Estimation of Directional Wave Spectra from Multicomponent Observations

T. H. C. HERBERS AND R. T. GUZA

Center for Coastal Studies, Scripps Institution of Oceanography, La Jolla, California

(Manuscript received 5 December 1989, in final form 4 April 1990)

ABSTRACT

An improved method for estimating the directional spectrum of linear surface gravity waves from in situ observations is presented. The technique, a refinement and extension of the inverse method of Long and Hasselmann, is applicable to multicomponent wave measurements at fixed locations in constant or slowly varying depth water. On a frequency band by frequency band basis, an estimate of the directional distribution of wave energy $S(\theta)$ is obtained by minimizing a roughness measure of the form $\int d\theta[d^2S(\theta)/d\theta^2]^2$ subject to the constraints: (i) $S(\theta)$ is nonnegative with unit integral, (ii) $S(\theta)$ fits the data within a chosen statistical confidence level, and (iii) $S(\theta)$ is zero on any directional sectors where energy levels are always relatively low because of the influence of geographic surroundings. The solution to this inverse problem is derived through a variational formulation with Lagrange multipliers.

A series of simulations using the new estimator show the fundamental limitations of sparse array data and the importance of using all available data-independent information [i.e., constraints (i) and (iii)] for achieving optimal estimates. The advantages of smoothness optimization are illustrated in a comparison of the present and Long and Hasselmann methods. The present method yields smooth estimates where Long and Hasselmann obtained rough estimates with multiple spurious peaks. A smooth solution to the inverse problem that has only truly resolved features is both easier to interpret and more readily evaluated numerically than wildly spurious solutions. The examples also demonstrate the subjectivity of intercomparing estimation techniques.

A few illustrative examples are presented of $S(\theta)$ estimates obtained from a two-dimensional array (aperture 120 m × 96 m) of 14 pressure transducer in 6 m water depth. Estimates using the full array and no geographic constraints are smooth and exhibit the expected refractive columnation of shoreward propagating energy towards normal incidence. Additionally, reflection from the mildly sloping beach 310 m shoreward of the center of this array is very weak at wind wave and swell frequencies. Estimates of $S(\theta)$ made using only the sensors on a longshore line, and a constraint of no reflected energy, are very similar to $S(\theta)$ obtained with the full array and no constraint.

1. Directional spectra estimation as an inverse problem

The frequency (σ)-directional (θ) spectral density $E(\sigma, \theta)$ is commonly used to statistically describe the sea surface excursion $\eta(\mathbf{x}, t)$ of natural wind-generated surface gravity waves. Often $E(\sigma, \theta)$ is estimated from data acquired with a small number of fixed instruments that sample sea surface elevation, surface slopes, subsurface pressure, or velocity components. Both point-type (for example the pitch and roll buoy, Longuet-Higgins et al. 1963; and others) and array-type (Barber 1963; and others) instrument configurations have been used for collecting directional wave data. In situ observations are inevitably spatially sparse. Sophisticated analysis is needed that effectively utilizes all the information contained in the data. Usually, time series can be collected that are sufficiently long to accurately resolve the cross-spectra in the frequency domain with Fourier analysis. The cross-spectrum $H_{nm}(\sigma)$ of a pair of array components with indices n and m can be expressed as a linear functional of $E(\sigma, \theta)$:

$$H_{nm}(\sigma) = \int_0^{2\pi} d\theta g_n(\sigma, \theta) g_m^*(\sigma, \theta) E(\sigma, \theta) \quad (1)$$

where the functions g_n and g_m (*denotes complex conjugate) relate the observed variables to the surface elevation η at a chosen array origin $\mathbf{x} = 0$ using linear theory. Since $E(\sigma, \theta)$ will be estimated on a frequency band by frequency band basis, we drop the frequency dependence from here on. The problem of estimating directional spectra from multicomponent wave data has been addressed by various authors. In this paper we give further refinements and extensions of the variational method of Long and Hasselmann (1979) applicable to array-type and point-type measurement systems of arbitrary sensor composition.

In the earliest attempts to estimate $E(\theta)$ from observed cross spectra $\tilde{H}_{nm}$ (Longuet-Higgins et al. 1963; Barber 1963; and others), simple analytic forms were arbitrarily chosen for the estimate $\hat{E}(\theta)$ (for example a truncated Fourier series or cosine power law) and the adjustable parameters were fit to the data. In the

Corresponding author address: Dr. Thomas H. C. Herbers, Scripps Institution of Oceanography, University of California, San Diego, Center for Coastal Studies, A-009, La Jolla, CA 92093-0209.

more systematic and "model independent" MLE method (Davis and Regier 1977) $\hat{E}(\theta)$ is expressed as a linear sum of the cross spectra with factorable weights:

$$\hat{E}(\theta) = \sum_n \sum_m \alpha_n(\theta)\alpha_m^*(\theta)\tilde{H}_{nm}; \qquad (2)$$

$\hat{E}(\theta)$ is minimized subject to the constraint that a plane wave is exactly resolved with perfect data. The numerically simple MLE method has been widely used. A fundamental shortcoming of both the MLE method and alternative methods proposed by Davis and Regier (1977) is that the differences between the observed $\tilde{H}_{nm}$ and the cross spectra $\hat{H}_{nm}$ associated with the estimate $\hat{E}(\theta)$:

$$\hat{H}_{nm} \equiv \int_0^{2\pi} d\theta g_n(\theta)g_m^*(\theta)\hat{E}(\theta) \qquad (3)$$

are not explicitly constrained to be smaller than some tolerance level for errors in the observations. If these "data misfits" ($\hat{H}_{nm}-\tilde{H}_{nm}$) are much larger than statistical and instrumental errors in the data ($\tilde{H}_{nm}-H_{nm}$) then the data is not optimally utilized and a tighter fit will generally improve the estimate.

Pawka (1983) gave an algorithm that iteratively modifies the MLE estimate until it fits the data ($\hat{H}_{nm} = \tilde{H}_{nm}$). Pawka's method and other similar techniques (Oltman-Shay and Guza 1984; Donelan et al. 1985; Marsden and Juszko 1987) return a directional spectrum that is "consistent" with the data. However, for a finite number of sensors there is not a unique $E(\theta)$ satisfying Eq. (1). A serious drawback of the algorithms of Pawka and others is that they are not based on an optimization principle, but on undefined assumptions about the shape of $E(\theta)$. Therefore it is not clear how other $E(\theta)$ that fit the same data differ from the selected estimate $\hat{E}(\theta)$. Insight in the resolving power of array data can only be gained through exhaustive model tests on a wide variety of directional spectra. Simulations with a limited class of $E(\theta)$ tailored to the algorithm may lead to unrealistic expectations of the resolving power of fundamentally limited wave observations.

In practice the observed $\tilde{H}_{nm}$ always contain errors due to the finite length of data records. Depending on the array geometry there may not be an $E(\theta)$ that fits the data exactly. For example, if $g_n(\theta)g_m^*(\theta) = g_p(\theta)g_q^*(\theta)$(e.g., two identical lags in an array of surface elevation gauges), then the misfits $\hat{H}_{nm}-\tilde{H}_{nm}$ and $\hat{H}_{pq}-\tilde{H}_{pq}$ cannot be zero simultaneously [Eq. (1)] unless $\tilde{H}_{nm} = \tilde{H}_{pq}$. In other cases the $E(\theta)$ that fit imperfect data exactly may be so rough that exact fit estimators are numerically unstable. Unless the estimator is restricted to particular measurement systems (e.g., Lygre and Krogstad 1986) some misfit between $\hat{H}_{nm}$ and $\tilde{H}_{nm}$ must be allowed for. Donelan et al. (1985) do allow for data misfits but their misfit tolerance is not based on data error statistics.

In this paper we adopt the Backus and Gilbert (1967) "inverse method," first applied to directional spectra estimation by Long and Hasselmann (1979), that does not have the drawbacks of the methods mentioned above. This approach is fundamentally different from that of Davis and Regier (1977), where estimation of the continuous function $E(\theta)$ is treated as a "parameter estimation problem," solving for $\hat{E}(\theta)$ independently at each angle θ. The inverse method solves a "function estimation problem," using integral constraints to solve for $\hat{E}(\theta)$ simultaneously at all angles. Although considerably more complex the "inverse method" approach seems more natural since integral properties of $E(\theta)$ are observed [Eq. (1)]. All information on the directional distribution $S(\theta)$ [$E(\theta)$ normalized by the frequency spectral density E], available from wave data and any other sources is cast in the form of constraints. Examples of constraints are

(i) $\hat{S}(\theta)$ fits the data $\tilde{H}_{nm}/\tilde{E}$ within a certain tolerance level for observation errors;

(ii) $\hat{S}(\theta)$ is a nonnegative function with unit integral on the interval $[0, 2\pi)$.

Only directional distributions that satisfy all constraints are considered acceptable estimates. The continuous function $S(\theta)$ cannot be specified by a finite number of these constraints so that in general there are infinitely many acceptable directional distributions. A unique estimate $\hat{S}(\theta)$ is found by minimizing some "nasty" property of $S(\theta)$. This measure of "nastiness" is subjective, but the assumptions on which $\hat{S}(\theta)$ is based are now clearly defined. More importantly, by solving for an extremum it is known that the entire class of $S(\theta)$ tolerably consistent with the observations are "nastier" than the selected $\hat{S}(\theta)$.

