An Assessment of Wave Observations from Ships in Southern Oceans

A. K. LAING

New Zealand Meteorological Service

(Manuscript received 3 May 1984, in final form 19 November 1984)

ABSTRACT

Observations of wind waves and swell from ship reports are investigated. Comparisons are made between estimates of wave parameters made from ships in southern oceans by calculating correlations as a function of ship separation, and analyzing the resulting series. It is shown that there is very little consistency in the reporting of wind wave and swell periods and swell directions. Heights fare considerably better, although it is shown that one observer still fails to account for at least 50% of the variance in the observations of another observer. Further, some comparisons of ship-reported wave heights with measurements show a high bias in the observations of at least 0.5 meters. Some weak quality control criteria fail to be met by a large number of observations. Despite the inconsistencies, the intercomparisons show that the data is representative of many of the physical characteristics of wavefields, and therefore can be useful in climatological studies.

1. Introduction

The monitoring of waves on the ocean surface is an important aspect of marine science and is essential for maintaining oceanographic and marine meteorological services. Wave data is often scarce, especially in the South Pacific. Some measurements exist, but these are confined to coastal sites where very specific engineering or operational requirements are involved. It is difficult to use this coastal wave data to extrapolate along a coast: shallow water refraction, diffraction, and coastal wind inhomogeneities locally distort the incident wavefield. In open ocean areas measurements are rare, and unfortunately, much of the generation of waves that affect our coasts occurs in these areas. The only other present data source, and the one on which we must rely, is the set of visual reports of wave parameters made from casual shipping (ships of opportunity).

Descriptions of waves and winds, by observation or measurement, are the input a forecaster requires for analyzing the state of the sea, producing forecasts and assessing these forecasts. Hindcasts and forecasts of the sea state can be made, and usually are, based on wind data alone. However, without "sea truth" these calculations can never be verified. This is particularly pertinent to the tuning and verification of numerical wave models.

Another important aspect of wave studies is the compilation of wave climatologies. In coastal areas this information can often be derived from a mix of measurements and visual observations (e.g., Pekerill and Mitchell, 1979, for the New Zealand coastline). However, in open ocean areas there is generally a paucity of suitable data. In such cases, ships' observations of waves are often used; in raw form (Hogben and Lumb, 1967; Reid and Collen, 1983; U.S. Navy, 1979), or calibrated against available measurements (Andrews et al., 1983). In some cases the wave climate has been estimated from numerical hindcasts (Hydraulics Research Station, 1977; U.S. Navy, 1983).

In this paper a data set of visual wave observations will be assessed with a view to its usefulness as a verification base for wave modeling, in wave analysis, and wave climatologies.

A measure of consistency of the full set of visual wave observations from ships can be obtained by comparing reports from ships sufficiently close together to be experiencing the same sea states. In open ocean, away from coastal sheltering of the wind and shallow water effects, the horizontal scale of major variations in the sea state is of the order of several hundred kilometers. For this study, coastal waters are defined as those within the marked areas around land as shown in Fig. 1, and have been chosen to minimize the above effects without necessarily eliminating fetch effects and local small-scale distortions of the wave field. Oceanic currents and small-scale atmospheric activity also affect the sea, although the lag in response of the sea to atmospheric forcing filters out the smallest variations from the latter. Therefore, we can not expect ships to experience exactly the same wave conditions unless they are very close together (within a few hundred meters). Unfortunately, there is very little quantitative information on the spatial variation of wave parameters, although recently remote sensing techniques have enabled some consideration of the problem (Challenor, 1983).

Observations should give a more smoothed view of the sea state. It is usually assumed that statistically

© 1985 American Meteorological Society
homogeneous turbulence in the wind over an infinite ocean gives rise to a statistically stationary and homogeneous wave field. If appropriate statistics of the wavefield are under observation, then similar observations would be expected from observers close together. Even if this assumption fails (for small fetches), good observing practice should give a picture of a large enough area of sea to smooth out inhomogeneities. Under good practice, observations will be made over an interval of several minutes. The propagative nature of waves then ensures spatial representativeness over distances of the order of a kilometer or more.

A previous study (see Laing, 1982) found a high degree of discrepancy between observations from neighboring ships. Presumably there are two main contributions to this error: the effect of a real error in observing the sea state, and the effect of a real difference in sea state, the latter resulting from upstream wind variations over distances similar to or less than those between the ships. The problem is to identify the respective contributions.

2. Data and procedure

The source of data used for this study is a group of vessels which have been recruited by New Zealand as "voluntary observing ships." Observations from these vessels are passed on to the New Zealand Meteorological Service and entered into the climatological archive. Transmission errors are avoided by using the archive derived from the ships' logs rather than from the reports transmitted to the shore.

The data set used here spans the period from 1968 to 1981 inclusive. Only vessels within the region
bounded by 10°N to the north, 40°E to the west, and 130°W to the east were included: in total 205 203 reports. Vessels in coastal waters (see Fig. 1) were for most purposes excluded, since the variability of the sea state was expected to be greater due to changes in roughness length at the coast, the orography and possible bathymetric influences. This reduced set has 153 138 reports. All reported wave parameters were considered.

Since the exercise is to compare the reported wave parameters, it is necessary to define acceptance and rejection criteria for these reports. These wave observations form part of the standard meteorological ship reports. A summary of the code groups used for describing waves is given in Appendix A. The form of this code changes from the time to time, but was consistent during the period of this study (1968–81). Changes did occur on 1 January 1968 and again in 1982, when there may have been some confusion due to the conversion, but any such instances are unlikely to have seriously contaminated results taken from a total of 13 years of data.

