikes were defined as data points which deviate by more than twice the ensemble standard deviation from the ensemble average. After rejection of the spikes, the ensemble mean was recalculated over the residual data points and discarded if the remaining standard deviation exceeded 10% of the ensemble mean. In this way, the data set was smoothed. The collocation of the data set to WAM data uses linear interpolation of the model data in two spatial directions and one time direction. Collocated pairs were only formed if the surrounding WAM grid points lay in ice-free water. We analyze the Fast Delivery Product (FDP) of H_s^E instead of the off-line preprocessed Ocean Product (OPR) data, because the former are the data which are assimilated operationally into WAM at ECMWF.

Modeled H_s are provided by from WAM (cycle 4) with a global $3^\circ \times 3^\circ$ grid and with forcing by ECMWF 6-hourly wind fields. The modeled H_s fields are stored every 6 hours. The major improvement of WAM cycle 4 with respect to WAM cycle 3 is the dynamic coupling between the wave-induced stress and the atmospheric stress [Komen et al. (1994)]. The next step toward improving the operational wave prediction is the assimilation of FDP H_s^E into WAM. Every six hours the modeled wave data are corrected through an optimal interpolation scheme [Lionello et al. (1995)]. All H_s^E observations within a certain correlation length scale (~ 200 km) from a model grid point and within a 6-hour interval centered at the analysis time are incorporated to determine corrections of the windsea and swell partition at a model grid point.

Time periods of the analyzed data sets and their latitudinal boundaries, which depend on the different satellite orbits and on the model grid are shown in Table 1. Only data from regions without sea-ice are considered. Since August 1993, the WAM model has accounted for the actual ice boundaries inferred from sea surface temperature analysis provided by the National Center for Environmental Prediction (NCEP).

3 Statistical Methods

In this section we briefly consider the statistical method used in the present study. Detailed derivations of the statistical parameters are beyond the scope of this pa-

Table 1. Latitudinal coverage and time periods of used collocated H_s data sets

data set	latitudinal cov.	period
WAM - ERS-1	69° N - 60° S	Jan 1 to Jul 31, 1993
WAM - ERS-1	78° N - 78° S	Aug 1 to Dec 31, 1993
WAM - Topex	66° N - 60° S	Jan 1 to Jul 31, 1993
WAM - Topex	66° N - 66° S	Aug 1 to Dec 31, 1993

per. The reader is referred to statistical textbooks.

Collocated data sets of H_s are commonly described by linear regression analysis and displayed in a scatter diagram. But the use of the linear regression analysis is based on the assumption that one data sets is known precisely and the other contains random errors. This assumption does not hold for the H_s data sets analyzed here. Bivariate distributions, where both data sets are affected by random uncertainty, are more adequately described with the method of the principal component analysis (PCA) [Preisendorfer (1988)]. The eigenvectors of the data covariance matrix yield the major and minor axes of the elliptical distribution. The slope

$$b_p = \tan \Theta \quad (1)$$

of the major axis

$$\hat{y}_p = b_p x + a_p \quad (2)$$

and its intercept a_p on the y axis are derived from the first PCA angle

$$\Theta = \frac{1}{2} \arctan \frac{2s_{xy}}{s_{xx} - s_{yy}}, \quad (3)$$

where

$$s_{xy} = \langle (y - \langle y \rangle)(x - \langle x \rangle) \rangle = s_{yx} \quad (4)$$

is the covariance and

$$s_{xx} = \langle (x - \langle x \rangle)^2 \rangle \quad (5)$$

$$s_{yy} = \langle (y - \langle y \rangle)^2 \rangle \quad (6)$$

are the variances of the two collocated data sets ($\langle \rangle$ denotes the ensemble mean). The minor principal axis lies perpendicular to the major axis. The principal axes cross at the point $P_c = (x_c, y_c)$ with $x_c = \langle x \rangle$ and $y_c = \langle y \rangle$. The first and the second eigenvalues of the covariance matrix represent the maximum and minimum variance along the axes, respectively. The PCA rms deviations σ_{p1} and σ_{p2} , i.e., the square roots of the eigenvalues of the data covariance matrix, are

$$\left. \begin{array}{l} \sigma_{p1} \\ \sigma_{p2} \end{array} \right\} = \left(\frac{1}{2} \{ (s_{xx} + s_{yy}) \pm [(s_{xx} - s_{yy})^2 + 4s_{xy}^2]^{\frac{1}{2}} \} \right)^{\frac{1}{2}}, \quad (7)$$

where σ_{p1} (σ_{p2}) are computed with the + (-) sign. The parameter σ_{p1} measures the significant range of variation of two data sets and σ_{p2} measures the decorrelated noise. The lower σ_{p2} , and the closer Θ to 45° , the better the agreement between the collocated data sets.