Long and Hasselmann (1979) solve for the $S(\theta)$ that is "as close as possible" to a preferred model $S_m(\theta)$ by minimizing $\int d\theta[S(\theta) - S_m(\theta)]^2$ subject to constraints i) and ii). In many situations the selection of $S_m(\theta)$ cannot be based on data-independent physical grounds, and the sensitivity of $\hat{S}(\theta)$ to the choice of $S_m(\theta)$ must be used to identify model-dependent structure in the estimates $\hat{S}(\theta)$ that is not necessary to explain the observations. When $S_m(\theta)$ is selected arbitrarily, only features present in $\hat{S}(\theta)$ irrespective of the choice for $S_m(\theta)$ are considered to be truly resolved features of the wave field. Features in $\hat{S}(\theta)$ that depend on $S_m(\theta)$ may or may not be real. These features of doubtful authenticity will hereafter be referred to as spurious. Routinely evaluating the effect of $S_m(\theta)$ on $\hat{S}(\theta)$ is rather cumbersome and in practical applications (e.g., Hasselmann et al. 1980; Lawson and Long 1983) $S_m(\theta)$ is a simple, smooth parametric expression fit to the data $\hat{H}_{nm}$. This somewhat biases the estimates to smooth solutions and suppresses wildly spurious structure in $\hat{S}(\theta)$ that can result from choosing an $S_m(\theta)$ that is very different from the true directional distri-

bution. The resulting $\hat{S}(\theta)$ cannot be interpreted as an extremum among all $S(\theta)$ consistent with the observations and thus provides limited insight in the resolving power of the data.

In this paper we explore an alternative formulation that has been applied to other oceanographic and geophysical inverse problems (e.g., Provost and Salmon 1986; Constable et al. 1987). We seek the "smoothest" $S(\theta)$ consistent with the observations by minimizing a roughness measure of the form $\int d\theta[d^2S(\theta)/d\theta^2]^2$ subject to constraints (i) and (ii). This approach, mentioned but not pursued by Long and Hasselmann (1979), has several advantages. The first advantage is rejection of spurious structure in $\hat{S}(\theta)$. The spatial extent of in situ wave observations is often very limited so that the kernels $g_n(\theta)g_m^*(\theta)$ in Eq. (1) are smooth functions. The gross features of $S(\theta)$ are constrained by such data, but very rapid oscillations in $S(\theta)$ are not resolved. That is, H_{nm} is rather insensitive to rapid fluctuations in $S(\theta)$ and such structure in an estimate is essentially arbitrary. These fundamentally unresolvable features tend to be suppressed in a smooth $\hat{S}(\theta)$. By minimizing the roughness of $S(\theta)$ we attempt to find the most featureless $\hat{S}(\theta)$ that satisfies all constraints. Ideally $\hat{S}(\theta)$ will have only features required by the data and no other peaks. The smoothest solution to the inverse problem is of interest because the interpretation of directional wave data is often focused, not on the detailed shape of $S(\theta)$, but on identifying and explaining peaks.

The second advantage of roughness minimization is that no a priori guess of $S(\theta)$ is required and $\hat{S}(\theta)$ provides an objective interpretation of the data. Unlike Long and Hasselmann (1979), the $\hat{S}(\theta)$ we obtain does not depend on a preferred model. The variational problem is considerably more complex than Long and Hasselmann (1979), but the numerical evaluation of $\hat{S}(\theta)$ is very stable. Like Long and Hasselmann (1979) the solution $\hat{S}(\theta)$ to the nonlinear inverse problem can only be evaluated iteratively. If $\hat{S}(\theta)$ is a very complex (rapidly fluctuating) function then iterative schemes may fail. Lawson and Long (1983) report convergence problems if the model $S_m(\theta)$ is very different from the true $S(\theta)$ so that $\hat{S}(\theta)$, attempting to fit this "wrong" model, is a very "rough" function. Seeking the smoothest solution to the inverse problem thus has the advantage that $\hat{S}(\theta)$ is a well behaved function that is easier to obtain numerically.

It is sometimes possible to supplement the wave data with independent physical constraints on $S(\theta)$. In cases where wave data is collected in the vicinity of land masses, $S(\theta)$ may be strongly affected by topographic effects. Refraction in shallow water, very weak reflection from nearby mildly sloping beaches, short fetches for wind wave generation, and wave blocking by nearby islands or headlands could result in shadow zones in $S(\theta)$ where energy levels are relatively very low regardless of meteorological forcing. It is shown below that such information can be effectively used by adding the constraint

(iii) $\hat{S}(\theta)$ vanishes on an arbitrary number of, a priori defined, directional sectors.

Data-independent constraints of the form (iii), based on a sound physical principle, can significantly improve the estimate $\hat{S}(\theta)$.

Several of the issues discussed above, and the properties of the estimator developed below, are illustrated by estimates of $S(\theta)$ based on simulated error-free cross spectra from a pitch and roll buoy. The orthogonalized kernels $[g_n(\theta)g_m^*(\theta)]$ of Eq. (1) are $\cos\theta$, $\sin\theta$, $\cos2\theta$ and $\sin2\theta$, and the basic resolving power of the buoy is very poor. An estimate $\hat{S}(\theta)$ consistent with such fundamentally limited data does not necessarily closely resemble the true directional distribution $S(\theta)$. As an example, Fig. 1 shows very different $S(\theta)$ that fit exactly (to computer accuracy) the same pitch and roll buoy data. Given this data an estimator that returns a bimodal distribution does not have "superior resolving power" to one that returns a unimodal distribution. The cross-spectral data alone, even without errors, are inherently inadequate to distinguish between the two. A "high resolution estimator" may arbitrarily select a bimodal distribution and the two peaks in this $\hat{S}(\theta)$ could be falsely interpreted as truly resolved (i.e., necessary to explain the observations) features of the wave field. With the method presented in this paper this possible overinterpretation of the data is avoided. The $\hat{S}(\theta)$ obtained through roughness minimization is in fact the unimodal distribution in Fig. 1. This is the "smoothest" and in this sense simplest $S(\theta)$ consistent with the data. On the other hand, if there were an a priori justification for assuming no energy in a sector around say $\theta = 180°$, blocking by an offshore island for example, then a topographic constraint can be included in the analysis. The dotted line in Fig. 1 illustrates that with the constraint $\hat{S}(\theta) = 0$ on $175° < \theta < 185°$ the smoothest distribution is now bimodal.

In section 2 the relation between the directional spectrum and the statistics of observed cross-spectra is discussed for the general case of an array of arbitrary composition and geometry in water of variable depth. A simple approximation is proposed for coastal applications on a seabed that is gently sloping in the on–offshore direction and uniform in the longshore direction. Redundancy in array data is also discussed. In section 3 the constraints on $\hat{S}(\theta)$ are defined, and the "smoothest" $\hat{S}(\theta)$ consistent with all constraints is obtained through variational calculus. To demonstrate that the new technique effectively rejects spurious structure in $\hat{S}(\theta)$, examples of model tests with artificial data are presented in section 4. The examples illustrate the limitations of sparse array data and the importance of data-independent information. An extensive array of pressure sensors was deployed in approximately 6-m deep water near Cape Canaveral, Florida in July

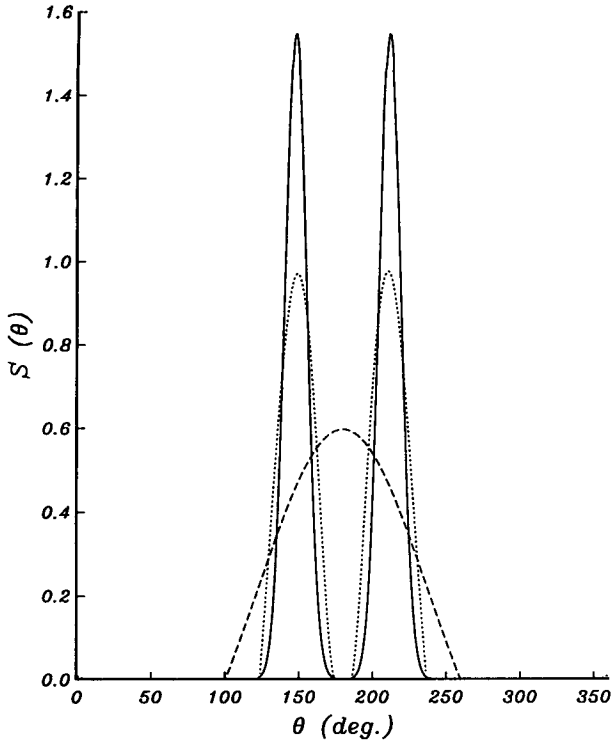


FIG. 1. Bimodal (solid and dotted lines) and unimodal (dashed line) directional distributions that fit the same cross-spectra of pitch and roll buoy measurements. The unimodal distribution is the "smoothest" $S(\theta)$ consistent with the data alone. The dotted bimodal distribution is the "smoothest" $S(\theta)$ consistent with both the data and a constraint of no energy on $175° < \theta < 185°$.

1988. Examples of directional spectra obtained from these observations are presented in section 5. The preliminary analysis of this data shows very weak reflection from the gently sloping beach, and supports the validity of a topographic constraint of the form (iii) in a typical coastal environment. In section 6, estimates obtained with the new technique are compared to model test results reported by Long and Hasselmann (1979) to highlight the differences in optimization principles. The subjectivity of estimator intercomparisons is discussed. The results are summarized in section 7.