For a given reporting parameter, an observation is rejected only if it does not stand alone as a sensible value. Later in this paper some quality controls, which take into account dynamic relationships between wave parameters, are applied. For now, wave heights are limited to less than 20 meters, directions to 10–360 degrees (in units of 10 degrees), and periods to less than 20 seconds. It is possible that errors in reporting location occasionally occur, but short of tracking every vessel from reporting time to reporting time there is no way of automatically detecting this.

According to the observation manuals (WMO, 1976, 1983) wave heights are of the larger, well-formed waves within a system (i.e., wind sea or swell) being observed, and the statistic “significant wave height” (H1/3), which is the average of the highest one-third of waves, was originally defined to resemble observed wave heights. However, it has been noted that there is some bias in reported heights when compared to significant wave heights derived from measurements (WMO, 1976). A further unobserved parameter, the combined wave height, has been included in this study. This is defined as the square root of the sum of the squares of the heights of the wind waves and all swell trains.

The basis of the method is in intercomparing ships’ observations by estimating the correlation statistic for a given parameter and determining how it varies with ship separation. First, for a given reporting time, the data was searched for all pairs of vessels whose separation was within a particular range. Initially these ranges were in 74 km (40 nautical mile) bands: 0–74 km, 74–148 km, etc. These bands will be referred to as separation classes. For a nominated parameter a pair was rejected if both reports did not include that parameter. These observation pairs were randomly reversed to ensure that there was no chance of climatological variations with latitude influencing the results, and it was noted that this indeed gave an unbiased sample.

The correlation coefficient is the most suitable statistic for comparing the elements within the observation pairs. To consider it in any depth it is preferable to have the observational parameters independently sampled from normally distributed populations. A better approximation to normality is obtained if the pairs of reported values (R1, R2) are replaced by their logarithmic transforms (lnR1, lnR2) (see Kendall and Stuart, 1963, Chap. 37.5).

As a first approximation, it is reasonable to expect the observations of wave parameters to behave as a first-order autoregressive, or “red-noise,” process. Obviously the coherence of the atmospheric conditions which generate wind waves decreases with distance, at least on subsynoptic scales, and we would naturally expect wind waves to behave similarly. There may be some lower bound to the separation range over which this coherence increases, created by the effect of observing error. Also, at synoptic scales we might also expect some positive response matching the average scale of separation of wave systems.

An exponential model was fitted to the correlation estimates. Assuming the form ρ(ρ) = ρ0 exp(–κx), where x is the distance of separation in kilometers, a least-squares fit was made so as to minimize the error

\[ d^2 = \sum \{\ln(\hat{\rho}(x_i)) + kx_i - \ln(\rho_0)\}^2. \]

The value of x used for each separation class is not quite the midrange value for the class, but is calculated by assuming (ideally) an isotropic distribution of vessels in a two-dimensional space. For each vessel the number of pairs within a radius r would be expected to increase as r². Given a separation class between m and n kilometers, the number of pairs for all vessels would be proportional to m² – n², and the mean separation x_i = [(m² + n²)/2]¹/₂.

To test for significant departures from this model we need a distribution for each correlation estimate \( \hat{\rho} \) (in the following a caret superscript will indicate an estimate), and from this we can determine whether the estimates are significant deviations from the model. The standard practice in correlation estimation is to use the Fisher variance stabilizing transformation defining z by z = 0.5 ln[(1 – \( \hat{\rho} \))/(1 + \( \hat{\rho} \))] (see Kendall and Stuart, 1963, Chap. 16.3). If n observations contribute to the estimate \( \hat{\rho} \), the distribution of z is close to normal with variance 1/(n – 3). A confidence range can then be estimated for the spatial constant 1/k and the intercept value \( \rho_0 \), using the distributions of the coefficients of the regression \( y = -kx + b + \epsilon \), where \( y = \ln(\rho) \), \( b = \ln(\rho_0) \), and \( \epsilon \) is distributed normally (see Zehna, 1970, Chap. 12; Kendall and Stuart, 1963).
3. Combined wave heights

a. Visual ship observations versus visual ship observations

The correlation estimates $\hat{\rho}(x)$ for the transformed combined wave height pairs in each separation class are displayed in Fig. 2. Also shown are the number of pairs used in each calculation and the 95% confidence limits on $\hat{\rho}$ (dotted lines). There is obviously an increase in correlation as the separation decreases, rising to a value of 0.54 for the class with least separation (0–74 km). The resulting exponential fit has $\rho_0 = 0.61$ and $k = 0.00119 \text{ km}^{-1}$. This curve is also plotted in Fig. 2, and it can be seen that the fit is acceptable at all points.

These results are consistent with a previous study (Laing, 1982) in which a sample of data (February 1981–April 1982) received in real time via the telecommunications network was used. For all pairs of ships within 111 km of one another the combined wave heights gave a correlation coefficient of 0.51. The root-mean-square error from the best linear fit to the pairs was quite high at about 1.0 m. From this it was suggested that a maximum of only about 30% of the variance in a ship's observations would be explained by a neighboring ship's observations.

Real physical differences may contribute to the low correlations in the present study. With the choice of a given separation range, most of the effects on sea state of atmospheric inhomogeneities and oceanic currents at scales greater than this range will have been filtered. Subdivision of this range should refine the filter. The closest separation class was therefore repartitioned into groups with ship separations in the ranges 0–18 km, 18–37 km, etc. The new results showed only a slight improvement, with correlations up to about 0.60, and no significant changes to the model. As the width of the separation classes is reduced the numbers of pairs of observations in each class decreases, and the confidence in the estimate of correlation decreases.