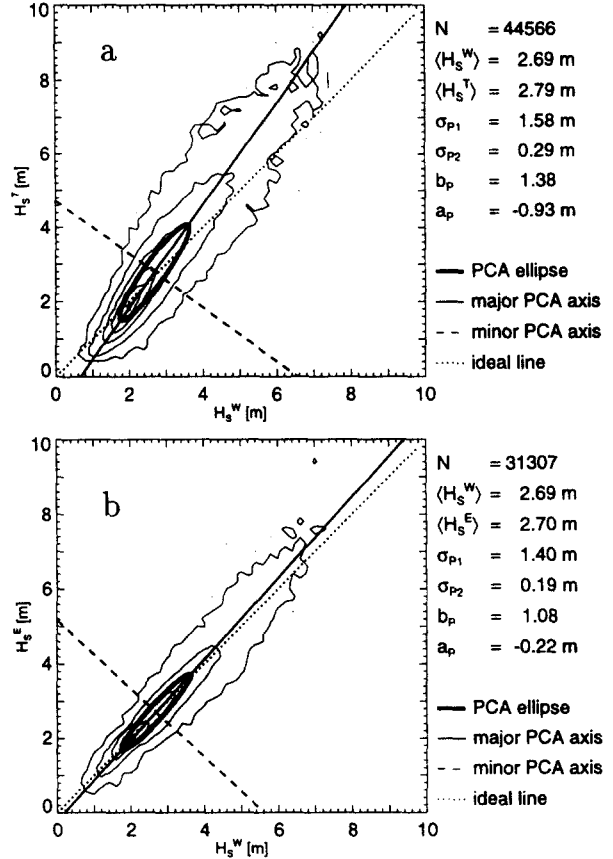
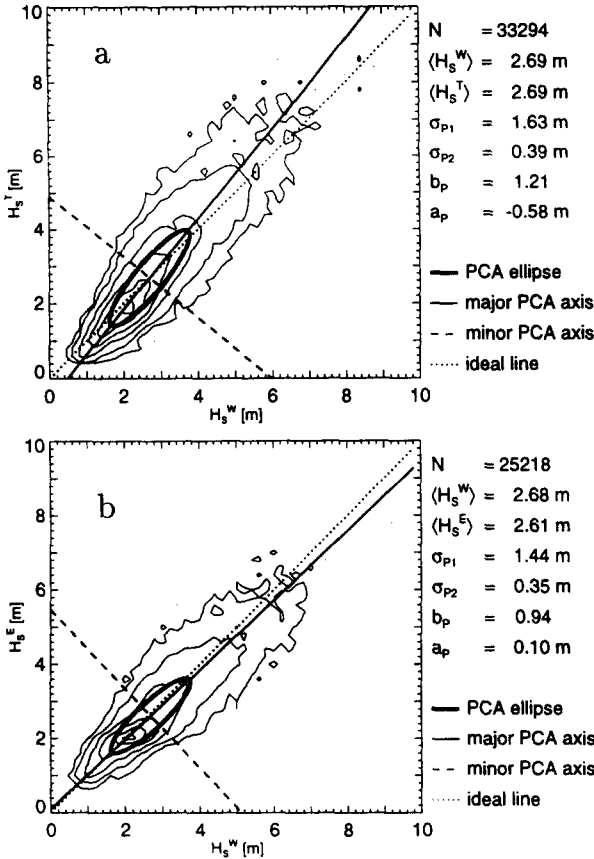


Fig. 1. Scatter diagrams of global H_s from WAM (H_s^W) collocated to a) Topex altimeter (H_s^T) and b) ERS-1 altimeter (H_s^E) for February 1993. Data are binned into 0.2×0.2 m boxes, and contour lines are for: 3 entries, and 2, 5, 10, 30, 60, and 90 % of the maximum number of entries. The major (solid) and minor (dashed) principal axes and their coefficients are shown together with the ideal line (dotted). The half-axes of the ellipse are equal to σ_{p1} and σ_{p2} .