2. Forward modeling

To formulate directional spectra estimation as an inverse problem, a physical model is needed that relates the directional spectrum to multicomponent wave observations. We assume that the water depth $h(\mathbf{x})$, with $\mathbf{x} = (x, y)$ the horizontal position relative to an arbitrary origin (Fig. 2), is slowly varying on wavelength scales. We also assume that the sea surface excursion $\eta(\mathbf{x}, t)$ is a linear superposition of independent Gaussian nearly plane wave components that satisfy the linear dispersion relation

$$\sigma^2 = gk(\mathbf{x}) \tanh[k(\mathbf{x})h(\mathbf{x})]; \quad k = |\mathbf{k}| \qquad (4)$$

with $\mathbf{k} = (k\cos\theta, k\sin\theta)$ the vector wavenumber (Fig. 2) and g the acceleration of gravity. The direction θ of a wave component can be expressed as a function of $\mathbf{x}$ and θ_0, the direction at $\mathbf{x} = 0$:

$$\theta = \theta(\mathbf{x}, \theta_0). \qquad (5)$$

We define $S(\theta_0)$ as the directional distribution of wave components with frequency σ at $\mathbf{x} = 0$. The cross spectrum (at frequency σ) H_{nm} of a pair of instruments with indices n and m (Fig. 2) normalized by E, the frequency spectral density (at frequency σ) of η at $\mathbf{x} = 0$, is related to $S(\theta_0)$ by

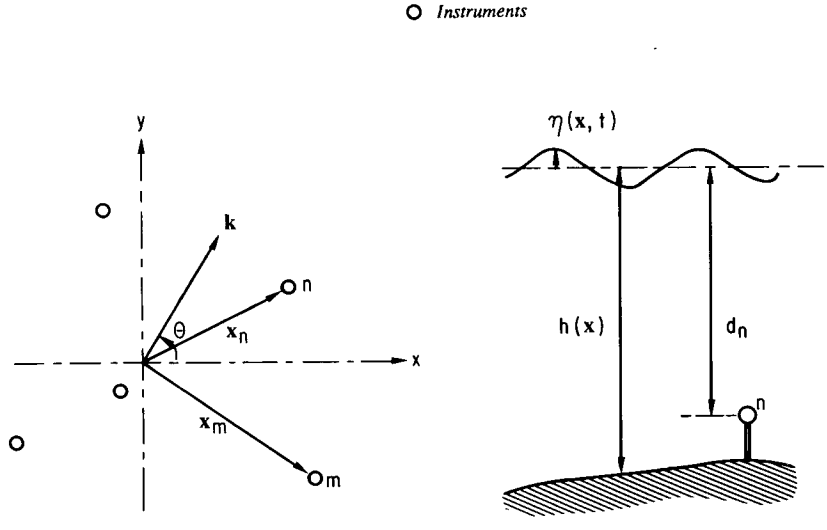


FIG. 2. Definitions of variables and coordinates.

$$\frac{H_{nm}}{E} = \int_0^{2\pi} d\theta_0 G_n[\theta(\mathbf{x}_n, \theta_0)] g_m^*[\theta(\mathbf{x}_m, \theta_0)] \times A(\mathbf{x}_n, \theta_0) A(\mathbf{x}_m, \theta_0) \exp\{-i[B(\mathbf{x}_n, \theta_0) - B(\mathbf{x}_m, \theta_0)]\} S(\theta_0) \quad (6)$$

where the vector $\mathbf{x}_n = (x_n, y_n)$ defines the horizontal position of instrument n (Fig. 2) and G_n is the linear transfer function that relates the observed variable to $\eta(\mathbf{x}_n, t)$. For a few commonly used instruments $G(\theta)$ is given in Table 1. The functions A and B describe the spatial evolution of the amplitude and phase of a wave component with direction θ_0, unit amplitude ($A(0, \theta_0) = 1$) and zero phase ($B(0, \theta_0) = 0$) at $\mathbf{x} = 0$. Simple approximations for θ, A, and B are discussed in appendix A.

Following Long and Hasselmann (1979) we express Eq. (6) in the general form:

$$\frac{\mathbf{d}}{E} = \int d\theta_0 \mathbf{b}(\theta_0) S(\theta_0) \quad (7)$$

where $\mathbf{d}$ and $\mathbf{b}(\theta_0)$ are a real data vector containing all auto-, co-, and quadrature spectra and the corresponding kernel vector, respectively, and $S(\theta_0)$ vanishes outside the limits of integration (Appendix A). The elements of Eq. (7) are not always linearly independent. For example, in water of uniform depth the co- and quadrature spectra of two identical instrument pairs with equal lags, each yield two identical equations. The sum of the autospectra of $\partial\eta/\partial x$ and $\partial\eta/\partial y$, and the autospectrum of η, obtained with a pitch and roll buoy (Table 1) also yields two linearly dependent equations. By definition

$$\int d\theta_0 S(\theta_0) = 1 \quad (8)$$

so that autospectra of pressure or sea surface elevation measurements are redundant in water of uniform depth. Equations (7) and (8) can be reduced to a linearly independent set through the diagonalization of the matrix $\mathbf{C}$ (e.g., Parker 1977):

$$\mathbf{C} \equiv \int d\theta_0 [\mathbf{b}(\theta_0) - \bar{\mathbf{b}}][\mathbf{b}(\theta_0) - \bar{\mathbf{b}}]^{\mathrm{T}} \quad (9)$$

TABLE 1. The transfer function $G(\theta)$ for measurements of surface elevation η, surface slopes $\partial\eta/\partial x$ and $\partial\eta/\partial y$, pressure p, and velocity components V_x and V_y. ρ is the density of sea water. x and d are the horizontal position and the depth of submergence of the instrument, respectively.

	$G(\theta)$
η	1
$\partial\eta/\partial x$	$ik(\mathbf{x})\cos\theta$
$\partial\eta/\partial y$	$ik(\mathbf{x})\sin\theta$
p	$\rho g \cosh[k(\mathbf{x})(h(\mathbf{x}) - d)]/\cosh[k(\mathbf{x})h(\mathbf{x})]$
V_x	$\sigma\{\cosh[k(\mathbf{x})(h(\mathbf{x}) - d)]/\sinh[k(\mathbf{x})h(\mathbf{x})]\}\cos\theta$
V_y	$\sigma\{\cosh[k(\mathbf{x})(h(\mathbf{x}) - d)]/\sinh[k(\mathbf{x})h(\mathbf{x})]\}\sin\theta$

with

$$\bar{\mathbf{b}} \equiv \frac{\int d\theta_0 \mathbf{b}(\theta_0)}{\int d\theta_0}. \quad (10)$$

Equation (7) is transformed to an orthonormal set [and orthogonal to Eq. (8)], $\mathbf{d}'$ with the eigenvectors $\mathbf{v}^n$ and eigenvalues λ^n of $\mathbf{C}$:

$$\mathbf{d}' \equiv \mathbf{W}^{\mathrm{T}}\left(\frac{\mathbf{d}}{E} - \bar{\mathbf{b}}\right) = \int d\theta_0 \mathbf{W}^{\mathrm{T}}[\mathbf{b}(\theta_0) - \bar{\mathbf{b}}] S(\theta_0) \quad (11)$$

where the columns of $\mathbf{W}$ are the normalized eigenvectors

$$\mathbf{W} \equiv \left\{\frac{\mathbf{v}^1}{(\lambda^1)^{1/2}}, \frac{\mathbf{v}^2}{(\lambda^2)^{1/2}}, \cdots\right\}. \quad (12)$$

A redundant element of Eq. (7) will manifest itself as a zero eigenvalue and the corresponding element of the transformed set of equations is removed. Additional eigenvalues of $\mathbf{C}$ are often very close to zero. The corresponding moments of $S(\theta_0)$ are extremely sensitive to the slightest inaccuracy in the data, the sensor positions, or the physical model. In practice the dataset does not contain useful information about these moments. For perfect data, an inverse $\hat{S}(\theta)$ that satisfies Eq. (11) exactly will be extremely sensitive to the data $\mathbf{d}$ (e.g., Parker 1977) if very small eigenvalues are retained, and the evaluation of $\hat{S}(\theta)$ may be numerically unstable. For imperfect data an inverse $\hat{S}(\theta)$ that satisfies Eq. (11) within a proper tolerance for errors may be insensitive to the data. However, the statistics of errors in the data are often only approximately known and the imposed data misfit tolerance criteria on $\hat{S}(\theta)$ may not always be adequate to prevent numerical instability. To obtain stable solutions to the numerically complex inverse problem discussed below we discard elements of Eq. (11) corresponding to eigenvalues λ^k that are smaller than a threshold value

$$\lambda^k < q\lambda^1 \quad (13)$$

with λ^1 the largest eigenvalue. The choice of q is somewhat arbitrary and is essentially a trade-off between numerical stability of the inverse model and a loss of data. Based on model tests we chose $q = 0.001$ as a truncation criterion that provides stable solutions with insignificant loss of data.