We have suggested that some coherence might be expected on a synoptic scale typical of wind-wave systems. This is not evident for combined waves. However, the picture is complicated somewhat by the contribution of swell to a combined wave. In fact, swell mixes the wave energy from wind-wave systems as it propagates from one active fetch into other active fetches, and this mixing will reduce any coherence in the combined waves at synoptic scales. The effect is to improve the conformity of combined wave heights to a first-order autoregressive model.

The spatial constant, $1/k$, is the e-folding distance for the function, and is a useful measure of the scale of event represented by the data. For the combined wave height this is about 840 km, but since the combined wave is an artifice, this scale has no physical meaning. The 95% confidence ranges for $\rho_0$ and $1/k$ were calculated as 0.59–0.63 and 805–885 km respectively, where 12 values were used in the estimates.

b. Visual ship observations versus measurements

To get a feel for the bias and scatter of ship reports of waves, where possible the observations have been compared with measurements. The observations and measurements are required to be independent of one another. Since measurements are limited to a few sites, and the incidence of reporting vessels at these sites is even more limited, few comparisons are possible. It should be noted that the observations used in these comparisons are probably of better quality than those normally received from shipping: the presence of the vessel usually indicates an interest in the operation involving the collection of data. It should also be noted that in this section the spatial variation of wave parameters is relatively unimportant, because for the most part the observations and visual measurements are coincident. The effects of proximity to the coast will only affect the comparisons indirectly through the overall climatological difference between coast and open ocean and the marginal distributions of observational error given the wave climate. Section 9 of this paper does include some intercomparisons between visual observations made within the coastal zones.

1) CASE 1—MAUI-A SITE: SOUTH TARANAKI BIGHT

For a period of twenty months from 1 September 1980 to 18 May 1982, a record was made of the
significant wave heights derived from measurements made by the Waverider buoy owned by Shell BP and Todd Oil Services Ltd. and located in the vicinity (6.5 km north until April 1981 then 1.5 km west) of the production platform at the Maui-A site (173.4°E, 39.6°S—see Fig. 1). Also, combined wave heights were calculated from the visual observations of wave and swell heights made by the vessel servicing the production platform. Only observations during which the vessel was within 33 km of the platform were used. Usually this distance was less than 1 km. Obviously misreported data were ignored (e.g., coding errors giving wave heights above 20 meters).

For this period, 2138 data pairs were useable and are displayed in a scatter diagram (Fig. 3). Calculations revealed a root-mean-square difference between the values of 1.30 meters, and a mean difference (visually measured) of 0.99 meters. Nevertheless, the correlation of 0.83 suggests a reasonably linear relationship, and a linear regression of the ship reports on the measured values gives the line

\[ H_s = 1.17H_b + 0.60, \]

where \( H_s \) is the combined wave height derived from the ship report and \( H_b \) is the measured significant wave height. The residual variance is 0.77 m². Monthly analyses of the data showed no great variation in these statistics over the period studied. The distributions of the coefficients are approximately normal with standard errors of 0.07 and 0.03 respectively. The most notable feature of this regression is the exaggeration of the ship reports by about 20% on top of a half-meter high bias. If the measurements are regarded as having error, then neither the measurements nor the observations can be regarded as the independent variable. In this case a best linear fit to the data is more appropriate, but it was found that although the actual line was different \( (H_s = 1.55H_b + 0.26) \) the residual error was insignificantly lower than for the \( H_s \) on \( H_b \) regression used above. Whichever way the measurements are regarded, the exaggeration by the visual observers is apparent.

2) CASE 2—SEDCO 445 DRILLING SHIP: WEST COAST NEW ZEALAND

For the 1981 operating period of the SEDCO 445 oil exploration vessel (operated by Shell BP and Todd Oil Services Ltd.), a comparison was made between significant wave heights from a Waverider buoy deployed near the vessel and combined wave heights derived from visual wave reports made from the vessel itself. During the period the vessel and the buoy were at four different locations, viz: “TANGAROA” (38.2°S, 173.3°E), “KIWA” (39.8°S, 172.7°E), “MIKONUI” (42.7°S, 170.1°E), and “WAINUI” (38.5°E, 173.3°W) (see Fig. 1). From these locations 1787 data pairs were collected and plotted in a scatter diagram (Fig. 4). The data received the same processing as the data in the previous case. A correlation coefficient of 0.92 showed a good linear relationship, namely:

\[ H_s = 1.05H_b + 0.42, \]

where \( H_s \) and \( H_b \) are as in case 1. The residual variance is 0.37 m². As with case 1 there is about a half-meter high bias in the reported combined wave
height. In contrast, there is appreciably less scatter and a slope very close to unity. However, there is a suspicion that in this case the visual reports were not always independent of the measurements, as the Waverider was monitored on board the observing vessel in contrast to the arrangement in case I.

3) CASE 3—WEATHER SHIPS: NORTH ATLANTIC

Jardine (1979) conducted a similar comparison in the North Atlantic using visual data from the weather ship “I” and measurements from the shipborne wave recorder. A total of 3901 data pairs were used. The best straight line fit to the data, weighted inversely by variance, was given as:

$$H_s = 0.98H_b + 0.5,$$

with $H_s$ and $H_b$ as before. It is notable that the linear relationship implies a bias of a half-meter but little other systematic departure from the line $H_s = H_b$. The residual variance was still appreciable, being given as 0.85 m$^2$ at 1.0 meter measured height, 1.05 m$^2$ at 2.0 meters, and 1.35 m$^2$ at 3.0 meters.

