ELSEVIER

Coastal Engineering 37 (1999) 235–257

COASTAL ENGINEERING

www.elsevier.com/locate/coastaleng

Methods for intercomparison of wave measurements

Harald E. Krogstad [a,*], Judith Wolf [b], Stephen P. Thompson [c], Lucy R. Wyatt [c]

[a] Department of Mathematical Sciences, NTNU, Trondheim 7034, Norway
[b] Proudman Oceanographic Laboratory, Bidston Observatory, Birkenhead, Merseyside L43 7RA, UK
[c] School of Mathematics and Statistics, Applied Mathematics Section, The University of Sheffield, Sheffield S3 7RH, UK

Abstract

The paper reviews methods for quality assessment and intercomparison of ocean wave data. The sampling variability for conventional time series recordings is summarized and compared to less common area measuring measurements. The sampling variability affects the scatter seen in simultaneous observations, and variability in excess of the sampling variability signifies real differences between the instruments. Various means of intercomparing wave parameters and spectra are discussed and two somewhat unconventional ways of deriving regression and calibration relationships are also shown. The methods are illustrated using data from SCAWVEX, focusing mainly on wave data from the HF radars and Directional Waveriders. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Method; Intercomparison; Wave measurement

1. Introduction

Ocean wave information is gathered with a variety of instruments for a multitude of needs. Over the years, a standard for measuring wave data with conventional instruments has evolved, and current recommendations for sampling and routine processing of in-situ data have been published by Tucker (1993). Remote sensing techniques are beginning to be used operationally (altimeter) and experimentally [Synthetic Aperture Radar (SAR), X-band marine radar, and HF coastal radar]. It is appropriate to both

* Corresponding author. Fax: +47-73-59-3524; e-mail: harald.krogstad@math.ntnu.no

PII: S0378-3839(99)00028-9

review standard intercomparison methods and to identify techniques that are most useful in assessing the accuracy of these new types of data.

There is often a need for merging data obtained from different sources into unified datasets for evaluation of wave climatology or for design criteria studies. Assessments, intercomparisons and subsequent calibrations of the data are important parts of such a process. Often, there are inherent instrument limitations due to the measurement principle which show up in the frequency domain by spectral distortion, high or low frequency cut-off, or in the time domain as, e.g., the horizontal motion of buoys giving inferior profiling capabilities. The present paper discusses techniques for wave data intercomparison with a focus on sea-state parameters and not on measurements of individual wave properties. The processing from the recorded signal to the final wave information may be quite complicated, e.g., for remote sensing techniques like the HF radar (Wyatt, 1990), but even for conventional instruments like the subsurface pressure cell, it has been demonstrated that the result may be considerably improved by proper processing (Wolf, 1997). For spatial arrays, the analysis has to assume linear wave theory and thus fails to record non-linear aspects of the waves.

With the ocean surface itself being a stochastic field, a certain sampling variability of the estimated parameters is unavoidable. Knowing the sampling variability is essential for assessing whether observed differences between the measurements are statistically significant. The theory for the sampling variability will be briefly reviewed, including a discussion of spatial vs. temporal measurements.

The paper continues with various ways of presenting data intercomparisons and discusses methods for optimal regression which is typically needed when different data sources are merged. The techniques are illustrated using data from SCAWVEX, in particular, data from Directional Waveriders and HF radars.

2. General remarks

Intercomparison of wave data highlights the problems encountered in comparing any dataset. The data are often collected from different kinds of instruments, using different sampling strategies and different analysis procedures. The data may be available in the form of directional spectra, or in the form of parameters derived from the spectra, leading to the question of which parameters are best to use in the comparisons. Various statistics must be chosen in an optimal way to illustrate the similarities and differences. An important consideration is the significance or confidence limits which can be applied to these statistics.

Some of the methods discussed for the intercomparison of wave model results with observations are also relevant to intercomparison of different instruments. Willmott et al. (1985) discuss various difference measures such as root-mean-square error, mean absolute error and index of agreement and apply the methodology to vector as well as scalar data. A set of complementary difference measures is recommended and also the 'bootstrap method' of assessing confidence and significance. Zambresky (1989) provides a useful list of standard wave parameters and statistics. Guillaume (1990) recommends some redefined wave direction variables and finds the mean relative error to be more useful than a scatter index for comparison of significant wave height. She also uses comparisons of frequency spectra between models and buoy data including confidence limits obtained from the buoy data.

Generally speaking, measurements of ocean waves involve estimation of parameters of random models. A central assumption about the random model is that it is stationary or homogeneous — a property which is never strictly attained in practice. Even if there are optimal space/time windows in which the wave field is stationary and the parameters can be estimated with maximal accuracy, no instrument existing today is close to reaching such accuracy for common sea-state parameters like the significant wave height.

Consider two wave instruments recording the same sea-states independently. The basic task will often be to reveal any systematic differences between the instruments, based on the actual measurements. Associated with each measurement, there are independent sampling errors; both instruments have in general systematic offsets (e.g., calibration errors) depending on the sea-state, and there may be some temporal and spatial offsets between the recordings. In addition, the underlying sea-states vary according to a certain natural variability which is beyond our control. We are thus facing several potential problems which have to be analysed and resolved properly:

Difference in measurement principles;

Inherent limitations of the measurement principles;

Systematic off-sets due to incomplete calibration;

Inherent and in general different sampling variability;

Temporal and/or spatial offsets.

Different instruments have different applications and, as long as they are known, inherent instrument limitations are not a problem. Whereas buoys are known to be excellent for measuring sea-state parameters, their surface profiling capability (for crest height, wave skewness, etc.) is less satisfactory. Subsurface instrumentation like the current meters/pressure cells or bottom-mounted pressure transducers have limited high frequency sensitivity simply due to the wave action attenuation with depth. Spatial arrays are in many respects different from point measurements with a sensitivity that may be dependent both on the frequency and direction of the incoming waves. Also, a spatial array is essentially limited to wave lengths longer than its size. Another feature of spatially extending instruments is that most analysis techniques need to assume linear wave theory. Many remote sensing measurement techniques such as the Synthetic Aperture Radar (SAR) suffer from a lack in basic understanding of the mechanisms. In these cases, there are also limitations due to the resolution which determines the minimum wavelength that can be observed at all.

A proper calibration of the instruments is essential for unbiased measurements as discussed in (Barstow et al., 1985). When measuring waves with heave/pitch/roll buoys, one typically has to consider several types of calibrations. The heave motion itself has a resonance which is dependent on the geometry and weight of the buoy. For medium-sized buoys, the resonance is above the main wave frequencies although some resonant enhancement and phase shift may extend into the high frequency range of the wave spectrum. On the contrary, the pitch and roll resonant motion is typically situated at an important range of the wave spectrum. This motion, which can be approximately modelled as a harmonic oscillator driven by a random force, may be strongly out-of-phase with the actual surface slope. In addition, electronic filters, e.g., integration of acceleration measurements and anti-aliasing filters, are frequently involved. As long as the response is linear, it is simple to make corrections by applying appropriate transfer functions to the spectrum, but often, one encounters a non-linear response for which it is very difficult to correct.

3. Wave parameters

3.1. Definitions

Linear stochastic wave theory is based on the concept of the *directional wave spectrum* which may be expressed either as a wavenumber spectrum, $\Psi(\boldsymbol{k})$, or a frequency/direction spectrum $S(f)D(\theta,f)$. Here, f is the frequency, $\boldsymbol{k}$ the wavenumber, θ the direction of $\boldsymbol{k}$, S is the frequency spectrum, and D the directional distribution. The two forms of the spectrum are connected by the dispersion relation:

$$S(f)D(\theta,f) = \Psi(k(f),\theta)k(f)\frac{\mathrm{d}k}{\mathrm{d}f}, \quad (2\pi f)^2 = gk(f)\tanh(k(f)h).$$

The directional distribution is, in general, dependent on frequency and may be expressed as the Fourier series:

$$D(\theta,f) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}\{a_n(f)\cos n\theta + b_n(f)\sin n\theta\}.$$

We refer to Tucker (1991) for a more detailed description of the stochastic wave model and the sea-state parameters defined from the spectra.

Main sea-state parameters used in the following are defined in Table 1.