As an example, Fig. 3 shows the eigenvalues (ranked according to size) for a "slope array," four pressure sensors arranged in a square with dimension L, for various kL. The λ^9 through λ^{16} are identically zero because of the redundancy in the array geometry. In the limit $kL \to 0$ (Herbers and Guza 1989), λ^1 and λ^2 are proportional to $(kL)^2$ and correspond to $\int d\theta_0(\cos\theta_0, \sin\theta_0)S(\theta_0)$, and λ^3 and λ^4 are proportional to $(kL)^4$ and correspond to $\int d\theta_0(\cos 2\theta_0,$

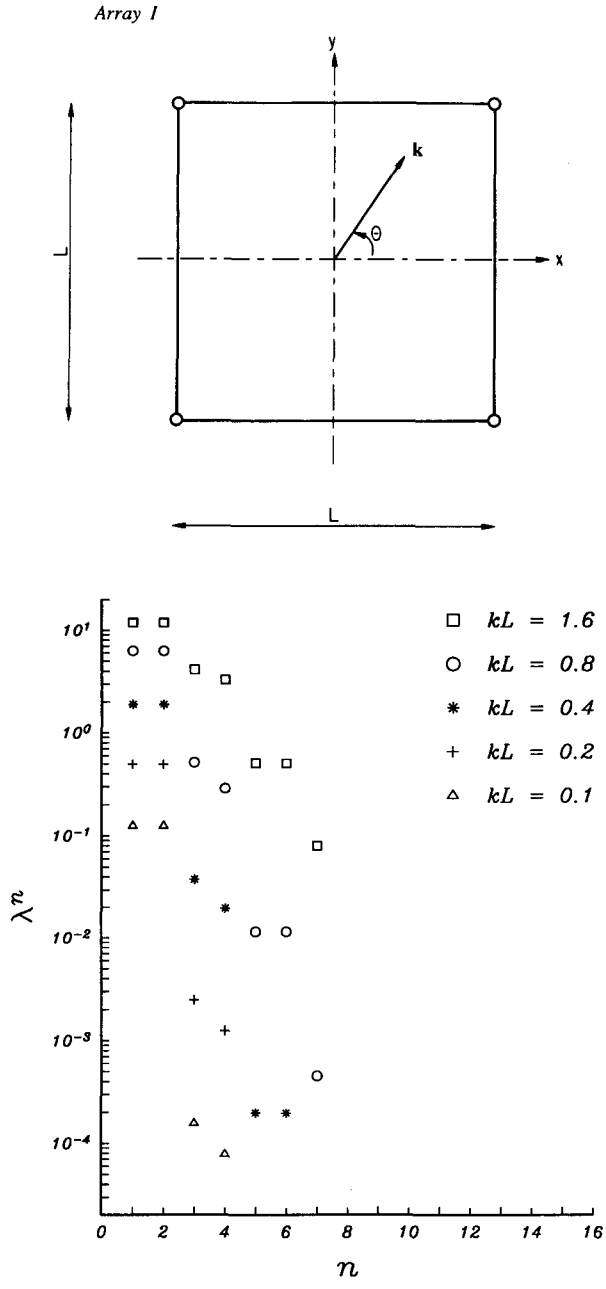


FIG. 3. Eigenvalues of the matrix C for a square slope array. (a) Array geometry. (b) Eigenvalues ranked according to size for various kL. Eigenvalues smaller than $2 \cdot 10^{-5}$ are not shown.

$\sin 2\theta_0) S(\theta_0)$. For $kL \ll 1$ the remainder of the eigenvalues contribute practically no information. For $kL = O(1)$ the curvature of the sea surface between the array elements is significant and, in principal, more information can be extracted from the slope array.

Data $\tilde{\mathbf{d}}$ obtained from observations will inevitably contain errors due to the finite length of data records. For a sufficiently narrow frequency bandwidth and a sufficiently large number of degrees of freedom υ, the cross-spectral estimates are approximately unbiased and joint Gaussian. The elements of the covariance matrix $\mathbf{Z}$

$$\mathbf{Z} \equiv \mathbf{E}\{(\tilde{\mathbf{d}} - \mathbf{d})(\tilde{\mathbf{d}} - \mathbf{d})^{\mathrm{T}}\} \tag{14}$$

with $\mathbf{E}$ the expected value are simply related to the true cross spectra (Jenkins and Watts 1968):

$$\mathbf{E}\{(\tilde{H}_{pq} - H_{pq})(\tilde{H}^*_{rs} - H^*_{rs})\} = \frac{2}{\upsilon} H_{pr} H^*_{qs}. \tag{15}$$

We assume that we can obtain an unbiased estimate of the frequency spectrum E with a linear combination of the elements of $\tilde{\mathbf{d}}$:

$$\tilde{E} = \mathbf{c}^{\mathrm{T}} \tilde{\mathbf{d}} \tag{16}$$

where $\mathbf{c}$ is an a priori specified coefficient vector (for example an average of the autospectra of pressure sensors located near $\mathbf{x} = 0$, converted to surface elevation spectra with Table 1). We can approximate the covariance matrix $\mathbf{V}$ of the normalized data

$$\mathbf{V} \equiv \mathbf{E}\left\{\left(\frac{\tilde{\mathbf{d}}}{\mathbf{c}^{\mathrm{T}}\tilde{\mathbf{d}}} - \frac{\mathbf{d}}{\mathbf{c}^{\mathrm{T}}\mathbf{d}}\right)\left(\frac{\tilde{\mathbf{d}}}{\mathbf{c}^{\mathrm{T}}\tilde{\mathbf{d}}} - \frac{\mathbf{d}}{\mathbf{c}^{\mathrm{T}}\mathbf{d}}\right)^{\mathrm{T}}\right\} \tag{17}$$

by expanding $\tilde{\mathbf{d}}/\mathbf{c}^{\mathrm{T}}\tilde{\mathbf{d}}$ for small errors (Long 1980):

$$\frac{\tilde{\mathbf{d}}}{\mathbf{c}^{\mathrm{T}}\tilde{\mathbf{d}}} \approx \frac{\mathbf{d}}{\mathbf{c}^{\mathrm{T}}\mathbf{d}} + \frac{\tilde{\mathbf{d}} - \mathbf{d}}{\mathbf{c}^{\mathrm{T}}\mathbf{d}} - \frac{\mathbf{d}\mathbf{c}^{\mathrm{T}}(\tilde{\mathbf{d}} - \mathbf{d})}{(\mathbf{c}^{\mathrm{T}}\mathbf{d})^2}. \tag{18}$$

Substitution of Eq. (18) in Eq. (17), using Eq. (14) yields:

$$\mathbf{V} \approx \left(\mathbf{I} - \frac{\mathbf{c}\mathbf{d}^{\mathrm{T}}}{\mathbf{c}^{\mathrm{T}}\mathbf{d}}\right)^{\mathrm{T}} \frac{\mathbf{Z}}{(\mathbf{c}^{\mathrm{T}}\mathbf{d})^2} \left(\mathbf{I} - \frac{\mathbf{c}\mathbf{d}^{\mathrm{T}}}{\mathbf{c}^{\mathrm{T}}\mathbf{d}}\right) \tag{19}$$

where $\mathbf{I}$ is the unit matrix. The covariance matrix $\mathbf{V}'$ of the transformed data vector $\tilde{\mathbf{d}}' = \mathbf{W}^{\mathrm{T}}(\tilde{\mathbf{d}}/\tilde{E} - \bar{\mathbf{b}})$ is

$$\mathbf{V}' = \mathbf{W}^{\mathrm{T}}\mathbf{V}\mathbf{W}. \tag{20}$$

There are alternative ways of normalizing the data. For example, a cross spectrum of surface elevation or pressure measurements in uniform depth can be normalized by the product of the autospectra of the two instruments. For a pitch and roll buoy, Long (1980) gives a normalization that eliminates the wavenumber dependence in Eq. (7). The sensitivity of directional spectra estimates to measurement errors may be affected by the way data is normalized, and the formulation presented here may not be optimal for every application. However, unlike the alternatives mentioned above, the more general Eq. (16) can be applied to an array of arbitrary composition in nonuniform water depth.

3. Inverse modeling

The relation between the directional distribution $S(\theta_0)$ and multicomponent data of surface gravity waves in deep or uniform depth water can be expressed in the general form [Eq. (11)]

$$\tilde{\mathbf{d}}' = \int_0^{2\pi} d\theta_0 \mathbf{b}'(\theta_0) S(\theta_0) + \epsilon \tag{21}$$

with $\tilde{\mathbf{d}}'$ the transformed data vector $\mathbf{W}^T(\tilde{\mathbf{d}}/\tilde{E} - \bar{\mathbf{b}})$, $\mathbf{b}'(\theta_0)$ the transformed kernel vector $\mathbf{W}^T \times (\mathbf{b}(\theta_0) - \bar{\mathbf{b}})$, and ϵ is an error vector. For varying water depth Eq. (21) takes the same form but with different integration limits (Appendix A). This case will be discussed later. The elements of ϵ are joint Gaussian with zero means and covariance matrix $\mathbf{V}'$ [Eq. (20)]. In the remainder of this paper we drop the primes on $\tilde{\mathbf{d}}$, $\mathbf{b}(\theta)$, and $\mathbf{V}$ and the subscript 0 on θ.

We seek an estimate $S(\theta)$ that is consistent with the data at a γ-confidence level, i.e., the misfit $\hat{\epsilon}$ between $\hat{S}(\theta)$ and the data

$$\hat{\epsilon} \equiv \tilde{\mathbf{d}} - \int_0^{2\pi} d\theta \mathbf{b}(\theta) \hat{S}(\theta) \tag{22}$$

must satisfy the constraint (Long and Hasselmann 1979)

$$\rho^2 \equiv \hat{\epsilon}^T \mathbf{V}^{-1} \hat{\epsilon} \leqslant \rho_\gamma^2 \tag{23}$$

where ρ_γ^2 is the γ-confidence level of the χ_N^2 (N is the dimension of $\hat{\epsilon}$) variable ρ^2. The matrix $\mathbf{V}$ is not known exactly. Long and Hasselmann (1979) obtain an approximation $\hat{\mathbf{V}}$ through (iterative) substitution of the cross-spectra corresponding to $\hat{S}(\theta)$ in Eq. (15). We have instead chosen an approximation $\tilde{\mathbf{V}}$ based on the observed cross-spectra. Although for small ϵ the matrices $\mathbf{V}$, $\hat{\mathbf{V}}$, and $\tilde{\mathbf{V}}$ are only slightly different, the inverse of the (often ill conditioned) covariance matrix may be sensitive to the chosen approximation. Fortunately $\mathbf{V}^{-1}$ does not appear in the solution for $\hat{S}(\theta)$ and extensive model tests indicate that $\hat{S}(\theta)$ is insensitive to small errors in $\tilde{\mathbf{V}}$. The confidence level γ is somewhat arbitrary but it provides an appropriate trade-off between resolving power and rejection of spurious structure in $\hat{S}(\theta)$. All the results presented in this paper are based on a 75% confidence level.