Hogben and Lumb (1967) compared observations and measurements made at weather ships “ALPHA” (62°N, 33°W), “INDIA” (59°N, 19°W), “JULIETTE” (52.5°N, 20°W), and “KILO” (45°N, 16°W) from 1955–61. Only reports in which one wave group (i.e., wind sea or swell) was reported are used. For 317 observations a correlation of 0.86 was calculated, and the linear regression given as

$$H_s = 0.88H_b + 1.23,$$

with a root-mean-square error of 1.27 meters.

The high residual variance in the above cases illustrates the considerable scatter in visual observations. Jardine (1979) has suggested that inaccuracies in the measuring devices may contribute significantly to the discrepancies at low wave height. However, in the New Zealand cases the measurements were made using Waverider buoys which have different response characteristics to the shipborne recorder. Also, the errors are present at all wave heights, and reference to Figs. 3 and 4 show that the scatter actually increases with wave height, indicating at least as much discrepancy at this end of the range.

A possible source of error is confusion in reporting an observed sea state as wind sea, swell, or both. A tendency in confused seas to allocate the same energy to more than one wave feature would give a high bias. Further, there may be a tendency to report only the very prominent waves within a feature, thus biasing the combined wave heights towards a height statistic more comparable with, say, $H_{1/3}$ (average of the highest one-sixth of waves). It is possible to separate the contributions from these possible causes by noting the fact that the latter contribution should not affect the correlation coefficient, as the various fractional height statistics akin to $H_{1/3}$ have constant ratios (Longuet-Higgins, 1952). However, by selecting the maximum wave feature (wind sea or swell) from the wave observation, rather than the combined wave height for comparison with the measured significant wave height, we may be able to get some indication of the “cross observation,” or contamination of observations of one wave feature by other features. In fact, this exercise proved inconclusive, increasing the correlation coefficient for case 1 by only 0.01, a figure which is significant only at the 85% level. Since at least one wave train has been removed from the observation, there is a substantial reduction in the overall bias.

4. Wind-wave heights

The reported wind-wave heights were treated in the same manner as the combined waves. The correlations, together with the exponential curve $\rho(x) = 0.58 \exp(-0.00160x)$ and the 95% confidence limits on the correlation estimates, were found (see Fig. 5). The intercept at 0.58 (with a 95% confidence range of 0.51–0.66) is much the same as for the combined waves, while the curvature is greater. That is, the characteristic scale at 625 km (with a 95% confidence range of 550–730 km) is appreciably less than for combined waves. This ties in with what we would expect: the scale of combined wave features being increased by the contribution from swell.

The exponential fit is not particularly good at small distances, degrading the fit over much of the range. Inspection of Fig. 5 shows that a better fit would be obtained if the first two values were ignored, and omitting these values the curve is $\rho(x) = 0.68$
\[ \times \exp(-0.00182x), \] with a characteristic scale of 550 km. It should be noted that an ideal box-shaped fetch has a sawtooth wind-wave height distribution, and this complicates attempts to interpret the e-folding distances in terms of fetch lengths, although real fetches will have considerably smoother wind and wave height distributions.

Since there is some departure from a first-order autoregressive model, and the shape of the correlogram in Fig. 5 suggests the flattening characteristic of a second-order autoregressive process (see Jenkins and Watts, 1968, Chap. 5.2), an attempt was made to define such a process to fit the correlations. An oscillating form for \( \rho(x) \), rather than a purely exponential form, minimizes the residual error; i.e.,

\[ \rho(x) = \rho_0 \left( \frac{[e^{-r \cos \theta} \sin(rx \sin \theta + \theta)]}{\sin \theta} \right). \]

The results gave \( \rho_0 = 0.46 \), \( r = 0.00242 \ km^{-1} \), and \( \theta = 0.61472^\circ \), with a residual error \( d^2 = 0.0015 \) (reduced from 0.0128). As indicated by the lower residual error, the fit given by this process is much better than that of the first-order process. The period of oscillation is 4200 km, which may represent an average separation between synoptic events. The parameter \( 1/r = 410 \ km \) represents a relative measure of the extent of each fetch. Note that whilst more than one fetch is present within a synoptic event they are often adjoining, and indistinguishable using the wind-wave heights alone.

Alternatively, from Fig. 5 it can be seen that the curve is drawn to the bottom of the 95% significance range by some very low correlations for the first two separation classes. This also means that the intercept value of 0.58 is well above that implied by the raw data. In fact there appears to be a plateau in the correlations at about 0.45. Subdivision of the first separation class into 18 km classes confirms this, with values of about 0.45 except for the very closest separation class. For this class there is a marked improvement with a correlation of 0.60. Although the number of observations contributing to this calculation is relatively small (145), and the 95% confidence limits (0.51–0.71) wide, the levels are high enough to indicate a significant trend. An even finer resolution was attempted with bandwidths of 9 km for the first 74 km, and 18 km for the next 90 km. The results appear in the inset of Fig. 5. The plateau in the correlations is evident for separations down to about 60 km, and closer than this there is the noticeable increase in the correlation.

A suggested explanation is that the variability of the wind over small distances (100–200 km) within a roughly uniform fetch has enough effect locally on the waves to limit the coherence of reports. When the separation of vessels becomes small enough (less than 60 km), the scale of variation is approached and there is the improvement noted. Thus, there appears to be a more positive response here to the effects of gustiness, mesoscale or frontal activity than for combined waves, which are partly insulated from such effects by the contribution of swell (which is expected to have characteristics changing little over the smaller distances). The sharpest of these variations in winds are probably those effective when a front lies between two observing vessels. The relative contribution of these events to the pattern in the correlations above could be tested by noting the wind directions and wind speed at all observing ships. However, the present study was not extended to assessing the accuracy of marine wind reports, and so this detail was not sought.