Fig. 2. Same as Fig.2 but for December 1993 and with assimilation of H_s^E into WAM.

4 Statistics of bivariate distributions of H_s

To assess the impact of assimilation of H_s^E into WAM, scatter diagrams of global H_s^W collocated to H_s^T and with H_s^E are presented as an example for February 1993 (Fig.1) and for December 1993 (Fig.2). In February 1993 data assimilation was not implemented, but in December 1993 H_s^E was assimilated into WAM. The major (solid line) and minor (dashed line) PCA axes are shown with their coefficients. The dotted line is the ideal line with slope 1. Values of H_s below 0.5 m are discarded, because they are below the precision of the altimeter measurement.

Fig.1 a) displays the bivariate distribution for H_s^T and H_s^W . The slope of the major PCA axis is $b_p = 1.21$. The corresponding intercept with the y-axis is $a_p = -0.58$ m. This indicates that the major PCA axis of the distribution

g

is slightly rotated in an anti-clockwise direction with respect to the ideal diagonal around the point P_c of the bivariate distribution. Taking into account that $\langle H_s^W \rangle = \langle H_s^T \rangle = 2.7$ m, H_s^W is seen to be lower than H_s^T for H_s higher than 2.7 m and vice versa for H_s lower than 2.7 m.

Before assimilation, in February 1993, the distribution of H_s^E and H_s^W (Fig.1 b) has slope $b_p = 0.94$, which is closer to 1 than inferred from H_s^T and H_s^W . But b_p is lower than 1, which indicates that H_s^E is less than H_s^W for H_s greater than 2 m. The parameter $\sigma_{p2} = 0.35$ m for H_s^E and H_s^W is lower than $\sigma_{p2} = 0.39$ m for H_s^T and H_s^W , indicating a better agreement between H_s^E and H_s^W . It is also remarkable that $\sigma_{p1} = 1.44$ m for the H_s^E and H_s^W data is lower than $\sigma_{p1} = 1.63$ m for the H_s^T and H_s^W data. This indicates that H_s^E represents a smaller portion of the natural range of H_s than H_s^T . Although H_s^E is measured over a higher latitude range than H_s^T , this obviously does not lead to a greater portion of high H_s values in the data set of H_s^E .

The impact of H_s^E assimilation into WAM is seen to

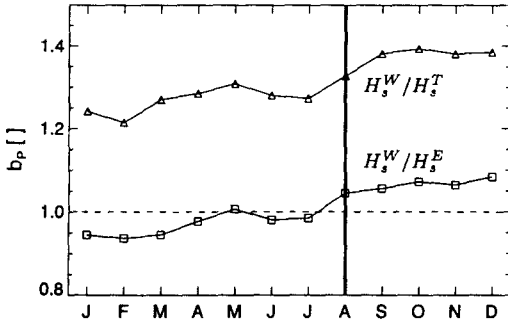


Fig. 3. Time series of monthly averaged slopes b_p of the major PCA axes of global H_s from WAM model collocated to Topex altimeter (triangles) and to ERS-1 altimeter (squares) for 1993. The vertical line indicates the start of wave data assimilation into WAM. The dashed line represents the ideal slope 1.

be positive through the enhanced correspondence between H_s^T and H_s^W in December 1993 expressed by a reduced σ_{p2} . However, the analysis of H_s^T and H_s^W (Fig. 2 a) shows a higher slope $b_p = 1.38$ (before 1.21) and a more negative intercept $a_p = -0.93$ m (before -0.58 m). Thus, the distribution is further rotated in an anti-clockwise direction suggesting a further reduction of the high H_s values in the H_s^W data set compared to H_s^T due to the assimilation of H_s^E .