Additional constraints on $\hat{S}(\theta)$ that always have to be satisfied are

$$\int_0^{2\pi} d\theta \hat{S}(\theta) = 1 \tag{24}$$

$$\hat{S}(\theta) \geqslant 0. \tag{25}$$

The non-negativity constraint (25), which renders the inverse problem nonlinear and not solvable with standard linear inverse methods (e.g., Parker 1977), is eliminated by setting

$$\hat{S}(\theta) = F(\theta) H(\theta) \tag{26}$$

with

$$H(\theta) = \begin{cases} 1, & \text{for } F(\theta) \geqslant 0 \\ 0, & \text{for } F(\theta) < 0 \end{cases}. \tag{27}$$

Substitution of Eq. (26) in Eqs. (22) and (24) yields

$$\hat{\epsilon} = \tilde{\mathbf{d}} - \int_0^{2\pi} d\theta \mathbf{b}(\theta) F(\theta) H(\theta) \tag{28}$$

$$\int_0^{2\pi} d\theta F(\theta) H(\theta) = 1. \tag{29}$$

We seek the "smoothest" directional distribution $\hat{S}(\theta)$ consistent with the data by minimizing a roughness measure of the function $F(\theta)$ subject to the constraints (23) and (29). We assume that $F(\theta)$ is well behaved with derivatives $d^n F/d\theta^n(\theta)$ that are continuous on the domain $0 \leqslant \theta \leqslant 2\pi$ and periodic ($d^n F/d\theta^n(0) = d^n F/d\theta^n(2\pi)$) for $n = 0$, 1, 2, and 3. We minimize the roughness measure R defined as

$$R \equiv \int_0^{2\pi} d\theta \left[\frac{d^2}{d\theta^2} F(\theta)\right]^2. \tag{30}$$

Note that negative regions of $F(\theta)$ also contribute to R. This is necessary because negative regions of the solution $F(\theta)$ would be undetermined unless roughness is penalized on the entire domain $0 \leqslant \theta \leqslant 2\pi$. Minimizing the roughness of $\hat{S}(\theta)$ directly would be a preferable approach but no method (if there is one) has been reported that converges to the optimal (nonnegative) solution of this much more difficult optimization problem (Long, personal communication). In practice the inclusion of negative $F(\theta)$ in R is not crucial because the precise definition of "roughness" is somewhat arbitrary. The remaining sections of this paper show that minimizing the roughness measure R defined in Eq. (30) very effectively rejects spurious structure in $\hat{S}(\theta)$.

Introducing Lagrange multipliers α [for (23)] and β [for (29)] we minimize the functional

$$\phi \equiv R + \alpha \hat{\epsilon}^T \tilde{\mathbf{V}}^{-1} \hat{\epsilon} + \beta \int_0^{2\pi} d\theta F(\theta) H(\theta). \tag{31}$$

Setting $\delta\phi = 0$ and solving for β, using Eqs. (28) and (29), yields the equation

$$\frac{d^4}{d\theta^4} F(\theta) = \alpha \hat{\epsilon}^T \tilde{\mathbf{V}}^{-1} \mathbf{B}_0(\theta) \tag{32}$$

where $\mathbf{B}_0(\theta)$ is defined as

$$\mathbf{B}_0(\theta) \equiv \left[\mathbf{b}(\theta) - \frac{\int_0^{2\pi} d\theta \mathbf{b}(\theta) H(\theta)}{\int_0^{2\pi} d\theta H(\theta)}\right] H(\theta) \tag{33}$$

and $\hat{\epsilon}$ can be expressed in the form

$$\hat{\epsilon} = \mathbf{D} - \int_0^{2\pi} d\theta F(\theta) \mathbf{B}_0(\theta) \tag{34}$$

with

$$\mathbf{D} \equiv \tilde{\mathbf{d}} - \frac{\int_0^{2\pi} d\theta \mathbf{b}(\theta) H(\theta)}{\int_0^{2\pi} d\theta H(\theta)}. \quad (35)$$

Now define the functions

$$\mathbf{B}_{n+1}(\theta) = \int_0^{\theta} d\theta' \mathbf{B}_n(\theta') - (2\pi)^{-1} \int_0^{2\pi} d\theta' \times \int_0^{\theta'} d\theta'' \mathbf{B}_n(\theta'') \quad n = 0, 1, 2, 3 \quad (36)$$

and integrate Eq. (32) twice

$$\frac{d^2}{d\theta^2} F(\theta) = \alpha \hat{\epsilon}^{\mathrm{T}} \tilde{\mathbf{V}}^{-1} \mathbf{B}_2(\theta). \quad (37)$$

Integration of Eq. (34) by parts twice and substitution of Eq. (37) for the second derivative of $F(\theta)$ under the integral yields

$$[\mathbf{C}_2 + \alpha^{-1} \tilde{\mathbf{V}}] \lambda = \mathbf{D} \quad (38)$$

where $\mathbf{C}_2$ is the matrix

$$\mathbf{C}_2 \equiv \int_0^{2\pi} d\theta \mathbf{B}_2(\theta) \mathbf{B}_2{}^{\mathrm{T}}(\theta) \quad (39)$$

and the vector λ is defined as

$$\lambda \equiv \alpha \tilde{\mathbf{V}}^{-1} \hat{\epsilon}. \quad (40)$$

Finally, integrate Eq. (37) twice using Eq. (29) and substitute Eq. (40) for $\hat{\epsilon}$, yielding

$$F(\theta) = \lambda^{\mathrm{T}} \mathbf{B}_4(\theta) + \frac{1 - \lambda^{\mathrm{T}} \int_0^{2\pi} d\theta \mathbf{B}_4(\theta) H(\theta)}{\int_0^{2\pi} d\theta H(\theta)}. \quad (41)$$

Substitution of Eq. (40) in Eq. (23) yields the condition for α

$$\alpha^2 \geq \frac{\lambda^{\mathrm{T}} \tilde{\mathbf{V}} \lambda}{\rho_\gamma{}^2}. \quad (42)$$

Equations (38), (41), and (42) determine the solution to the inverse problem. This nonlinear system of equations is very similar to that obtained by Long and Hasselmann (1979). The only essential difference is the more convoluted way in which a "masking function" $H(\theta)$ appears in the equations presented here. The optimal solution $F(\theta)$ is found iteratively. A start solution is obtained by substituting $H^0(\theta) = 1$ for $H(\theta)$ and solving Eq. (38), sweeping through α values to find the minimum roughness $R = \lambda^{\mathrm{T}} \mathbf{C}_2 \lambda$ subject to constraint Eq. (42). Substitution of the solution vector λ and $H^0(\theta)$ in Eq. (41) yields the start solution $F^0(\theta)$. An improved guess $H^1(\theta)$ is determined through substitution of $F^0(\theta)$ in Eq. (27). Solving Eqs. (38), (41), and (42) with $H^1(\theta)$ yields $F^1(\theta)$, et cetera until a convergence criterion

$$\frac{\int_0^{2\pi} d\theta [F^n(\theta) - F^{n-1}(\theta)]^2 H^n(\theta)}{\int_0^{2\pi} d\theta [F^n(\theta) + F^{n-1}(\theta)]^2 H^n(\theta)} \leq \mu^2 \quad (43)$$

is met. The start solution $F^0(\theta)$ is in fact the optimal $\hat{S}(\theta)$ if the nonnegativity constraint (25) were relaxed. Often $F^0(\theta)$ and the final solution $F^n(\theta)$ are not too dissimilar and the iterative scheme converges rapidly ($\approx$5 iterations for $\mu = 0.01$). Like the Long and Hasselmann (1979) scheme, the iterations have a tendency to overshoot, in particular if $|dF(\theta)/d\theta|$ is large at zero crossings. Stability of the algorithm is achieved by replacing $F^n(\theta)$ by a weighted average of $F^{n-1}(\theta)$ and $F^n(\theta)$ ($0.9F^{n-1} + 0.1F^n$ was used in the estimates discussed below) at each iteration step.

The above given solution to the inverse problem is based only on wave measurements. In coastal applications data-independent information is often available in the form of a constraint on the direction of wave propagation. Deep ocean directional spectra are columnated by refraction as they propagate into shallow water; and bounds on wave directions at an array in shallow water can be obtained with ray theory. For example, on a beach that is uniform in the longshore direction, the angle enclosed by $\mathbf{k}$ and the beach normal is bounded by $\arcsin(\sigma^2/gk)$. A weaker form of the constraint that allows for larger angles of incidence but assumes that there are no caustics in the immediate array vicinity (Appendix A) is more appropriate if the array is surrounded by complex bathymetry. On mildly sloping beaches wind waves and swell are strongly dissipated in the surf zone and $S(\theta)$ might be assumed to be negligible small at seaward propagating angles.

Here we consider the general constraint:

$F(\theta) = 0$ everywhere except on the open intervals

$$(\phi_m{}^1, \phi_m{}^2) \quad m = 1, 2, \cdots, M \quad (44)$$

with M an arbitrary number of intervals on $(0, 2\pi)$. In this case without loss of generality, a coordinate frame is chosen such that $S(0) = S(2\pi) = 0$. The inverse problem with the additional constraint (44) is treated in the same way as the "data only" case and yields a nonlinear system of equations that has the same form as Eqs. (38), (41), and (42), and can be solved with the same iterative scheme outlined above. The details of the variational formulation are given in Appendix B. Examples of applications of constraint (44) are given in the next sections.