Unfortunately, the second-order approach appears inconsistent with the sharp increase in correlation at small separations. The differences are a matter of scale. On coarse scales (greater than 60 km) it may be an adequate description, with a different model being required for the finer scales.

5. Wind-wave periods

a. Visual ship observation versus visual ship observations

The analysis of the correlations of observed wind-wave periods is displayed in Fig. 6, with the fitted exponential curve \( \rho = 0.36 \exp(-0.0090x) \). It is obvious that all of the correlations, except those for the widest separation classes, are much lower than those for wave heights. The uniformly low correlations also reduce the curvature with the result that the characteristic scale is very long at 1100 km.

The curve falls within the 95% confidence limits of the correlation estimates for all values, but the

![Fig. 6. As in Fig. 2 but for wind-wave periods.](image_url)
plateau in correlation which characterized the wind-wave height observations is once again evident. Subdivision of the lowest separation class confirms this plateau at a correlation of about 0.3, with an increase in the correlation for the 0-18 km separation class to ~0.4. The interpretation given above for wind-wave height observations is consistent here. Excluding the first two values (0-74 and 74-148 km classes) and recalculating the exponential function, the characteristic scale was 1005 km. This is still greater than the scale for the wind-wave heights.

Fitting a second-order process to the correlation estimates gives $\rho_0 = 0.32$, $r = 0.00197$ km$^{-1}$ and $\theta = -0.00134^\circ$, with a reduction in residual error to 0.0010 from 0.0018 for the first-order process. In this case the period of oscillation is meaningless, but the improvement in fit is marked, indicating that although the second-order process is hard to interpret quantitatively, it is a better description of the data than the simple red-noise process.

b. Visual ship observations versus measurements

Hogben and Lumb (1967), in the study referred to in Section 3b, included a comparison between measured periods and visually estimated periods $T_v$. There were 294 pairs of values. The measured periods used were the average period $T_a$ (equal to the ratio [area under spectrum/second moment of spectrum about origin]$^{1/2}$) and the modal period $T_m$ (peak of spectrum). Their linear regressions were:

$$T_a = 0.32T_v + 4.7$$
$$T_m = 0.76T_v + 4.1$$

with root-mean-square errors of 1.20 and 2.23 s respectively. In both cases the correlation was 0.50. It is evident that the visual estimates bear more resemblance to the average period although there is little to enthuse about the accuracy.

6. Swell heights

Having considered the parameters which respond to local conditions, we now look at the characteristics of the observations of the distantly generated feature: swell. Where there is the possibility of multiple swell trains, the swell height is taken as the superposed value (calculated by taking the square root of the sum of squares of the reported swell heights).

Swell should be relatively unaffected by small-scale activity such as gustiness, and it is reasonable to expect that in regions where there is little wave generation or surface motion due to ocean currents, the observations of swell should be well correlated. However, there are extreme difficulties involved in accurately assessing a swell train in the presence of a substantial wind-wave field.

The correlation estimates with their 95% confidence limits and the fitted exponential curve $\rho(x) = 0.57 \times \exp(-0.0012x)$ are displayed in Fig. 7. The correlations are very similar to those for wind-wave height observations, indicating that swell observations are by no means more accurately made. The scale of swell variation is 870 km with a 95% confidence range of 780-980 km. This is substantially greater than the scale of wind-wave variation. For fine resolution, with band widths of 9 km for the first 74 km and 18 km for the next 90 km, the results appear in the inset of Fig. 7. The plateau in the correlations appears for separations down to about 50 km, and closer than this there is an increase in the correlation of the observations. However the estimates for the smallest separations have only about 30 contributing pairs (only half the number for wind-waves) and there is very low confidence in the estimates. The hypothesis put forward for this pattern in wind-wave observations is obviously not valid here for swell, but a possibility is that wind waves contaminate the swell observations, thus ending the swell with some characteristics of wind waves.

The correlogram of Fig. 7 appears to have a regular bumpiness. It is not characteristic of a second-order autoregressive process, and fitting a second-order process to the correlations only reduces the residual error of the first-order process by 30-40% (for the wind-wave model it was reduced by 80-90%). An alternative approach, given the regularity of the departures from the exponential, is to try a Fourier analysis. The Fourier transform of a correlation function gives a power spectrum, and the spectral estimates will give an idea of the distribution of power over wavenumber, indicating prevalent scales of variation (see Jenkins and Watts, 1968, Chap. 6).

FIG. 7. As in Fig. 5 but for swell heights.
Given the correlations $\rho(x)$, we have a series truncated at $x = 1111$ (km), with a lowest resolvable wavelength in the vicinity of 158 km. The effect is of a rectangular filter, limiting the wavenumber spectrum to wavenumbers in the range 0–7. It is relevant to also calculate the truncated Fourier series for the point values of the fitted exponential curve, i.e., for $\rho'(x) = 0.55 \exp(-0.00114x)$, $i = 1, 15$. This will be the discrete red-noise spectrum.

The spectrum is plotted in Fig. 8, along with the discrete red-noise spectrum from the fitted curve $\rho'(x)$ (the upper pair of lines). In this case, the spectral shape is typical of a red-noise (first-order autoregressive) process except for a slight peak at wavenumbers 4 and 5. This represents a slight periodicity on a scale of 220–280 km. Assuming 64 degrees of freedom (see Appendix B) the 95% confidence limits are 0.65 and 1.35 times our spectral estimates, and the 90% limits are factors of 0.71 and 1.29.