The distribution of H_s^E and H_s^W (Fig. 2 b) is rotated in an anti-clockwise direction like that of H_s^T and H_s^W , but by a smaller amount. The slope b_p of the major principal axis is increased to 1.08 and its corresponding intercept a_p is now negative (-0.22 m). The distribution width σ_{p2} is 33 % smaller for H_s^E and H_s^W than for H_s^T and H_s^W . σ_{p2} decreases to 0.19 m, indicating excellent agreement between H_s^E and H_s^W . This is not surprising because H_s^E is assimilated into WAM, but it proves that the assimilation scheme works and H_s^W is corrected in the right direction.

Further details on the differences between the distributions of H_s^W and H_s^T and of H_s^W and H_s^E are seen by means of time series from monthly mean statistical parameters obtained throughout 1993. During 1993, no modifications in any of the data processing methods are known, except the start of wave data assimilation in August 1993. Therefore, we can assess the impact of H_s^E assimilation on H_s^W with the help of the independent data set of H_s^T .

The slopes b_p of the major PCA axes from the distributions of H_s^W and H_s^T are constantly larger than those of H_s^W and H_s^E by about 0.3 (Fig. 3). This gives evidence that H_s^T contains more and/or higher values of H_s than H_s^E . Staabs [1994] found that H_s^T reaches higher values on average in regions of high sea states than H_s^E . In regions of low sea states, such as the tropical oceans, H_s^E values were found to be higher than H_s^T values. In Fig. 3, the slopes b_p for both bivariate distributions are nearly constant through the first half of 1993. In Au-

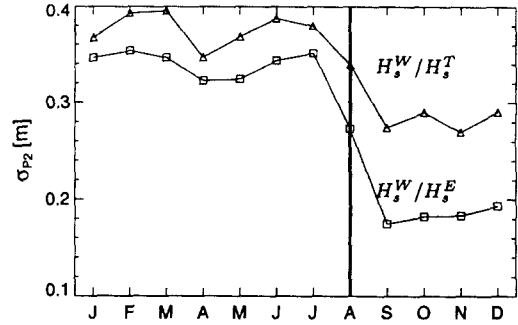


Fig. 4. Time series of monthly averaged minor principal rms deviations σ_{p2} for bivariate distributions of global H_s from WAM model collocated to Topex altimeter (triangles) and to ERS-1 altimeter (squares) for 1993. The vertical line indicates the start of wave data assimilation into WAM.

gust 1993, b_p increases by about 0.15 to 1.1 and 1.4 for the distributions of H_s^W and H_s^E and of H_s^W and H_s^T , respectively. The wave data assimilation causes a reduction of the high H_s^W values and an increase of the low H_s^W values, which results in a slight rotation of the bivariate data distribution.

The higher correspondence of model and satellite data after assimilation is indicated objectively by the reduction of σ_{p2} (Fig. 4). With the onset of H_s assimilation σ_{p2} falls off significantly from 35 cm to 18 cm for the distribution of H_s^W and H_s^E and from 38 cm to 27 cm for the distribution of H_s^W and H_s^T .

5 Conclusions

Principal component analysis, which is more suitable for the comparison of two data sets both containing noise than a linear regression analysis, is applied to global H_s data from ERS-1 and Topex altimeter measurements collocated to H_s from WAM model calculations. The slope b_p of the major PCA axis of the elliptical distribution suggests that H_s^T is overestimated at high sea states compared to H_s^W . The assimilation of H_s^E into WAM increases this effect since the slope b_p increases for both collocated data sets.

The PCA rms deviation σ_{p2} along the minor PCA axis measures the decorrelated noise of the bivariate distribution. Therefore, the degree of agreement between two collocated data sets is more adequately described by σ_{p2} than by the correlation coefficient r , which measures rather the agreement with a linear relationship. Throughout 1993, σ_{p2} shows smaller values for the distribution of H_s^E and H_s^W than for the distribution of H_s^T and H_s^W , confirming the higher correspondence between H_s^E and H_s^W . Note that a recent study shows that H_s from Topex are, on average, 5 % lower than those from the buoys, with rms scatter about the mean relation of 30 cm [Gower (1996)]. Cotton and Carter (1994) found both H_s^E OPR and H_s^E FDP to be smaller

than H_s^T , which seems to confirm our results. Taking into account that Queffelec and Lefevre (1992) found an underestimation of H_s^E against TOBIS buoys, one may conclude that both H_s^E FDP and H_s^W tend to underestimate H_s .