3.2. Sampling variability

The sampling theory for time series in the spectral domain is well-known (see, e.g., Brillinger, 1975; Tucker, 1991). The *periodogram*, I, is the squared magnitude of the discrete Fourier transform of the time series and, with smooth spectra and reasonably short correlation times as in the present case, the Central Limit Theorem for discrete Fourier transforms ensures that the periodogram values, $I(f_k)$, are (scaled) χ^2-distributed variables with 2 degrees of freedom (DOF). Bias in the periodogram, often called *spectral leakage*, is reduced by data tapering. In the following, we discuss estimates of wave spectra and wave parameters. We shall use a *caret* (^) to signify an estimate, $\hat{a}$, of the parameter a.

Many of the sea-state parameters dependent on the frequency spectrum are derived from the spectral moments, m_r, defined in Table 1. The estimates of the spectral moments, $\hat{m}_r$, are sums over the periodogram for which it follows that expectation, $E(\hat{m}_r) \approx m_r$, and the covariance:

$$\mathrm{Cov}(\hat{m}_r,\hat{m}_s) = \frac{1}{T}\int_{f=0}^{\infty} f^{r+s}S^2(f)\mathrm{d}f + O(N^{-2}),$$

Table 1
Main sea-state parameters

Name	Symbol	Definition
Significant wave height	$Hm0$ (Hs)	$Hm0 = 4m_0^{1/2},\ m_k = \int_{f=0}^{\infty} f^k S(f)\mathrm{d}f$
Mean zero-crossing period	$Tm02$ (Tz)	$Tm02 = (m_0/m_2)^{1/2}$
Mean wave period	$Tm01$	$Tm01 = m_0/m_1$
Peak period	Tp	$Tp = 1/f_\mathrm{p}$, $\max_f S(f) = S(f_\mathrm{p})$
Mean wave direction	θ_1	$\theta_1(f) = \mathrm{atan2}(b_1(f), a_1(f))$
Directional spread	σ_1	$\sigma_1(f) = (2(1-r))^{1/2},\ r = [a_1^2(f) + b_1^2(f)]^{1/2}$
Direction at the spectral peak	θ_p	$\theta_\mathrm{p} = \theta_1(f = 1/Tp)$
Main wave direction	MDIR	MDIR $= \mathrm{atan2}(b, a)$ where a and b are spectrally weighted averages of a_1 and b_1.
Spread at the spectral peak	σ_p	$\sigma_\mathrm{p} = \sigma_1(f = 1/Tp)$

where T is the recording interval and N the number of points in the time series (see, e.g., Krogstad, 1982). In practice, moments and covariances may also be estimated from a smoothed spectrum estimate, $\hat{S}$, which is a χ^2-distributed variable fulfilling:

$$E\left(\hat{S}(f)\right) \approx S(f),$$

$$\mathrm{Var}\left(\hat{S}(f)\right) \approx (2/\nu) S^2(f),$$

$$E\left(\hat{S}^2(f)\right) \approx (1 + 2/\nu) S^2(f),$$

where ν is the DOF equal to the number of periodogram values involved in the smoothing. There is a trade-off between the resolution of the smoothed spectrum and the requirement of independent spectral estimates. For a sampling frequency f_s, the recording interval is $T = N/f_\mathrm{s}$ and the periodogram frequency resolution is f_s/N. The maximum frequency resolution in a smoothed spectrum with ν DOF which maintains independent spectral estimates is therefore $\Delta f = (\nu/2) f_\mathrm{s}/N$.

A Taylor expansion technique may be applied to determine the variance for estimates of parameters which are functions of the spectral moments:

$$\widehat{Hm0} = 4\hat{m}_0^{1/2}, \quad \mathrm{Var}\left(\widehat{Hm0}\right) = 4 m_{00}/m_0,$$

$$\widehat{Tm01} = \hat{m}_0/\hat{m}_1, \quad \mathrm{Var}\left(\widehat{Tm01}\right) = \frac{m_{00}}{m_1^2} - 2\frac{m_0 m_{01}}{m_1^3} + \frac{m_0^2 m_{11}}{m_1^4},$$

$$\widehat{Tm02} = \left(\hat{m}_0/\hat{m}_2\right)^{1/2}, \quad \mathrm{Var}\left(\widehat{Tm02}\right) = \frac{1}{4}\left(\frac{m_{00}}{m_0 m_2} - 2\frac{m_{02}}{m_2^2} + \frac{m_0 m_{22}}{m_2^3}\right),$$

where $m_{rs} \equiv \mathrm{Cov}(m_r, m_s)$ (see, e.g., Krogstad, 1982). For the HF radar (Wyatt et al., 1998), the DOFs in the frequency spectrum vary with frequency. The relations are easily modified to handle such a case.

The sampling variability for the directional Fourier coefficients and directional parameters obtained from heave/pitch/roll buoy data was derived by Long (1980). Virtually identical expressions exist for displacement buoy data or for any single point triplet measurements, since the definitions of the directional Fourier coefficients from the cross-spectra of the time series are similar. Long's expressions are independent of frequency and the variability is inversely dependent on the DOF in the cross-spectra. The expressions are rather complicated, but it turns out that the precision in estimates of the mean direction and the directional spread is strongly dependent on the peakedness of the directional distribution itself, as shown in Fig. 1. It is therefore meaningless to give an absolute statement about the directional resolution of the Directional Waverider or a heave/pitch/roll buoy. A well-defined mean direction in the directional distribution leads to a small sampling error in the estimates.

When comparing frequency spectra from two different instruments, it is convenient to consider their *ratio*:

$$r(f) = \frac{S_Y(f)}{S_X(f)}.$$

If the spectra for the two systems have been computed with ν_X and ν_Y DOF, an unbiased estimate for the spectral ratio is:

$$\hat{r}(f) = \frac{\hat{S}_Y(f)}{\hat{S}_X(f)} \frac{\nu_X - 2}{\nu_X}.$$

This spectral ratio is a (scaled) Fisher distributed variate with ν_Y and ν_X DOF and a variance equal to:

$$\mathrm{Var}(\hat{r}(f)) = r^2(f) \frac{2(\nu_X + \nu_Y - 2)}{\nu_X(\nu_Y - 4)}.$$

An alternative to the ratio is to consider the logarithm since $U = \log(\hat{S}/E(\hat{S}))$ has a probability density which is easily expressed in terms of the χ_ν^2-distribution and is more symmetric about its mean. In particular:

$$E(U) = \psi(\nu/2) - \log(\nu/2),$$

$$\mathrm{Var}(U) = \psi'(\nu/2),$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the *di-Γ-function* (Abramowitz and Stegun, 1965, Section 26.4.37). From the asymptotic expansion of ψ, we the obtain to the first order:

$$E\left(\log\left(\frac{\hat{S}_Y(f)}{\hat{S}_X(f)}\right)\right) \approx \log\left(\frac{S_Y(f)}{S_X(f)}\right) - \frac{1}{\nu_Y} + \frac{1}{\nu_X},$$

$$\mathrm{Var}\left(\log\left(\frac{\hat{S}_Y(f)}{\hat{S}_X(f)}\right)\right) \approx \frac{2}{\nu_Y} + \frac{2}{\nu_X}.$$

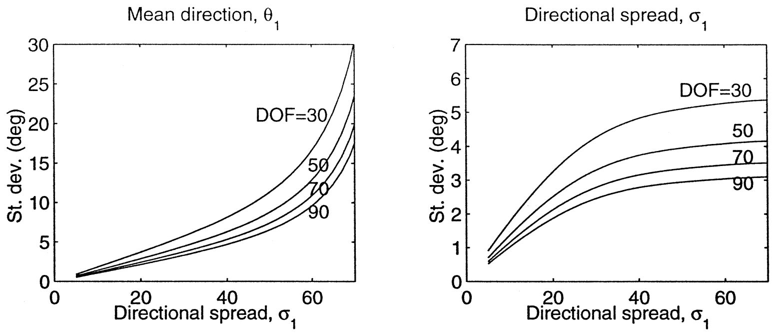


Fig. 1. Sampling variability for the mean direction (left) and directional spread (right) as a function of the directional spread when the directional distribution is a cos-2s-distribution.

Modern statistical methods frequently use computer simulation to reveal the sampling variability. For given spectra, there are simple and effective ways of simulating multivariate Gaussian time series. However, it is often more convenient to simulate the cross-spectra directly. The complex Wishart distribution of cross-spectral estimates is easily simulated by summing squares of independent complex Gaussian variables (Krogstad, 1989).