Applied to the spectrum in Fig. 8, there are no departures from the red-noise model significant at the 95% level, but at the 90% level the estimate at wavenumber 3 (wavelength of 370 km) is significantly different. Figure 8 suggests a characteristic red-noise spectrum with lower power than that calculated above, but with a significant departure at scales of 220–280 km. The apparent increased coherence in observations of swell height at multiples of this distance has no obvious cause but could be a response to the influence of frontal activity on existing swell fields or contamination of swell observations by wind-wave observations. The wavenumber spectrum was also calculated for the wind-wave height data but, in this range, showed no obvious scales of increased coherence.

7. Swell periods

The correlation estimates, their 95% confidence limits, and the fitted exponential curve, $\rho(x) = 0.11 \times \exp(-0.0015x)$, for the swell periods appear in Fig. 9. To overcome the problem raised with multiple swell trains, the reports of swell periods were used only if a single swell train was observed. It is obvious that there is very little correlation between observations. The characteristic (e-folding) spatial scale indicated by the fitted exponential curve is 660 km with 95% confidence range 390–2180 km. This is reasonably consistent with the scale for swell heights (880 km) although there is little confidence in the estimate. However, as with the swell heights, there is some evidence of periodicity in the correlation estimates and the curve lies outside the 95% confidence limits at two places.

A Fourier decomposition was once again performed to obtain a power spectrum. Figure 8 shows these results also (the lower pair of lines), and it is obvious that there is considerable departure from the fitted red-noise model (significant at the 95% confidence level using a chi-square test with 64 degrees of freedom). The two areas of departure are in the 370–550 and 180–220 km ranges. The higher wavenumber variation is possibly noise in a region where the spectral resolution is approaching its limits. The low wavenumber variation is on a scale more typical of wind waves and could be consistent with some confusion in the observation of swell and wind waves (i.e., the aforementioned "cross observation").

8. Swell directions

To compare swell directions there are several problems. First, we can only consider those reports in
which a single swell train is reported. Second, since no statistical correlation can be derived for cyclic quantities such as direction, the comparisons are limited to the relative frequency of differences between observations as a function of ship separation. For this purpose Fig. 10 has been constructed. Since the differences between observed directions include both positive and negative differences, the zero difference frequency has been weighted by a factor of 2 to give the correct distribution pattern.

It can be seen from Fig. 10 that even in the minimum separation class (0–74 km) there are 38% of all pairs of observations differing by 30 deg or more. This indicates considerable error, since within this separation class swell has little chance to vary its direction by any physical means. As the separation increases, the distribution of observed differences with separation eventually tends towards something like a uniform distribution (affected slightly by climatological factors introduced by the unbalanced north–south distribution of reports).

Fitting a surface of the form \( f = f_0 \exp(-Kx + Lb) \) we get \( f_0 = 0.293, K = 0.00065 \text{km}^{-1} \), and \( L = 0.0297 \text{deg}^{-1} \). The value of \( 1/K \) implies a scaling distance of 1540 km, and \( 1/L \) a characteristic observational spread of 34 deg on each side of the consensus angle.

9. Other features

In addition to the comparisons between each of the observed parameters, comparisons were made based on other criteria such as region, time of day and proximity to coast. A slightly different data set was used; namely, the reports from the years 1978–81 expanded by including data reported in real time from vessels recruited by other nations. Since the best observations appear to be those of height, attention was limited to observations of significant wave heights.

a. Regional differences

It is evident that the nature of weather systems will affect the eventual character of the wavefields generated. A mix of wave features will be involved. For example, we might expect that wind-wave scales in low latitudes are mostly influenced by the very long fetches associated with the southeast trades, and the occasional tropical storm. In midlatitudes we encounter mostly a mixture of northwest and southwest fetches, and at higher latitudes some longer westerly fetches. The comparisons so far have extended over a range of latitudes but the density of shipping in middle to lower latitudes is greater and must therefore weight the observations. Since wavefields have characteristics dependent on their location, the effect of locality on the coherence of reports was considered.

Three regions were designated: a southern oceanic region being those waters south of 45°S; a northern region between 30°S and 10°N; and a Tasman Sea region between 30 and 45°S, 150 and 175°E. All regions exclude coastal waters. Figure 1 indicates the divisions. The correlations for the first four separation classes are shown in Fig. 11. The 95% confidence limits were calculated and the differences were not significant at this level.

b. Night observations

The difficulty involved in making visual observations during hours of darkness is obvious. The influ-
ence of these observations on the overall set was assessed by looking at the set of observations of wind-wave heights made between the hours of 0900 and 1500 GMT within the longitudes 140°E and 175°W. These observations are made in the hours of darkness for the given longitudes except for a few of the 0900 GMT observations made near the east coast of Australia in summer. The correlations for the first four separation classes are shown as the solid line in Fig. 12. The 90% confidence limits for the estimates are also shown. The correlations for the complementary set of observations (which unfortunately include some night observations from winter months) are shown in Fig. 12 as the dotted line. For all classes the nighttime observations give estimates of correlation which are considerably lower and the differences are significant at the 90% level in three out of four classes.

c. Observations in coastal waters

In all of the above considerations reports from ships within coastal waters (roughly shown in Fig. 1) were ignored so that small-scale variations in the wavefield were minimized. The reports of wind-wave height from coastal waters will now be considered and compared to open ocean reports. Since the extent of what we designate as “coastal” is generally not greater than 150 km, we limit our attention to the first two separation classes. The results indicate that for the first class (0–74 km) there is little difference, but for the larger separation in the second class (74–158 km) the coastal observations are marginally worse, the correlation dropping from 0.44 to 0.41 (significant at 90% level).