The nonnegativity constraint on $\hat{S}(\theta)$ [Eq. (25)] renders the inverse problem nonlinear and much more difficult to solve. However, as was shown by Long (1986) and is also illustrated with an example in section

4, this constraint is crucial in optimizing directional spectra estimates. If $S(\theta)$ is very narrow then constraint (25) greatly improves the accuracy of $\hat{S}(\theta)$ but also causes the matrix $\mathbf{C}_2$ [Eq. (39)] to become more ill-conditioned at every iteration step. Instability of the numerical scheme can be prevented by reorthogonalizing the kernel vector $\mathbf{b}(\theta)$ [Eqs. (9) through (12)], using $H^n(\theta)$ of the last iteration step as a weighting function in the matrix $\mathbf{C}$ [Eq. (9)].

When applying a topographic constraint on the wave direction θ we absorb this information in the forward model by substituting

$$\sum_{m=1}^{M} \int_{\phi_m^1}^{\phi_m^2}$$

for the integration limits in Eq. (9). A linear data-independent constraint of the form (44) generally "weakens" the nonlinear nonnegativity constraint (25) and enhances the stability and the efficiency of the numerical scheme.

4. Model tests

In this section we present a series of model tests applying the new directional spectrum estimator to simulated array data. The purpose of these tests is to show that minimizing the roughness measure R [Eq. (30)] indeed yields smooth directional distributions without spurious peaks. The examples also illustrate how the various constraints contribute to the estimate and, in particular, show the importance of data-independent information [$S(\theta) \geq 0$ and topographic constraints]. Two simple arrays, with four surface elevation (or pressure) gauges each, were chosen for the tests. The test arrays are a square with dimension L in deep (or uniform depth) water (Fig. 3a), and a linear array with spacings $2L$, L, $4L$, aligned with the depth contours and the x-axis (Fig. 4). Unlike the square array, the resolving power of the linear array is a strong function of wave direction. Wave components with direction θ and $2\pi-\theta$ cannot be distinguished by the linear array and the resolution is very poor for θ close to 0 or π. On mildly sloping (assumed to be weakly reflective) beaches, linear arrays aligned with the depth contours have been deployed to optimize the resolving power for directional spectra that are strongly columnated to near normal incidence ($\theta \approx \pi/2$) by refraction (e.g., Pawka 1983).

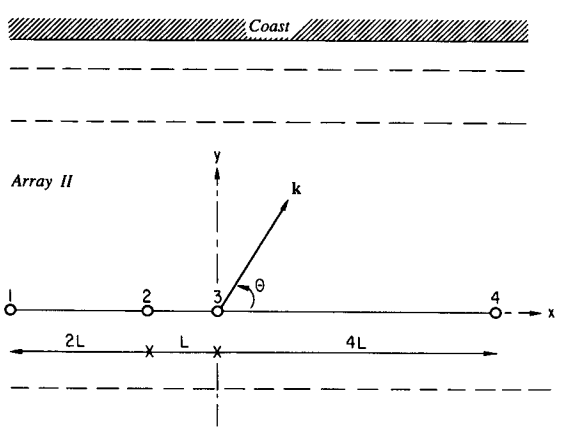


FIG. 4. Geometry of the simulated linear array. The dashed lines are depth contours.

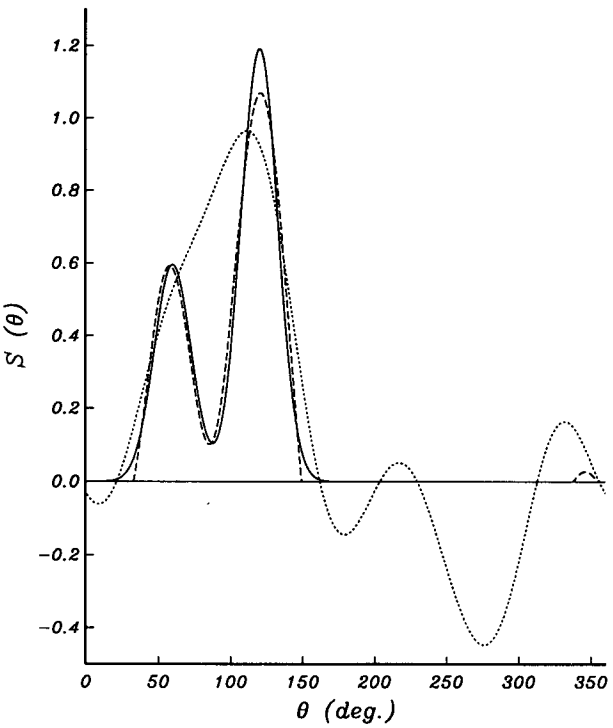


FIG. 5. Response of the square array (kL = 3.2) to a bimodal input directional distribution (solid line) of the form $S(\theta) \sim \frac{1}{2}\cos^{80}[(\theta-60°)/2] + \cos^{80}[(\theta-120°)/2]$. The dashed and dotted lines are estimates obtained from perfect data ($\nu = \infty$), with and without the constraint $\hat{S}(\theta) \geq 0$, respectively.

The procedure of the model tests is as follows. For a chosen input ("true") directional distribution $S(\theta)$, the "true" cross-spectra H_{nm} are computed with Eq. (6)(the value of E is arbitrary). Next the data covariance matrix $\mathbf{Z}$ is evaluated [Eqs. (14) and (15)] and random realizations $\tilde{H}_{nm}$ with ν degrees of freedom are generated (the procedure is described in Long and Hasselmann 1979). The simulated $\tilde{H}_{nm}$ are treated as "data" in exactly the same way as ocean observations, and an estimate $\hat{S}(\theta)$ of the directional distribution is obtained with the method described in the previous sections.

Figure 5 compares estimates for the square array (kL = 3.2) with and without the constraint that $S(\theta)$ must be a nonnegative function [Eq. (25)]. The estimates are based on perfect data ($\nu = \infty$ and error free sensors) of a bimodal directional distribution with

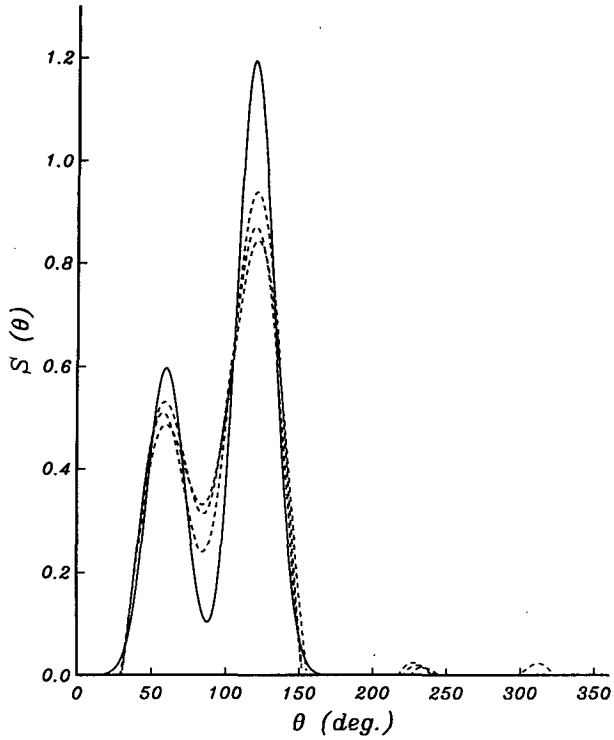


FIG. 6. Estimates of $S(\theta)$ obtained with the square array (kL = 3.2) for the same input directional distribution (solid line) as Fig. 5. The three dashed lines are estimates for 3 independent realizations of cross-spectra $\tilde{H}_{nm}$ with υ = 180.

peaks separated by 60°. The estimate utilizing the nonnegativity constraint accurately resolves the two peaks without any spurious structure. Relaxing the constraint yields a smoother unimodal distribution with negative side lobes. The data by itself evidently cannot resolve the bimodal structure of $S(\theta)$ and the nonnegativity constraint contributes important information. Adding the constraint Eq. (25) renders the optimization problem nonlinear and complicates the numerical effort required to evaluate $\hat{S}(\theta)$. However, this example (and similar tests, Long 1986) shows that a nonnegativity constraint is crucial for obtaining optimal results. All estimates in the remainder of this paper utilize the nonnegativity constraint. Additionally, the data will include realistic statistical errors.

For the same bimodal input directional distribution (Fig. 5) and array, Figs. 6 and 7 show estimates of $S(\theta)$ based on data records of finite length (υ = 180) for kL = 3.2 and 0.4, respectively. In each case estimates are shown for three independent data realizations. The estimates of Fig. 6 (same wavelength as Fig. 5) resolve the bimodal structure of $S(\theta)$ but, naturally, not as accurately as the perfect data estimate of Fig. 5. All three estimates are smaller than $S(\theta)$ at the peaks and larger in the valley. This smearing effect of the estimator is expected because the smoothness constraint is more dominant when less information can be extracted from the imperfect data. Figure 7 shows the same test with the array size reduced to only $\frac{1}{16}$th of a wave length. The resolving power is reduced by the small array aperture to the extent that the two peaks of $S(\theta)$ are smeared to a broad unimodal shape in the estimates. Figures 6 and 7 demonstrate that simple featureless solutions to the inverse problem can be obtained with the present optimization scheme. In both cases the estimates from three data realizations are similar. These examples (and numerous other tests, not reported here) indicate that the inverse method is insensitive to small errors in the covariance matrix $\tilde{\mathbf{V}}$ obtained from data.