In a simulation study of the sampling variability of sea-state parameters obtained from buoys like the Directional Waverider (Munthe-Kaas and Krogstad, 1985), the variability of parameters derived from the spectrum (e.g., $Hm0$, $Tm01$, $Tm02$) agreed well with the Taylor expansion technique using the expressions for the variability of the spectral moments. As compared to the more conventional zero-crossing parameters, the spectral definitions were generally favoured as having less sampling variability. More important for spectral definitions is, however, the easier calibration of systematic errors due to various filters affecting the time series. The simulation study also showed that the standard deviation of estimates of Tp is about 4–10 times larger than for $Tm02$. For the directional parameters, the expressions of Long (1980) seem to be accurate for narrow directional distributions whereas some deviations were observed for very broad directional distributions.

HF radar wave measurements are made through an inversion of a non-linear integral relationship between the radar power (Doppler) spectrum and the directional wave spectrum. The only way to determine the sampling variability of wave parameters is by the use of simulation techniques (Sova, 1995). A procedure has been implemented that determines the sampling variability as a function of frequency of the frequency spectrum, the mean direction, and the directional spread. This involves, firstly, determining the sampling variability of the Doppler spectrum, dependent as usual on the length of the dataset and the amount of averaging and overlapping in the spectral analysis. This sampling variability is then used to simulate a large number of realisations of Doppler spectra for a range of specified sea conditions. These are inverted and the DOF and variance in the parameter estimates are determined. This procedure has to be carried out in full for any change in operating frequency, sampling frequency or averaging procedure. However, only small differences were found between the sampling statistics for the Ocean Surface Current Radar (OSCR) and Wellen Radar (WERA) radars used in the SCAWVEX project.

In order to construct a full directional spectrum from conventional measurements like displacement or heave/pitch/roll buoys, the directional distributions have to be estimated from a small set of Fourier coefficients. There are several ways of obtaining the distributions. The simplest is beam-forming expressions of the form $\hat{D}(\theta,f)=\gamma(\theta)^H\hat{\Phi}(f)\gamma(\theta)$ where $\gamma(\theta)$ are fixed weight factors and $\hat{\Phi}$ is the frequency-dependent cross-spectral matrix from the data. Since the cross-spectral matrix will asymptotically have a complex Wishart distribution (Brillinger, 1975) it is easily proved that $\mathrm{Var}(\hat{D}(\theta,f))\approx 2D(\theta,f)^2/\nu$, where again ν is the DOF in the spectral estimates. The variability for non-linear estimates like the Capon maximum likelihood estimate or the maximum entropy (ME) estimate (Lygre and Krogstad, 1986) is harder to assess. Triplet data produce four Fourier coefficients for which the ME estimate takes the form:

$$\hat{D}(\theta,f)=\frac{1}{2\pi}\frac{1-\phi_1c_1^*-\phi_2c_2^*}{|1-\phi_1\mathrm{e}^{-i\theta}-\phi_2\mathrm{e}^{-i2\theta}|^2},$$

where * denotes complex conjugate, $c_i=a_i+ib_i$, and $\phi_1=(c_1-c_2c_1^*)/(1-|c_1|^2)$, $\phi_2=c_2-\phi_1c_1$ (the dependence of f is omitted for clarity). It may be shown that the variability of this estimate is three times the variability of the linear estimate for very flat distributions. Computer simulations have revealed that the variability may be much more for more complicated distributions (see Krogstad, 1991). The considerable variability of the resulting distributions suggests that more spectral smoothing should be used for the Fourier coefficients than is typically needed for the one-dimensional frequency spectrum. Some examples of the actual sampling variability of the final directional spectra are shown in Krogstad (1991).

3.3. Spatial vs. temporal measurements

With the development of remote sensing techniques, we are now having instrumentations which record the ocean surface properties over the area rather than time series in fixed points. It is therefore of interest to compare the sampling variability for these quite different ways of obtaining wave information. If we consider the measurement of significant wave height by an estimate of m_0 from an (so far hypothetical) instrument recording the instantaneous surface over a square region with area A, a similar technique as above shows that:

$$\mathrm{Var}(\hat{m}_0)=\frac{4\pi^2}{A}\int_{\mathbf{k}}\Psi^2(\mathbf{k})\mathrm{d}^2k.$$

Consider now two instruments where one measures a time series of surface elevation at a single point for a duration T and the other measures the elevation of the surface at a fixed instant of time for all points over an area A. Assume deep water for simplicity and write the wave spectrum as:

$$S(f)D(\theta,f)=\frac{Hm0^2}{16f_\mathrm{p}}S_0(f/f_\mathrm{p})D_0(\theta,f/f_\mathrm{p}),$$

where

$$f_p = 1/Tp \text{ and } \int_{x=0}^{\infty} S_0(x)\,\mathrm{d}x = \int_{\theta=0}^{2\pi} D_0(\theta, x)\,\mathrm{d}\theta = 1.$$

Since $\mathrm{Var}(Hm0) = 4\mathrm{Var}(m_0)/m_0$, we have for the *coefficient of variation*, COV, of the first instrument:

$$\mathrm{COV}_{\mathrm{time}} = \frac{\mathrm{std}(Hm0)}{Hm0} = \frac{\sqrt{\mathrm{Var}(\hat{m}_0)}}{2m_0} = \frac{1}{2}\|S_0\|_2 (f_p T)^{-1/2},$$

where $\|S_0\|_2 = (\int S_0(x)^2 \mathrm{d}x)^{1/2}$. Similarly, for the spatial instrument:

$$\mathrm{COV}_{\mathrm{space}} = \frac{\mathrm{std}(Hm0)}{Hm0} = \|D_0(\theta, f_p)\|_2 \|S_0\|_2 (d/\lambda_p)^{-1},$$

where d is diameter of the recording area and λ_p is the wavelength corresponding to Tp. With a typical JONSWAP-like frequency spectrum and a cos-2s-distribution with $s = 10$ at Tp for D, we obtain the following approximate expressions in deep water:

$$\mathrm{COV}_{\mathrm{time}} \approx 0.52(Tp/T)^{1/2},$$

$$\mathrm{COV}_{\mathrm{space}} \approx 0.33(\lambda_p/d).$$

which are shown graphically in Fig. 2. The behaviour with varying wave period is therefore quite different for the two instruments, obviously due to the quadratic dependence of wavelength on wave period. Shallow water will affect the relations somewhat, as is easily seen.

In this paper, we will consider the HF radar as if it were a temporal measurement. It is measuring the time series of backscatter but from an area of the surface and not a point. If we assume that both diagrams in Fig. 2 apply, then, for a 10-min duration, 1

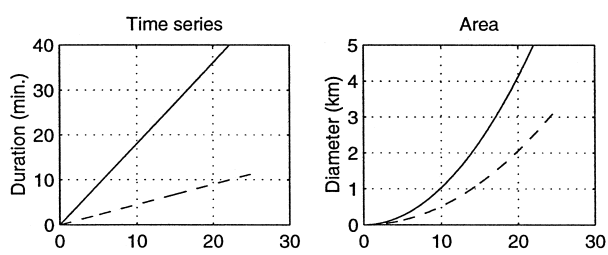


Fig. 2. Necessary recording time interval (left) and area diameter (right) as a function of the peak period Tp. Solid line corresponds to COV = 5%, dashed line to COV = 10%.

km^2 area, we anticipate that the parameters of such a system have large variability for very long waves.

4. Data intercomparisons

A first step in a data comparison will typically be joint occurrence tables, scatter plots and descriptive statistics like bias, correlation coefficients, average scatter, etc. Although simple to produce, such presentations alone are, in general, not sufficient to explain differences between the datasets and whether these are in excess of the sampling variability. Often, the high and low frequency parts of the spectra compare differently and direct intercomparison of spectra is the only way to assess differences that show up in the integrated parameters such as the significant wave height and the mean period. Below, we present various ways to intercompare data and illustrate the methodology using data from SCAWVEX.

4.1. Scatter diagrams and calibration relations

The scatter diagram suggests the computation of some kind of regression between the variables. If the variability of one variable (say X) is much less than the other, one-sided regression, e.g., $Y = aX + b$, is appropriate. Sometimes, it is feasible to let the line go through the origin in which case $b = 0$. A power relation $Y = bX^a$ becomes linear by a logarithmic transformation, and angular variables which are defined modulo 2π, may be fitted to $Y = X + b$.