![Figure 12](image)

**Fig. 12.** Correlations between observations of wind-wave height made at night and day. The error bars indicate 90% confidence limits.

<table>
<thead>
<tr>
<th>Category</th>
<th>Wind-wave period</th>
<th>Wind-wave height</th>
<th>Swell</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK</td>
<td>161 090</td>
<td>166 520</td>
<td>133 472</td>
</tr>
<tr>
<td>Inconsistent with other element</td>
<td>13 833</td>
<td>11 064</td>
<td>11 765</td>
</tr>
<tr>
<td>Doubtful</td>
<td>561</td>
<td>13</td>
<td>24 704</td>
</tr>
<tr>
<td>Erroneous</td>
<td>9</td>
<td>0</td>
<td>704</td>
</tr>
<tr>
<td>Missing</td>
<td>5 182</td>
<td>3 609</td>
<td>10 671</td>
</tr>
</tbody>
</table>

**Table 1.** Summary of quality classifications for reports of wind waves and swell. Total number of reports considered: 180 666.

d. Quality controls

A coarse quality control screen was applied. For data reported using the SHIP code, the World Meteorological Organization (1981) has set criteria so that data within a report can be marked, giving an indication of how much suspicion it should be viewed with. For observations of wind-wave parameters, their consistency with reported wind, for instance, is taken into account. Outer bounds on the parameter values are applied. Also, checks are made whether the reporting of one parameter is possible given that another is regarded as “unobervable.” There are five categories: OK, Inconsistent with another element, Doubtful, Erroneous and Missing. Table 1 summarizes the results of applying the quality controls.

Very few observations fall into the “erroneous” category since it is difficult to objectively claim that an observation is erroneous without reference to at least two other observational elements or to a set of wide observational bounds.

The data set was thus limited to those reports for wind-wave height made in the open ocean and which received an OK status; in this case 92% of the original number of reports. The correlations were greater than those for the whole set by an amount significant at the 95% level. For the first two separation classes the increase was from 0.47 to 0.52 and 0.44 to 0.50 respectively.

For swell reports the screening made very little difference. This is understandable, since there is no way of checking swell parameters against another element in the same way as wind waves can be checked against the wind.

10. Conclusions

The observations of wave parameters from ships have been studied primarily by investigating the consistency between reports as a function of separation. A reasonable correlation model was found for most observed wave parameters, and from this the various contributions to the differences between observations can be assessed.
The intercept values for the correlation coefficients from the first-order model for combined wave height, wind-wave height, swell height, wind-wave period and swell period were 0.61, 0.58, 0.57, 0.36 and 0.11, respectively, representing the correlations at zero separation. These imply that about 37, 34, 32, 15 and 1% respectively of the variance of one observation is explained by another nearby. This leaves 65–70% of the variance for wave heights, and 85–100% for periods, unexplained with the main contributing factor being "observing difficulties." Also, swell directions show discrepancies, with a 70 deg spread angle characterizing the variations.

The intercept values from the second-order model fitted to the wind-wave estimates were invariably lower than those from the first-order model, indicating the possibility of even larger unexplained variances. However, it must be added that for finer separation classes an increase in the correlations was noted for close vessels, particularly for wind-wave observations, suggesting that some of the unexplained variance may be due to small-scale variations or sharp changes in the wave fields. This is not consistent with a second-order model but it has been suggested that the applicability of the model depends on the scale of the event being observed.

The raw estimates of correlation, found by considering observations from ships within 9 km of one another, were 0.62 for wind-wave heights and 0.79 for swell heights, implying that 40 and 60% respectively of the variance is explained. Unfortunately, the small number of pairs of ships contributing to these calculations limits our confidence in them. Overall, taking a favorable view and taking into consideration the improvements noted when quality control measures were adopted, one observer will still fail to account for at least 50% of the variance in another observer's estimates of wave heights and 80% of the variance in the period. Unfortunately, this discrepancy must be attributed to observational error, and the observation of both wind-wave and swell periods must be viewed as unreliable.

Despite the errors, there is a reasonable information content in the reports (particularly of wave heights) usable for climatological purposes. The comparisons between observations show consistency with general physical characteristics of wavefields: correlations decrease with increasing separation between observing ships; the e-folding distances are consistent with values expected from the scales of generating events (e.g. swell heights correlate better at large distances than wind waves—as expected given the larger scale of swell events); greater variability is noticed near coasts; and observations made at night are slightly more suspect.

Attention should be paid to the high bias evident in the ship observations of wave heights. Comparisons between observations and measurements in both the Tasman Sea and the North Atlantic show a general overestimation of height. Also it should be noted that even the application of very simple quality controls can improve the consistency of the data.

For operational uses the data has qualitative value only. If it is available in large quantities for a given time, and the data density is high, then some quantitative value may be extracted by performing an analysis. However, care is required as the comparisons between observations and measurements show considerable scatters in the observations. Forecasters have long realized many of the problems with ship observations, and generally attempt to take this into consideration when using such data.

There is some evidence of confusion in the discrimination of wind waves and swell, with cross observation (i.e., the attribution of swell energy to the wind-wave field or vice versa) contaminating the observations. Hence, more notice should perhaps be taken of the somewhat artificial combined wave height, rather than the reported component wave heights.

Another potential use of these observations is in the assessment of wave models. However, as the data poorly represents the dynamic aspects of wavefields, which are very important in models designed for operational use, it can not be expected to provide a sound quantitative base for verification exercises.