Finally, in order to compare the data misfits of $\hat{S}(\theta)$ to the errors generated in the simulated cross-spectra, the "model" cross-spectra $\hat{H}_{nm}$ are evaluated through substitution of $\hat{S}(\theta)$ in Eq. (6). Figures 8 and 9 show the deviations of $\tilde{H}_{nm}/\tilde{E}$ ("data") and $\hat{H}_{nm}/E$ ("model") from the true normalized cross-spectra H_{nm}/E for the examples of Figs. 6 and 7, respectively. In both cases the errors in $\hat{H}_{nm}/E$ are comparable in size to the errors in $\tilde{H}_{nm}/\tilde{E}$. For kL = 0.4 the errors are considerably smaller than for kL = 3.2. This is due to the fact that the cross-spectra of different sensor pairs become increasingly coherent as kL decreases and the errors in $\tilde{H}_{nm}$ and $\tilde{E}$ ($=[\tilde{H}_{11} + \tilde{H}_{22} + \tilde{H}_{33} + \tilde{H}_{44}]/4$) tend to cancel (e.g., Herbers and Guza 1989). For very small values of kL errors in the observations due to other sources (for example instrument and digitization

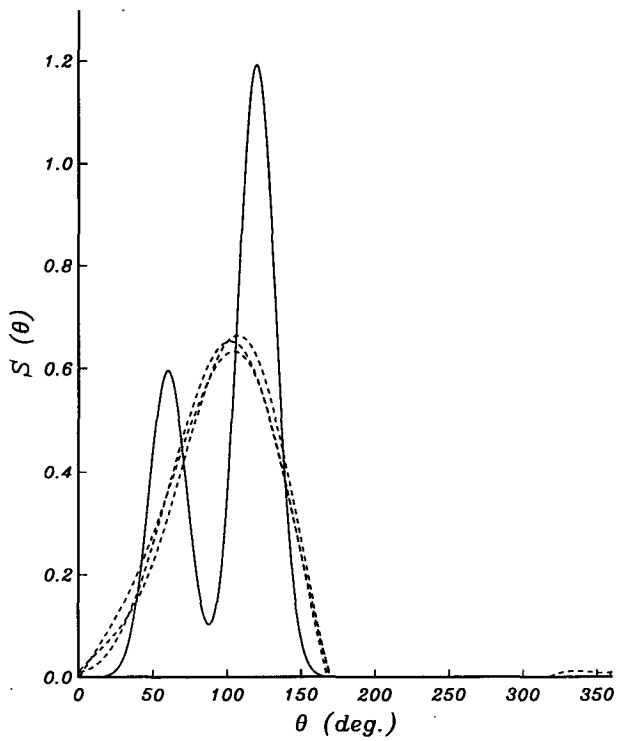


FIG. 7. Estimates of $S(\theta)$ obtained with the square array for the same input directional distribution as Fig. 5. kL = 0.4 and υ = 180 (same format as Fig. 6).

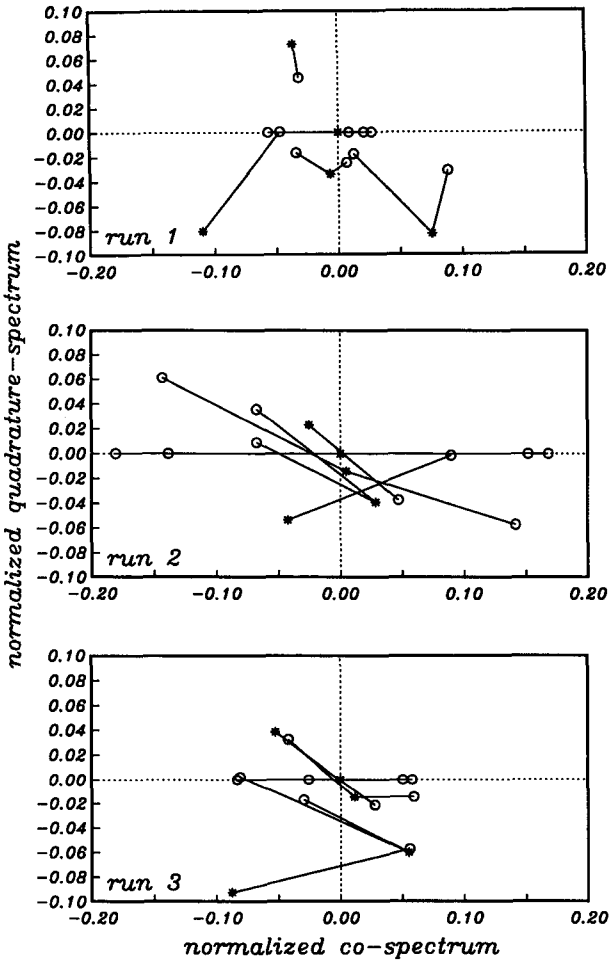


FIG. 8. Errors in the data $\tilde{H}_{nm}/\tilde{E} - H_{nm}/E$ (circles) and errors in the estimate $\hat{H}_{nm}/E - H_{nm}/E$ (asterisks) for the model test of Fig. 6. For every sensor pair (n, m) the two errors are connected by a straight line. Each panel shows the results for one data realization.

noise, turbulence, and uncertainty in sensor positions) are not negligible compared to the statistical uncertainty in finite length data records. The inverse model, imposing a too stringent data fit constraint, may become unstable for $kL \ll 1$ unless measurement errors are allowed for.

In the simulations presented so far, no geographic constraint [Eq. (44)] has been imposed on $\hat{S}(\theta)$. The next three examples (Figs. 10, 11, 12) show the utility of such a constraint for the linear array (Fig. 4) with dimension $kL = 2.4$. The input directional distribution is again bimodal but, in order to allow for refraction in shallow water, the peaks are narrower and separated by only 24°. As before, in each test estimates are given for 3 random realizations of cross-spectra having 180 degrees of freedom. The estimates $\hat{S}(\theta)$ in Figs. 10–12 are constrained to vanish for $180° < \theta < 360°$, based on the assumption that reflection from the beach is negligible. The estimates in Fig. 10, using all sensors in the array, fairly closely resemble the true directional distribution and resolve the bimodal structure. Figure 11 shows results for the same test, but with sensor 2 omitted from the array (Fig. 4) to simulate the effect of an instrument failure. Without sensor 2, the smallest sensor separation in the array exceeds one surface wavelength (for $kL = 2.4$) and the wave field is severely undersampled. The main feature of $S(\theta)$, all wave components close to normal incidence ($\theta = 90°$), is not resolved by the estimates. In the "smoothest" solution to the inverse problem, the two peaks in $S(\theta)$ are completely aliased to large angles of incidence where the array has less resolving power. The data by itself is evidently grossly inadequate, but if combined with additional data-independent information, an accurate estimate of $S(\theta)$ may be obtained. Figure 12 shows estimates of $S(\theta)$ for the same data as Fig. 11 but with the additional constraint that wave incidence angles are reduced by refraction to $60° < \theta < 120°$. The estimates agree very well with the bimodal input directional distribution. The crucial information miss-

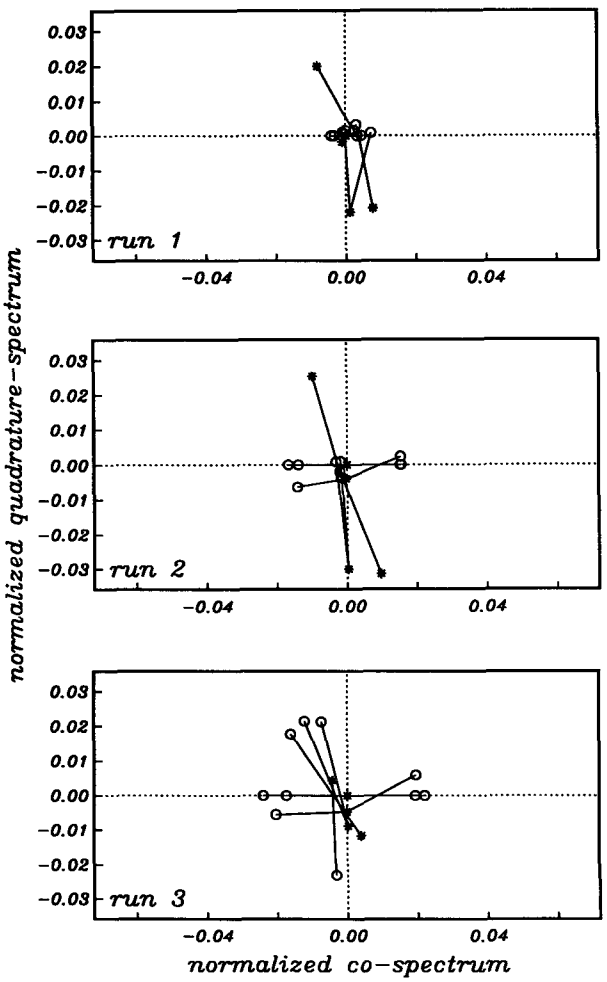


FIG. 9. Errors for the model test of Fig. 7 (same format as Fig. 8).