In the present case, both variables may have comparable variability and symmetric methods, e.g., principal component regression, are better. If we assume that the variances for both instruments and each measurement are equal and we seek a line through the origin, the slope of the line can be determined by minimising the function $\sum_{i=1}^{n}(Y(i) - bx(i))^2/(1 + b^2)$. This can be interpreted as either minimising the sum of squares of the perpendicular distances of datapoints from the line $y = bx$, or as a least-squares problem weighted by the sum of variances. In the latter case, the variance should appear in the denominator of the function being minimised but since it is constant, it does not affect the resulting expression for b. This method is referred to as two-sided regression below.

However, the sampling variabilities may be highly variable and different for the two variables. This situation can be accounted for in a ML approach suggested by Sova (1995). Assuming that the two measurements of interest are Gaussian with a linear relationship between their means, i.e., $X_j \sim N(\mu_j, \sigma_{X_j}^2)$, $Y_j \sim N(a + b\mu_j,\ \sigma_{Y_j}^2)$, the parameters (slope, b, and intercept, a) can be determined by maximising the log likelihood function:

$$\mathrm{loglik}\left(a, b, \mu | X, Y, \sigma_X^2, \sigma_Y^2\right)$$
$$\propto -\sum_{j=1}^{n}\left\{\log\left(\sigma_{X_j}^2\sigma_{Y_j}^2\right) + \frac{\left(X_j - \mu_j\right)^2}{\sigma_{X_j}^2} + \frac{\left(Y_j - a - b\mu_j\right)^2}{\sigma_{Y_j}^2}\right\}.$$

(Note that μ_j, $j = 1, \dots, n$, can be derived explicitly and inserted in the expression such that only a and b remain). The relative bias between the two instruments and its 95% confidence interval is given by:

$$y_{\mathrm{rel}} = \frac{\hat{a} + (\hat{b} - 1)x}{x},$$

$$y_{\mathrm{rel}} \pm 1.96\left(\mathrm{Var}(\hat{a})/x^2 + \mathrm{Var}(\hat{b}) + 2\mathrm{Cov}(\hat{a},\hat{b})/x\right)^{1/2}.$$

(Again, the *caret* signifies the estimate of the parameter.) A power relationship can be handled using logarithms, and in the directional case, one assumes that the directional difference is normal, i.e., $\psi_j = \theta_{1j} - \theta_{2j} \sim N(\mu, \sigma_j^2)$.

For a real stochastic variable X, the cumulative distribution function, F_X, is defined as $F_X(x) = P(X \le x)$. There is actually a simple and direct way of obtaining a *non-parametric* regression function $Y = h(X)$ between two arbitrary wave parameters X and Y from their cumulative distribution functions F_X, and F_Y. If we disregard sampling errors and the regression function is ever increasing:

$$F_X(x) = P(X \le x) = P(h(X) \le h(x)) = F_Y(h(x)),$$

from which it follows that $y = h(x) = F_Y^{-1}(F_X(x))$. For an observed dataset (X_i, Y_i), $i = 1, \dots, n$, the piecewise linear function defined by the *ranked* observations, $X_1^* \le X_2^* \le \dots \le X_n^*$ and $Y_1 \le Y_2^* \le \dots \le Y_n^*$, will be a simple estimate of h. Since sampling errors typically stretch the sampling distributions as compared to the exact distributions, this will introduce some bias if the sampling error is large as compared to the variations of the variables, or if the sampling errors are highly different for X and Y.

There are actually several more advanced approaches to the problem of finding a best possible calibration or regression curve with errors in both variables. *Principal curve regression* aims to determine the curve with the least possible distance to the data subject to smoothness constraints (Hastie and Stuetzle, 1989). Since the sampling variability in the present case is typically strongly dependent on the magnitude of the variables, a weighted regression is however preferred (Boggs and Rogers, 1990; Fan and Truong, 1992). Application of these methods to wave data is not known.

The ML methods are already somewhat more complicated to implement than simple regression because they need estimates of the sampling variability. One question that arises is whether they provide additional information of use in assessing a particular measurement system or is the additional rigour providing results that are in fact very similar to those achieved with the more straightforward methods and hence, simply confirming their validity. Another question concerns the validity of assuming a linear relationship. A subjective judgement has to be made based on the nature of the scatter plot or the non-parametric regression introduced above. Actually, the ML methodology may also be extended to piecewise linear or cubic splines relationships.

Some of these issues are discussed in more detail in (Wyatt et al., 1998) where the ML method has been applied for a detailed assessment of the accuracy of HF radar vs. buoy measurements. Below, we show some examples for the main wave parameters, *Hm*0, *Tm*01 and the spectrally averaged main wave direction (MDIR). The scatter plots and the ML regression lines for simultaneous and co-located data from the Holderness 2 and Petten experiments are shown in Fig. 3. Fig. 4 compares the non-parametric regression to the ML regression and Table 2 shows a number of statistics for the same parameters. If X is the wavebuoy and Y the radar measurement, the mean error is given by $N^{-1}\sum_{i=1}^{N}(Y_i - X_i)$, the relative error by $N^{-1}\sum_{i=1}^{N}(Y_i - X_i)/X_i$, and the two-sided regression uses a slope through the origin. For the maximum likelihood (ML) method, the slope, b, and the intercept, a, are shown together with the relative biases over specified waveheight or period ranges, as appropriate.

For significant waveheight, the mean error and its standard deviation are not very informative, they give the impression of much better agreement than can be justified from the scatter plot. The two-sided regression (with zero intercept) and the relative error estimate are significantly different for the two experiments but the slope of the ML line is similar. The increased bias shown at Petten is accounted for in the change in intercept of the line and is in fact larger at low than at high waveheights. Over most of the range of waveheights in this experiment, the bias found using this method is less than is inferred from the other two estimates. This consistency in slope from experiment to experiment in the ML description suggests that it might provide a more robust estimate of a relationship between the radar and wavebuoy measurements that could be used to calibrate the radar measurements. However, the validity of the assumption that there is a straight-line relationship for the entire waveheight range for Petten could be questioned. The non-parametric regression shows evidence of a change in the nature of

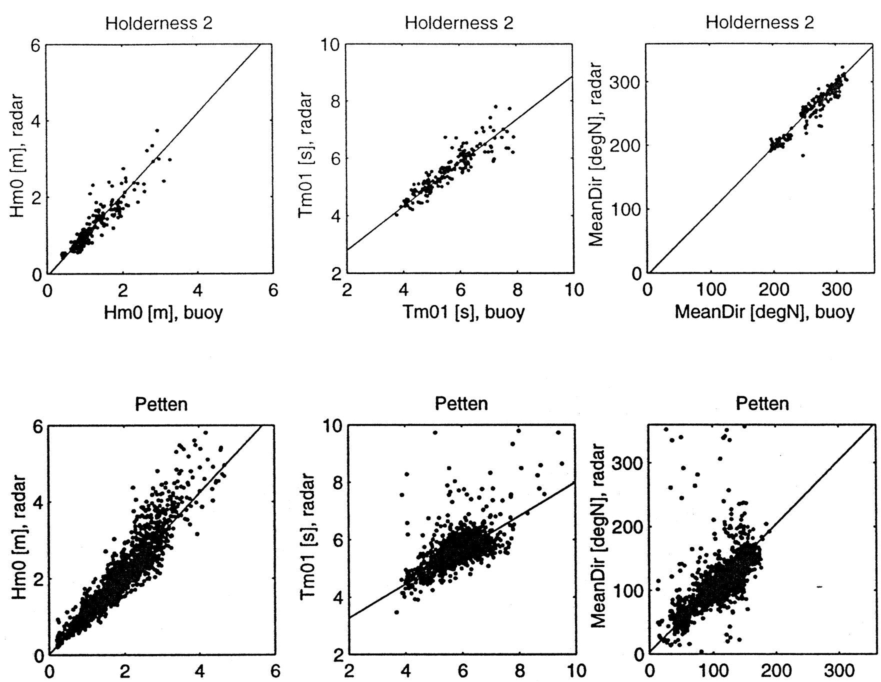


Fig. 3. Scatter plots of significant wave height, $Hm0$, mean wave period, $Tm01$, and integrated mean wave direction, MDIR, for buoy and radar data from Holderness (177 datapoints) and Petten (1663 datapoints).