Acknowledgments. I would like to thank Mr. S. W. Goulter, whose many useful suggestions and discussions have contributed to this work, and Mr. A. Penney, who helped with extracting the data from the archive. I am also grateful to Shell BP and Todd Oil Services Ltd. for allowing use of data from their waverider buoys.

APPENDIX A

Meteorological Code for Wave Observations

From 1 January 1968 until 1982 the code for the relevant parameters was as follows:

\[
99L_0L_0L_vQ_0L_0L_0L_0 YYGGx xddff xxxxx \ldots \]

\[
3P_wP_wH_wH_w d_wd_wT_wh_wh_w (d_wd_wT_wh_wh_w) \ldots ,
\]

where:

- \(L_0L_0L_v\) latitude in tenths of degrees (absolute value)
- \(Q\) octant in the globe
- \(L_0L_0L_0\) longitude in tenths of a degree
- \(YY\) day of the month
- \(GG\) time to nearest hour GMT
- \(dd\) direction from which the wind is blowing in tens of degrees
- \(ff\) wind speed in whole knots
- \(P_wP_w\) wave period in seconds (wind waves)
$H_w, H_w$ height of the wind waves in half-meters

d_w, d_w$ direction of swell in tens of degrees

$T_w$ swell period (0–3 = 10–13 s, 4 = 14 s or
greater, 5–9 = 5–9 s)

$h_w, h_w$ swell height in half-meters.

For more than one discernable swell train additional groups may be added.

There are possible anomalies in statistics compiled directly from reports made using this code. Specifically, in the half-meter resolution for wave heights, the one-second resolution for wave periods, and the 14-second limit on wave periods. For normally distributed variates this leads to almost zero error in means and variances if the resolution unit (half a meter or one second) is less than the standard deviation (Nicholson, 1979). In this case this criterion is easily satisfied. Although the distributions of the heights and periods are not normal, their logarithmic transforms approach normality, and although the transform distorts the resolution on the logarithmic scale, it is always much smaller than the standard deviation. The 14-second limit on periods may bias periods on the low side, but will also slightly increase the correlations. Since these are low enough as it is, and the amount of bias is insignificant, no problem is manifested here.

APPENDIX B

Confidence Limits for Spectral Estimates

If spectral values are calculated from the correlation series the estimate obtained is for a power spectrum (or the mean of the sample spectral distribution). It is well known that the ratio of the spectral estimate, $C_{xx}(k)$, to its expected value, $\Gamma_{xx}(k)$, is chi-squared [i.e., $nC_{xx}(k)/\Gamma_{xx}(k)$ is chi-squared with $n$ degrees of freedom; see Jenkins and Watts, 1968, Chap. 6.4], but it is not always obvious how many degrees of freedom should be quoted.

For a single-dimensional spatial series with single observations at 0, 40, 80, . . . 600 there are only 16 observations. The correlations for each distance of separation 40, 80, . . . would be based on 15, 14, 13, . . . 2 pairs of observations respectively, and the resulting spectral estimate would be chi-squared with 2 degrees of freedom. However, in our case we have thousands of pairs at each distance, although a single correlation is derived from them. Thus $n$ is considerably greater than 2.

There are perhaps two ways of looking at the problem. First, we can regard the data set as a composite of $n$ series, each one as previously described with 16 observations. The smoothed sample spectrum, $E_n[C_{yy}(k)]$, gained by averaging the $n$ sample spectra, $C_{yy}(k)$, from the $n$ series, has $2n$ degrees of freedom (Jenkins and Watts, 1968). To find $n$ we look to the category with the smallest number of pairs, which in all cases is the smallest separation class. In this case we have 1082 pairs in this class, implying at least 1082/15 = 72 series, or 144 degrees of freedom. Since the other separation classes have more pairs, this value is a minimum. The validity of this approach depends on whether the smoothed estimate of the sample spectrum converges to the single estimate from the combined data set,

$$E[C_{xx}(k)] = E_n[C_{yy}(k)] = \Gamma_{xx}(k), \quad \bigcup_{n} Y_n = X.$$

This is true if there are enough sample spectra contributing. Thus, it is not unreasonable to expect that with the large number of subsurfaces here the mean sample spectrum and the power spectrum from the combined set will in fact both converge to the population value. In reality, the subsurfaces do not exist, as the distribution of vessels is two-dimensional and is not sufficiently isotropic or dense enough in time to ensure unbroken sequences. A realistic value may therefore be something less than 144.

An alternative way of looking at the problem is to treat the distribution of vessels as an isotropic distribution on a two-dimensional grid. If the grid is 16 × 16 and has 74 km spacing then the number of pairs of vessels in each separation class can be calculated. These turn out to be reasonably in proportion to the actual number of pairs in the data set. In fact the ratio is about 1.5 to 1. This indicates that, if the data could be organized onto such a grid, about 1.5 grids would be required. Now in this grid there are 16 rows and 16 columns, each representing an idealized sequence, so it can be regarded as an arrangement of 1.5 × 16 independent spatial series. Thence, there are 48 degrees of freedom in the mean sample spectrum calculated from them, and the estimate of the power spectrum to which this converges must have at least 48 degrees of freedom. Once again this interpretation of the data distribution is not very real, since the data is from anything but a nice grid at a single time, and there are obviously many more than 256 vessels involved. The figure 48 would therefore be lower than a realistic value.

These arguments at best give us a rough idea, and it is reasonable to use a compromise of the two estimates. Thus, 64 degrees of freedom have been assumed.

REFERENCES