It is demonstrated that the assimilation scheme is able to update the calculations of WAM in the right way, which is expressed by a decreased σ_{p2} for H_s^W collocated to H_s^E . This is proven convincingly, because σ_{p2} decreases for H_s^W collocated to the independent H_s^T . However, the assimilation causes a rotation of the bivariate distributions which appears to overshoot the mark. Further investigations on the quality assessment of H_s data are currently carried out. These studies include the comparison against in-situ data [Cotton et al. (1997)], the use of improved wind analyses to run the WAM model [Sterl et al. (1997)] and the analysis of the univariate distribution functions [Bauer and Staabs (1997)]. From these studies more insight into the noise properties and the absolute precision of the different H_s data sources is expected.

Acknowledgements. This study is a contribution to the research program "Klimarelevante Prozesse im System Ozean-Atmosphäre-Kryosphäre" (SFB 318) funded by the Deutsche Forschungsgemeinschaft. It was also supported by the German Space Agency (DARA) under contract 50 EE 94 13. The data were provided by ESA under the ERS-1/2 projects PP2.D1 and AO2.D109.

References

- Barrick, D. E., Rough Surface Scattering Based on the Specular Point Theory, *IEEE Transactions on Antennas and Propagation*, AP-14, No. 4, 449-454, 1968.
- Bauer, E., S. Hasselmann, K. Hasselmann, and H. C. Graber, Validation and Assimilation of Seasat Altimeter Wave Heights Using the WAM Wave Model, *J. Geophys. Res.*, 97, 12,671-12,682, 1992.
- Bauer, E. and C. Staabs, Statistical properties of global significant wave heights and their use for validation, accepted by *J. Geophys. Res.*, 1997.
- Cardone, V.J., H.C. Graber, R.E. Jensen, S. Hasselmann, and M. J. Caruso, In search of the true surface wind field in swade IOP-1: Ocean wave modelling perspective, *Global Atmosphere Ocean System (GAOS)*, 3, 107-150, 1995.
- Cotton, P. D. and D. J. T. Carter, Cross calibration of TOPEX, ERS-1, and Geosat wave heights, *J. Geophys. Res.*, 99, 25,025-25,033, 1994.
- Cotton, P. D., S. K. Gulev, and A. Sterl, Intercomparison of the North Atlantic wave climatology from in-situ, voluntary observing, satellite data and modelling, accepted by *Physics and Chemistry of the Earth*, 1997.
- Gower, J. F. R., Intercalibration of wave and wind data from TOPEX/POSEIDON and moored buoys off the west coast of Canada, *J. Geophys. Res.*, 101, 3817-3829, 1996.
- Komen, G. J., L. Calvaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, Dynamics and modelling of ocean waves, *Cambridge University Press*, Cambridge, UK, 560 pp., 1994.
- Lionello, P., H. Günther, and B. Hansen, A sequential assimilation scheme applied to global wave analysis and prediction, *J. of Marine Systems*, 6, 87-107, 1995.
- Preisendorfer, R. W., Principal Component Analysis in Meteorology and Oceanography, *Elsevier Science Publishers*, 423 pp., 1988.
- Queffelec, P. and J. M. Lefevre, Validation of ERS-1 Altimeter Wave and Wind Fast Delivery Product, *ERS-1 Geophysical Validation Workshop Proceedings, ESA WPP-36*, (Penhors, Bretagne, France), 163-168, 1992.
- Staabs, C., Vergleich von Wellenhöhen berechnet mit dem Seeangangsmodell WAM mit Daten der Radarhöhenmesser an Bord der Satelliten ERS-1 und TOPEX/POSEIDON, Master thesis, Universität Hamburg, Hamburg, 88 pp., 1994.
- Sterl, A., G. J. Komen, and P.D. Cotton, 15 Years of global hindcasts using ERA winds, KNMI Report 97-07, 1997.
- WAMDI Group (S. Hasselmann, K. Hasselmann, E. Bauer, P. A. E. M. Janssen, G. J. Komen, L. Bertotti, P. Lionello, A. Guillaume, V. C. Cardone, J. A. Greenwood, M. Reistad, L. Zambresky, and J. A. Ewing), The WAM model - A third generation ocean wave prediction model, *J. Phys. Oceanogr.*, 14, 1775-1810, 1988.