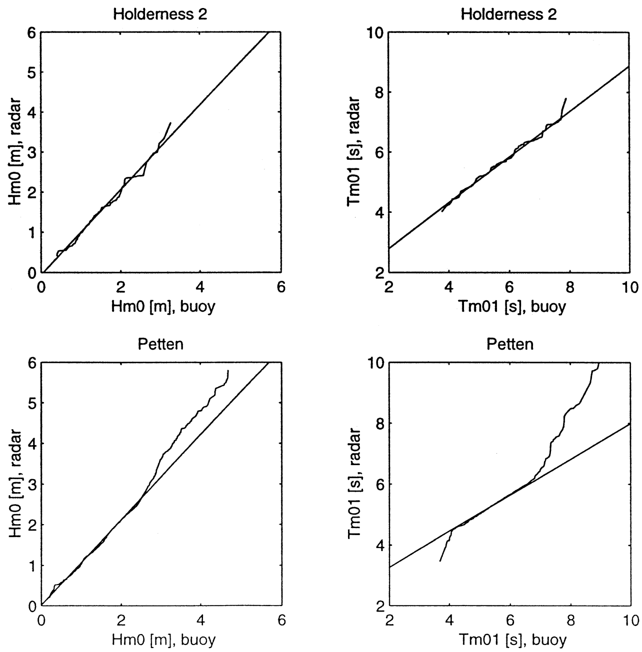


Fig. 4. The ML lines and non-parametric regression curves for wave height and period for the Holderness 2 and Petten datasets (same data as in Fig. 3).

the relationship in the higher sea-states which is driving the change in intercept and hence in the estimated bias.

There is a lot more scatter in the Petten data for mean period than is seen in the Holderness data. This is reflected in the standard deviation of the relative error which is once again a better descriptor than the mean error. The mean, relative and regression estimates of bias give small biases for both experiments in contrast to the ML method. The latter is showing a large negative bias when the period is large, as is also obvious from Fig. 4. In this paper a negative bias means the radar measurement is smaller than the buoy measurements, and a positive bias is the opposite. However, there are a number of measurements with large positive bias in the Petten data set. This partly compensates for the negative bias at the higher periods driving the relative error towards zero but these are of less significance in the ML calculation because the sampling variabilities associated with these measurements are larger. The large positive bias cases, in particular at the larger periods, also drive the non-parametric regression to imply an overestimation of this parameter by the radar. These cases are associated predominantly with low amplitude swell conditions when the radar measurements are contaminated by spurious contributions from antenna sidelobes and surface current variability (see the discussion in Sections 4.2 and 4.3 below). The larger negative bias at higher periods is identified more clearly with the ML method and is associated with an increased contribution to the radar measured waveheight spectrum at high frequencies in high sea-states. The larger biases in wave height in the Petten data are also due to this

Table 2
Statistics for the radar–buoy intercomparisons

Parameter	Method	OSCR at Holderness		WERA at Petten	
		Bias	Standard deviation	Bias	Standard deviation
*Hm*0	Mean error	1 cm	25 cm	23 cm	40 cm
	Two-sided regression	3%	18%	14%	28%
	Relative error	+0.8%	18%	+14%	28%
	ML parameters, relative bias	$b,a = 1.06, -0.06$		$b,a = 1.05, 0.02$	
		< 0.7 m: −5%			
		0.7–3.2 m: −5%–5%		5–10%	
*Tm*01	Mean error	−0.03 s	0.5 s	−0.07 s	0.7 s
	Two-sided regression	−1%	8%	−1%	13%
	Relative error	0.4%	8%	−0.5%	13%
	ML parameters, relative bias	$b,a = 0.77, 1.25$		$b,a = 0.6, 2.04$	
		< 5 s: 0–10%		< 5 s: 0–10%	
		5–7 s: −5–0%		5–6 s: −5–0%	
		7–8 s: −5–10%		6–8 s: −10%	
MDIR	Mean difference	2°	8°	11°	18°
	ML	−3.5°		3.5°	

See text for explanation.

phenomenon which is associated with an increase in the backscattered signal in higher sea-states that is not explained by the theory used for HF radar wave measurement. Note that the ML regressions for the mean period for the two experiments are very different and hence a consistent calibration for this parameter is not possible. However, as has already been said, the method does provide a more reliable measure of bias in the period parameter than the standard statistics or the non-linear regression.

The ML estimate of the direction difference is also similar although biased positive at Petten and negative at Holderness. This difference in sign is related to the different range of directions measured at the two sites which are on opposite coasts of the North Sea. Once again, the consistency is suggestive of a more robust relationship than is given by the mean difference.

When comparing similar instruments, the use of *confidence regions* in the scatter diagrams is useful (Allender et al., 1989). The confidence regions are here defined by two lines, which, for a certain model of the sampling variability, on the average should enclose a certain fraction of the datapoints. As an example, assume that both instruments should measure the same apart from the sampling variability, i.e., $E(X) = E(Y)$ for simultaneous pairs. Assume further that the sampling variabilities are equal but independent with a standard deviation that increases linearly with the expectation, $\mathrm{std}(X) = \mathrm{std}(Y) = aE(X)$. Two lines through the origin which approximately enclose a fraction p of the datapoints are then given by $y_{\pm} = \tan(\pi/4 \pm \delta)x$, where $\sin(\delta) = a\gamma_{\mathrm{p}}/2^{1/2}$,

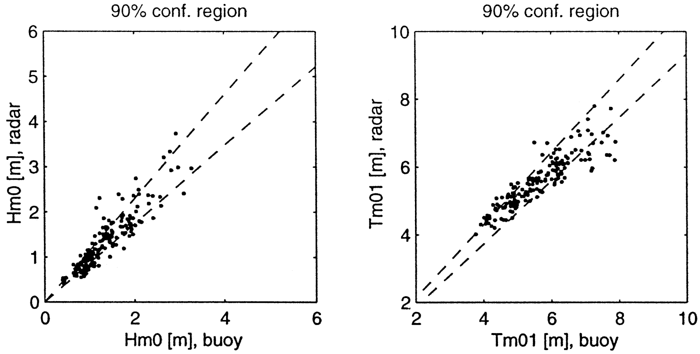


Fig. 5. Confidence regions based on pairs of ordered variables for wave height and period from Holderness.

$2\Phi(\gamma_p) - 1 = p$, and Φ is the cumulative standard normal distribution. Using the Holderness data as an example, the COV was computed to about 4–6% for the buoy wave height and about 3–5% for the radar (Wyatt et al., 1998). For the wave period, the COV was similarly computed to 2–3% for the buoy and 1–2% for the radar. Neglecting the bias for the moment, 90% confidence regions based on a common 6% COV for the wave height and 3% for the period are shown in Fig. 5. It turns out that the fraction of the datapoints inside the 90% limits is merely 57% for wave height and 66% for the wave period. We obtain approximately 90% enclosure by increasing the COV for wave height to 12% and to 4% for the wave period. We therefore conclude that the sampling variability explains almost all the scatter seen in the wave period despite the obvious bias discussed above, whereas wave height has additional variability not accounted for by the computed sampling variability.

Another example where computer simulation may be used to assess sampling variability is illustrated in Fig. 6. On the left is shown a scatter plot for the significant wave height between 1365 near coincidences of the Topex radar Ku altimeter on the Topex/Poseidon satellite and 13 NOAA buoys. How accurate are then the altimeter

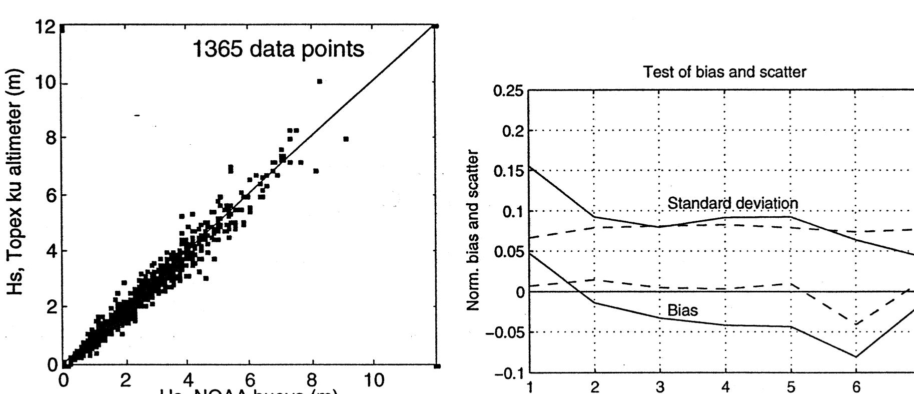


Fig. 6. Significant wave height from 13 NOAA buoys and the Topex altimeter (left). Bias and scatter for the actual dataset (solid) and an equivalent simulated buoy dataset (dashed) shown to the right.

measurements (in this case, the standard data product obtained from AVISO, France) as compared to the buoys? The sampling variability for the altimeter data is not known, and hence, the approach adopted for the HF radar measurements cannot be used, but we would assume that the NOAA data are processed with a sampling variability typical for 20 min buoy measurements. The idea is then to compare the bias and standard deviation of the differences in the observations to a *simulated* dataset consisting of an equally large and equally distributed dataset from two independent buoy systems. The sampling variability in the buoy measurements is based on the theory in Section 3.2. The interesting conclusion, as shown to the right in Fig. 6, is that the observed scatter is quite comparable to the scatter in the simulations above 2 m wave height, although some minor bias in the measurements is likely. The altimeter data for this study were averaged over three points, corresponding to about 20 km along the track, and, judging from the simulations, the data have a sampling variability quite comparable to buoy measurements. The somewhat increased variability below 2 m may probably be due to the spatial offsets between the altimeter and the buoy.

4.2. Comparisons of frequency spectra

The simplest way of comparing frequency spectra is to superimpose individual spectra on the same plot. However, the sampling variability makes it difficult to find consistent differences and the mean spectral ratio is useful when looking for consistent biases over fixed frequency ranges and hence to establish a spectral calibration that could be used. For this to be effective, the biases have to be consistent and not associated with occasional instrument problems as illustrated in the following example. Fig. 7 shows the mean spectral ratio at Holderness $\hat{r}(f) = \hat{S}_{\mathrm{hf}}(f)(\nu_{\mathrm{wb}} - 2)/(\hat{S}_{\mathrm{wb}}(f\nu_{\mathrm{wb}}))$ where hf refers to the HF radar measurement and wb to the buoy. The figure shows that, on average, the radar overestimates energy at the low frequencies. However, this is nearly all an overestimation at times when the wavebuoy energy is low in these frequencies, i.e., when these frequencies are below the spectral peak frequency. This is confirmed in Fig. 8 which shows the spectral ratio when only those frequencies at and above the wavebuoy peak are included. This shows a value of the ratio nearer to 1 with

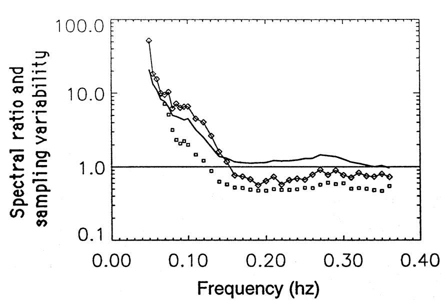


Fig. 7. The mean spectral ratio (radar/buoy) is shown as a solid line. The line with diamonds shows the measured variance which is a little higher than the variance that is due to sampling variability (shown with squares).

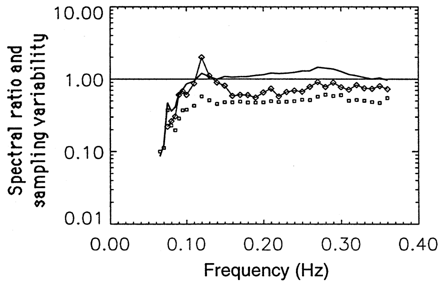


Fig. 8. Same as Fig. 7 except that energy below the buoy spectral peak is excluded from the calculation of the means.

an underestimation at the very low frequency end. The source of the problem is the occasional appearance of spurious contributions to the spectrum due to antenna side-lobes and/or current variability. Their presence or absence in the Holderness dataset depends on the prevailing wind direction and the phase of the tide. Since these are sporadic rather than consistent contributions to bias, the use of a calibration that includes such contributions is not advised. These can be identified clearly in the comparisons of directional spectra (see below) and are discussed in more detail in (Wyatt et al., 1998).

The graphs also illustrate the use of observed vs. intrinsic sampling variability in the spectral ratio. The observed variability is as expected somewhat in excess of the unavoidable intrinsic variability. However, in particular for high frequencies, it is not much larger. This indicates by and large that the ratio between the spectra is essentially constant for high frequencies. Around the spectral peak, the variability is larger, probably due to the fast variations in the spectra.

4.3. Comparisons of directional spectra

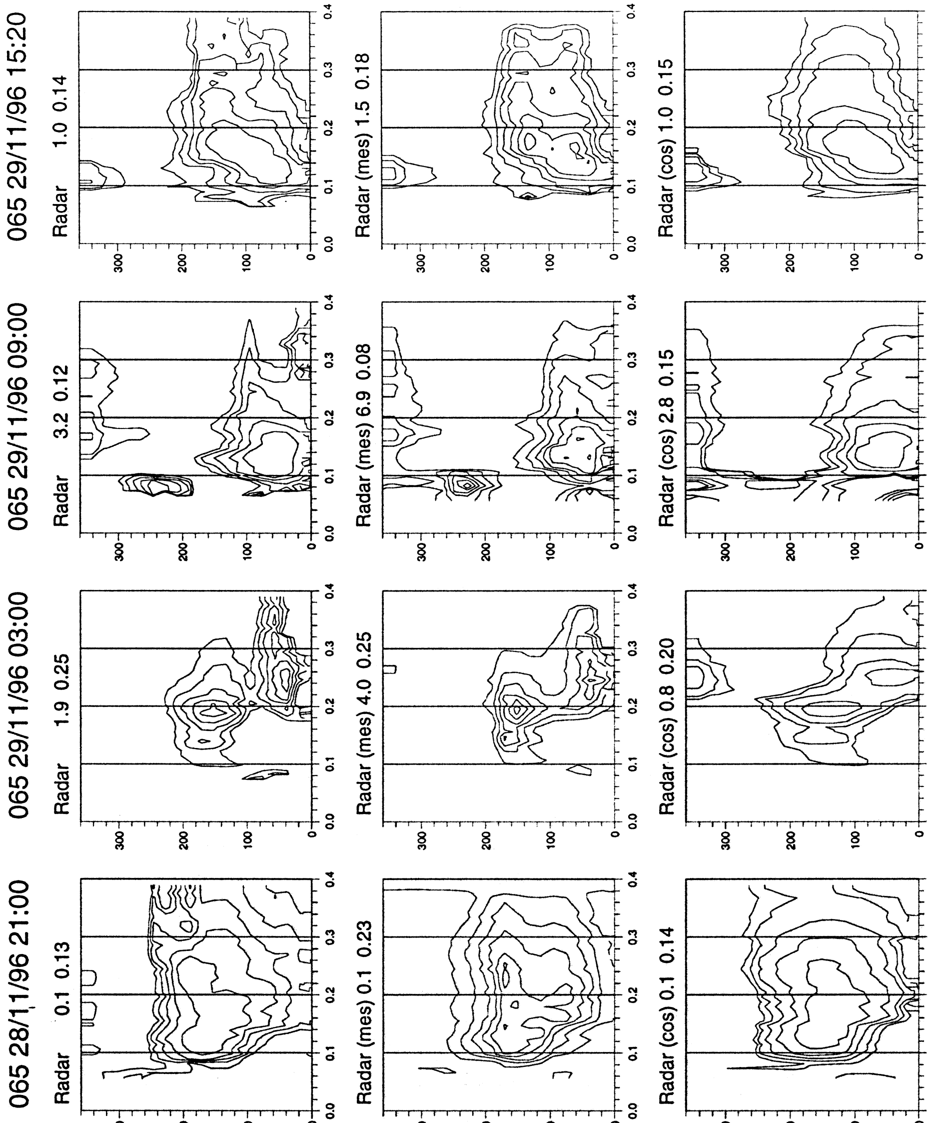
Due to the large dynamic range of directional spectra, direct comparisons are not straightforward. A simple compromise is to plot frequency-dependent parameters like the mean direction and the directional spread superimposed on plots similar to superimposed frequency spectra.

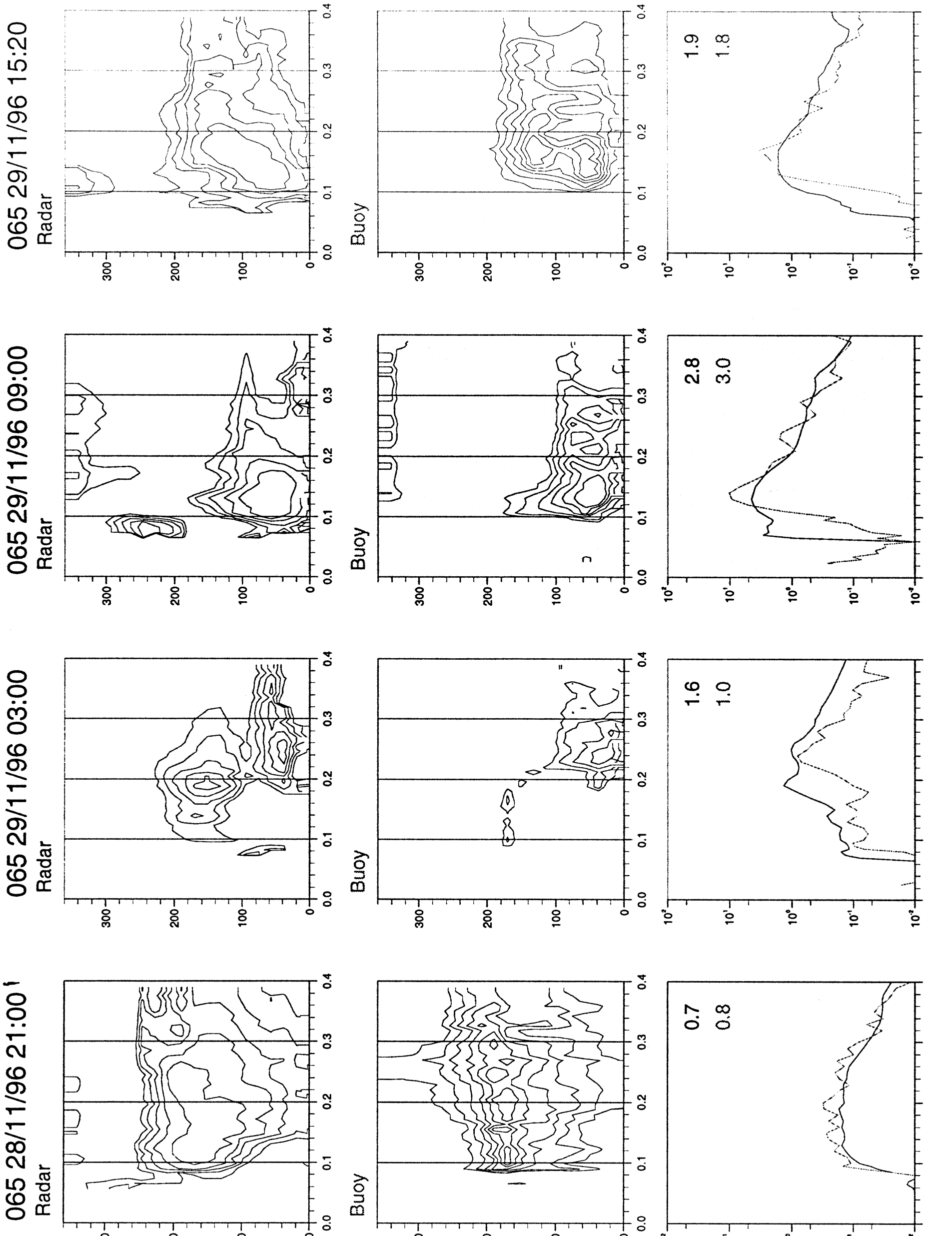
Nevertheless, direct comparisons are interesting. Remote sensing instruments like the HF radar or the satellite SAR tend to measure the wavenumber spectrum from which the directional spectrum and the wave parameters are determined by suitable transformations and integrations. The directional wave buoy data provide only the frequency spectrum and the first four frequency-dependent Fourier coefficients. The directional distribution has then to be constructed using, e.g., the ME mentioned above. The ME estimates can have a tendency to split peaks (Lygre and Krogstad, 1986), but in real situations, the lack of better measurements of the directional distribution makes it difficult to distinguish between real bimodal structures and artefacts of the estimate. Lygre and Krogstad compare the ME estimate with a ML method (Capon, 1969) and show that the ML method tends to produce broader distributions with poorer directional resolution. A similar comment on the ML method is given in (Donelan et al., 1996) where it is compared with a wavelet analysis method. The wavelet analysis method gives directional distributions with fairly narrow spreads even at quite high frequencies and wavenumbers.

The inversion process that generates wave measurements from HF radar Doppler spectra provides the directional wavenumber spectrum on a non-uniform grid. For wavebuoy comparisons, and indeed for all parameter extraction procedures, the data on the non-uniform grid are first averaged into wavenumber-direction bins. Alternatively, the spectra are first converted to directional frequency spectra using the shallow water dispersion relationship (with the same depth that is used during the inversion) and then binned into frequency bins. One method that can be used to assess the use of the maximum entropy method for directional spectra estimation is to determine the Fourier coefficients of the radar-measured spectrum and compare this measured spectrum with the ME spectrum obtained using the coefficients. Four examples are presented in Fig. 9. They show WERA measurements from Petten and have been selected because they show some directional bimodality as the sea responds to changing wind directions. Where this bimodality is very clear, e.g., at 0300 on 29/11/96 at ~ 0.2 Hz, the maximum entropy spectrum is a better approximation than the cosine model which is simply a cos-2s distribution fitted to the first pair of Fourier coefficients. There is some evidence of peak splitting (e.g., 1520 on 29/11/96) and increased amplitude at the peak (e.g., at 0300 on 29/11/96) but over most of the spectrum, the maximum entropy method has a similar shape to the radar measurement. This qualitative judgement suggests that peakedness in wavebuoy spectral estimates can be ignored but bimodality and other features of the shape are real and should be present in both wavebuoy and radar measurements.

Fig. 10 shows the spectra in Fig. 9 as compared with the buoy maximum entropy estimates at that time. The levels and normalisation are as before. Relative amplitudes between columns can now be seen in the lower panel which compares the frequency spectra (also with amplitude on a log scale). Note that the sampling variability of all these estimates is quite high. There are many differences, in particular the radar measurements show occasional discrete modes in the spectra that are not present in the buoy measurements at all (e.g., at 0.8 Hz, 250° at 0900, 29/11/96). These are the source of differences in mean direction estimates at low frequencies (Wyatt et al., 1998). Directional spectra comparisons have proved very important in revealing the origin of errors in HF radar-integrated parameter estimates. There is broad agreement in the shape of the main contributions to the spectrum with some evidence of bimodality at high (0900) as well as low (1520 in the buoy data) frequencies. These features can be seen

Fig. 9. Four directional spectra measured by the WERA HF radar at Petten (upper panel, time and date shown) compared with the maximum entropy spectra determined from the Fourier coefficients of the measurement (middle panel) and the spectrum estimated using a cos-2s-model with parameters determined from the Fourier coefficients (lower panel). The amplitude in each column is normalised by the maximum in the radar measurement which varies from column to column. Logarithmic amplitude levels are shown. There is sometimes some distortion at high and low frequencies due to limitations of the contouring at the boundaries. The figures above each picture indicate the maximum spectral amplitude and the frequency of the spectral maximum.


065 29/11/96 15:20
Radar 1.0 0.14
Radar (mes) 1.5 0.18
Radar (cos) 1.0 0.15
065 29/11/96 09:00
Radar 3.2 0.12
Radar (mes) 6.9 0.08
Radar (cos) 2.8 0.15
065 29/11/96 03:00
Radar 1.9 0.25
Radar (mes) 4.0 0.25
Radar (cos) 0.8 0.20
065 28/11/96 21:00
Radar 0.1 0.13
Radar (mes) 0.1 0.23
Radar (cos) 0.1 0.14


065 28/11/96 21:00
Radar
Buoy
0.7
0.8
065 29/11/96 03:00
Radar
Buoy
1.6
1.0
065 29/11/96 09:00
Radar
Buoy
2.8
3.0
065 29/11/96 15:20
Radar
Buoy
1.9
1.8

more clearly in an animation of 1 month's data that can be viewed via the SCAWVEX web site (http://www.sheffield.ac.uk/uni/academic/D-H/eoc/scawvex/home.-html). Note that the directional spectra comparisons confirm the view expressed earlier that a transfer function based on the mean frequency spectrum or spectral ratio as discussed in Section 3.2 is not necessarily appropriate for HF radar since differences are due to sporadic effects introducing spurious modes into the directional spectra rather than consistent biases in the measurement.

5. Conclusions

In this paper, we have discussed a number of different methods for the intercomparison of wave data. The most obvious conclusion is that the quality of a given measurement system depends on a range of factors that cannot be uniquely identified by the use of a single statistical approach.

Deviation between instruments may be due to general systematic offsets (e.g., calibration errors), temporal and spatial offsets, or merely caused by the sampling variabilities. The sampling variability of the main wave parameters is easily derived for conventional time series measurements, but is considerably harder to obtain for the HF radar and other remote sensing instruments. Computer simulation is very important in this respect.

The most central sea-state parameters are the significant wave height, the mean period, the spectral peak period, and the mean wave direction and the directional spreading of the waves.

The significant wave height is an integral of the spectrum and weights all spectral components equally. High and low frequency bias in the spectrum will therefore show up as a bias in $Hm0$ for low and high sea-states, respectively, assuming that these biases are not sea-state dependent. We have seen that for, e.g., HF radar, the high frequency bias is larger in high sea-states and is responsible for a larger $Hm0$ bias. For time series measurements, the sampling variability of $Hm0$ increases more than proportional to the value itself. Assessing errors in one measurement by considering mean differences is therefore not appropriate.

The mean wave period, $Tm01$, does not vary over a large range and has a rather low sampling variability for time series data. The parameter is affected by any high frequency bias of the spectrum and reveals the positive bias in the radar-measured spectrum at the higher sea-states. Similar conclusions also apply to the more commonly used spectral mean wave period, $Tm02$ although the high frequency dependency is larger.

The peak period, Tp, varies over a larger range than $Tm01$ and $Tm02$ and has larger variability. One obvious reason is the case of multi-modal spectra where there is a

Fig. 10. Directional spectra in Fig. 9 compared with the maximum entropy spectra determined from the Fourier coefficients of the wave buoy measurement (middle panel). Scaling similar to Fig. 9. The lower panel shows the two frequency spectra, radar solid line and wavebuoy dashed line with their corresponding significant waveheights (radar above wavebuoy).

chance that the observed maximum irregularly switches between two possible peak frequencies. As mentioned above, computer simulations have revealed that the sampling standard deviation is typically 4–10 times larger than for $Tm02$ even for single mode spectra. The sampling variability for HF radar measurements of Tp has not yet been determined and so it was not included in the ML analysis.

The mean direction is a function of frequency, and one often considers the directions for high and low frequency waves (wind sea and swell) separately. Picking the direction at the spectral peak is a popular choice although the randomness in Tp then also affects the direction. Spectrally averaged directions like the main wave direction avoids this problem, but being an integrated quantity, it is also less sensitive to particular features in the spectra. In the case of the HF radar, the main wave direction can be distorted by the presence of additional modes in the directional spectrum and in bimodal spectra, in general, it is not a particularly useful parameter since it measures the average direction rather then the particular directions of the component modes. The situation for the directional spread is similar to that for the mean direction, but the sampling variability is larger. Note that spread has only been assessed in a qualitative manner here through the use of plotted directional spectra.

Of the regression method considered above, the ML regression appears to be the preferred choice for a quantitative intercomparison between the two instruments when the sampling variabilities are known. Also the simple non-parametric regression is a useful tool for inspection of scalar data and could be used to indicate a need for a piecewise linear or a spline relationship that could also be determined using ML.

Deviations in sea-state parameters should, if possible, be traced back to deviations in the underlying spectra. The frequency spectrum is most easily investigated by the average spectral ratio, but computing the ratio should be limited to frequencies above the spectral peak. Frequency-dependent directional parameters like the mean wave direction and the directional spread may be studied similarly to the frequency spectrum, but full directional spectra are more difficult to compare. Note however, as illustrated for the HF radar, that it may nevertheless be necessary to assess the full directional spectrum to identify bimodality and for spurious contributions. The HF radar comparisons do, however, provide some independent support for the use of the maximum entropy technique for the analysis of buoy data

Acknowledgements

The data set consisting of the Topex-Poseidon altimeter data (provided by AVISO, France) and NOAA buoy data (obtained from the National Oceanographic Data Centre, USA) was compiled by Dr. Stephen Barstow, OCEANOR ASA. Markus Sova and Ralph Burton contributed to the HF radar statistical analysis. Louise Ledgard collected the Holderness radar data and the SCAWVEX group at the University of Hamburg the Petten radar data (Drs. Sova, Burton and Ledgard were all at the University of Sheffield at the time). The wavebuoy data at Holderness was provided by the Proudman Oceanographic Laboratory and at Petten by the Rijkswaterstaat.

References

Abramowitz, Stegun, 1965. Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables. Dover, New York.

Allender, J., Audunson, T., Barstow, S.F., Krogstad, H.E., Bjerken, S., Steinbakke, P., Vartdal, L., Borgman, L.E., Graham, C., 1989. The WADIC project: a comprehensive field evaluation of directional wave instrumentation. Ocean Engineering 16, 505–536.

Barstow, S.F., Krogstad, H.E., Torsethaugen, K., Audunson, T., 1985. Procedures and problems associated with the calibration of wave sensors. Adv. Underw. Tech. Offs. Eng. 4, 55–82.

Boggs, P.T., Rogers, J.E., 1990. Orthogonal distance regression. Contemp. Math. Am. Math. Soc 112, 183–194.

Brillinger, 1975. Time Series, Data Analysis and Theory.

Capon, J., 1969. Maximum likelihood spectral estimation, In: Heykin, S. (Ed.), Nonlinear Methods in Spectral Analysis, Topics in Applied Physics, Vol. 34. Springer, Berlin, pp. 155–179.

Donelan, M.A., Drennan, N.M., Magnussen, A.K., 1996. Nonstationary analysis of the directional properties of propagating waves. J. Phys. Oceanogr. 26, 901–914.

Fan, J., Truong, Y.K., 1992. Nonparametric regression with errors in variables. Ann. Stat. 21, 1900–1925.

Guillaume, A., 1990. Statistical tests for the comparison of surface wave spectra with applications to model validations. J. Atmos. Oceanogr. Tech. 7, 557–567.

Hastie, T., Stuetzle, W., 1989. Principal Curves. J. Am. Stat. Ass. 84, 502–516.

Krogstad, H.E., 1982. On the covariance of the periodogram. J. Time Series Anal. 3, 195–207.

Krogstad, H.E., 1989. Simulation of multivariate Gaussian time series. Commun. Stat. 18 (3), 929–941.

Krogstad, H.E., 1991. Reliability and Resolution of Directional Wave Spectra From, Heave, Pitch and Roll Data Buoys, In: Beal, R.C. (Ed.), Directional Wave Spectra. Johns Hopkins University Press.

Long, R.B., 1980. The statistical evaluation of directional spectrum estimates derived from pitch/roll buoy data and 10, 1980 buoy data. J. Phys. Oceanogr. 10, 944–952.

Lygre, A., Krogstad, H.E., 1986. Maximum entropy estimation of the directional distribution in ocean wave spectra. J. Phys. Oceanogr. 16, 2052–2060.

Munthe-Kaas, H., Krogstad, H.E., 1985. Sampling variability of sea-state parameters. ANODA Report No. 09, SINTEF.

Sova, M.S., 1995. The Sampling Variability and the Validation of High Frequency Wave Measurements of the Sea Surface. PhD Thesis, Univ. Sheffield.

Tucker, M.J., 1991. Waves in ocean engineering, measurement, analysis and interpretation. Ellis Horwood Series in Marine Science. Ellis Horwood.

Tucker, M.J., 1993. Recommended standard for wave data sampling and near real-time processing. Ocean Engineering 20, 459–474.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res. 90 (C5), 8995–9005.

Wolf, J., 1997. The analysis of bottom pressure and current data for waves. Proceedings of 7th International Conference on Electronic Engineering in Oceanography, Southampton, Publ. 439, IEE, June 1997, pp. 165–169.

Wyatt, L.R., 1990. A relaxation method for integral inversion applied to HF radar measurement of the ocean wave directional spectrum. Int. J. Remote Sens. 11, 1481–1494.

Wyatt, L.R., Thompson, S.P., Burton, R.R., 1998. Evaluation of HF radar wave measurements. Coastal Engineering, this issue.

Zambresky, L., 1989. A verification study of the global WAM model, December 1987–November 1988. GKSS Forschungszentrum Geesthacht, Technical Report No. 63, GKSS 89/E/37, 86 pp.