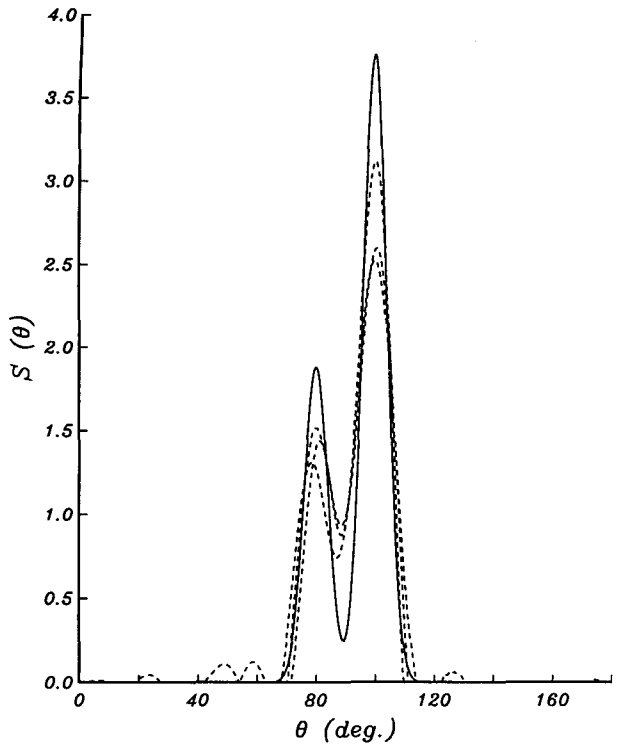


FIG. 10. Response of the linear array (kL = 2.4) to a bimodal directional distribution of the form $S(\theta) \sim \frac{1}{2}\cos^{800}[(\theta-78°)/2] + \cos^{800}[(\theta-102°)/2]$. The cross spectra have v = 180 and a "no reflection" constraint: $\hat{S}(\theta) = 0$ on $180° < \theta < 360°$ was imposed on the estimates (same format as Fig. 6).

ing when sensor 2 fails is provided by the geographic constraint. The example illustrates how the use of both array observations and geographical data in the inverse problem can provide more information than either one of them alone.

5. Examples of estimates obtained from field observations

During July 1988 a 14-element array of pressure sensors was deployed south of Cape Canaveral, Florida, to investigate reflection of surface gravity waves from a natural beach. A detailed description and analysis of these observations will appear elsewhere. The array (Fig. 13) spanned 120 m in the longshore direction and 96 m in the cross-shore direction, and the pressure sensing elements were about 0.5 m above the bed. The sample frequency was 4 Hz. At mean high water, the array center was located approximately 310 m from the shoreline in a water depth of 6 m. Extensive bathymetric surveys at the instrumented site (Kirby 1988) during, before, and after the experiment show that the seabed was uniform in the longshore direction and monotonic, gently sloping (1:30 at the shoreline to 1:200 at the array) in the cross-shore direction. Depth variations in the array vicinity are sufficiently small to assume a uniform depth (Appendix A). Using a few illustrative examples, we show that the estimator produces directional distributions that are physically realistic, and that reflection from the gently sloping beach is weak.

Although the array is too limited in size to very accurately resolve long swell, it has good omnidirectional response and can detect weak reflections from the coast. Figure 14 shows tests for a hypothetical example of bimodal shoreward propagating waves ($0° < \theta < 180°$), partially reflected from the beach, for both typical wind wave and swell frequencies ($f = \sigma/2\pi$ = 0.2 Hz and 0.1 Hz). No geographic constraints (Eq. 44) were imposed in these simulations. At both frequencies all peaks in $S(\theta)$ are essentially resolved. For 0.2 Hz wind waves, the array spans several wavelengths, and accurate estimates of $S(\theta)$ are obtained. For longer wavelength 0.1 Hz swell, the estimates are less accurate. Insufficient resolving power results in smoother $\hat{S}(\theta)$ with smaller peak values. However, even the smallest peak in $S(\theta)$ is present in all estimates. Additionally, in model tests (not shown) similar to Fig. 14 but with zero true reflected energy, the estimated reflected energies are very small. Thus the presence (or absence) of weak seaward propagating components in the directional spectrum can be detected with this array even if directional detail of the incident wave field is not well resolved.

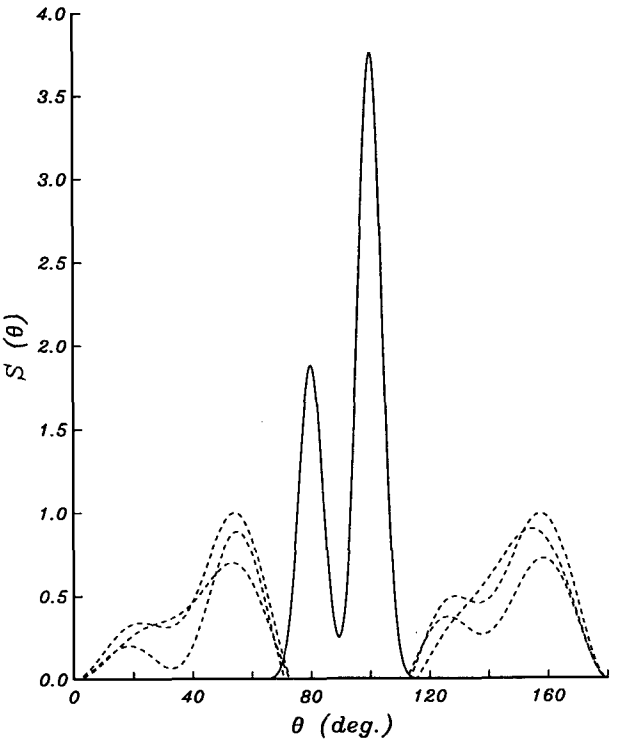


FIG. 11. Results for the same test presented in Fig. 10 but with sensor 2 missing in the array.

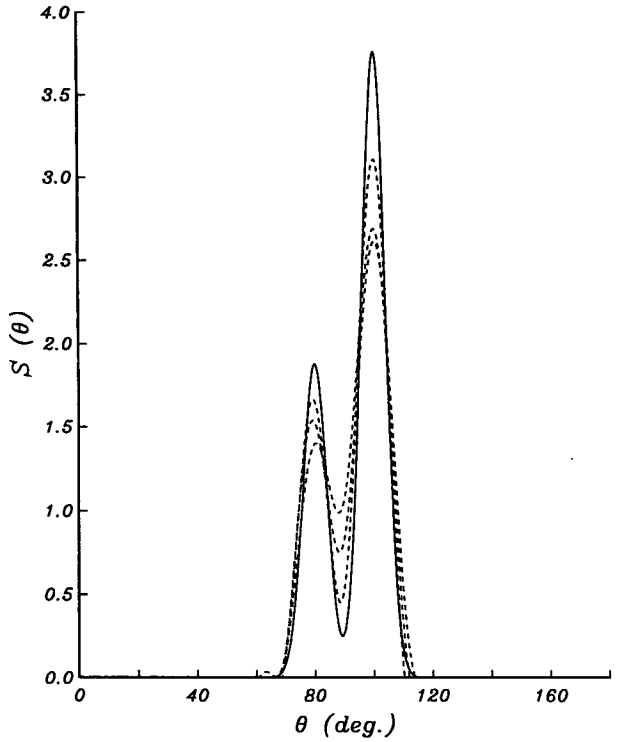


FIG. 12. Results for the same test presented in Fig. 11, but with the additional constraint that $\hat{S}(\theta)$ vanishes for large angles of incidence: $\hat{S}(\theta) = 0$ on $0° < \theta < 60°$ and $120° < \theta < 180°$.

Examples of energy and directional spectra obtained from the Cape Canaveral data are given in Figs. 15 and 16. No geographic constraints on the wave direction θ (Eq. 44) were imposed on these estimates. In

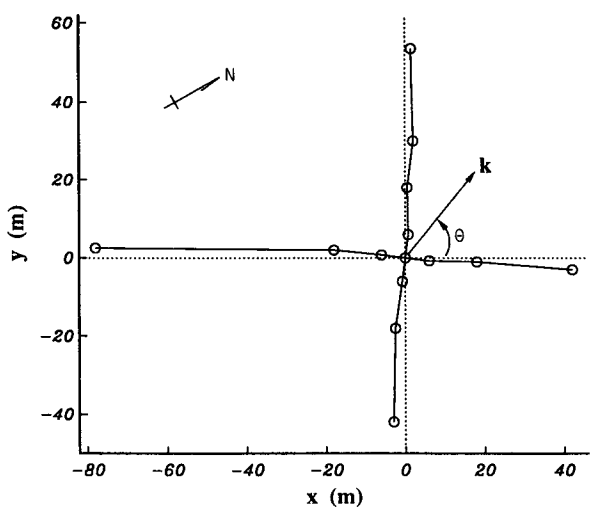
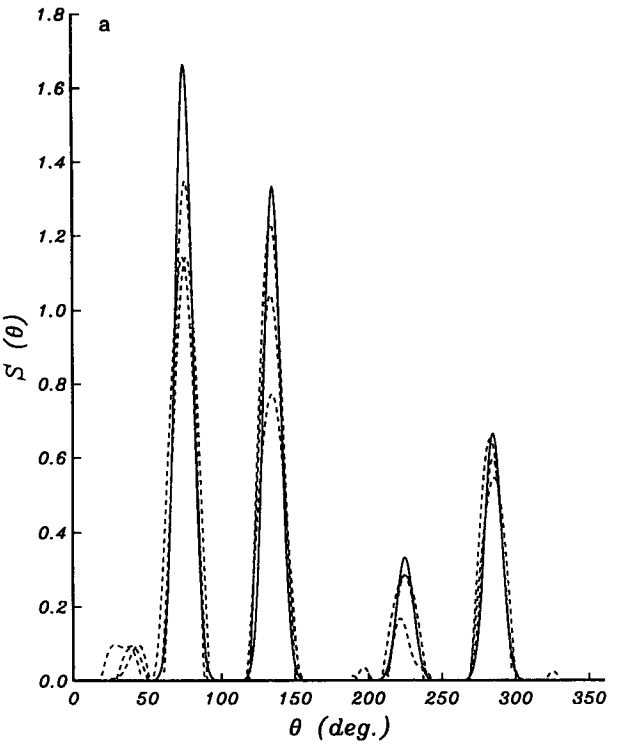


FIG. 13. Geometry of the pressure sensor array deployed approximately 1½ km north of the entrance to Port Canaveral in July 1988. The x-axis is aligned with the shoreline (33° relative to true north) and positive y is directed onshore.



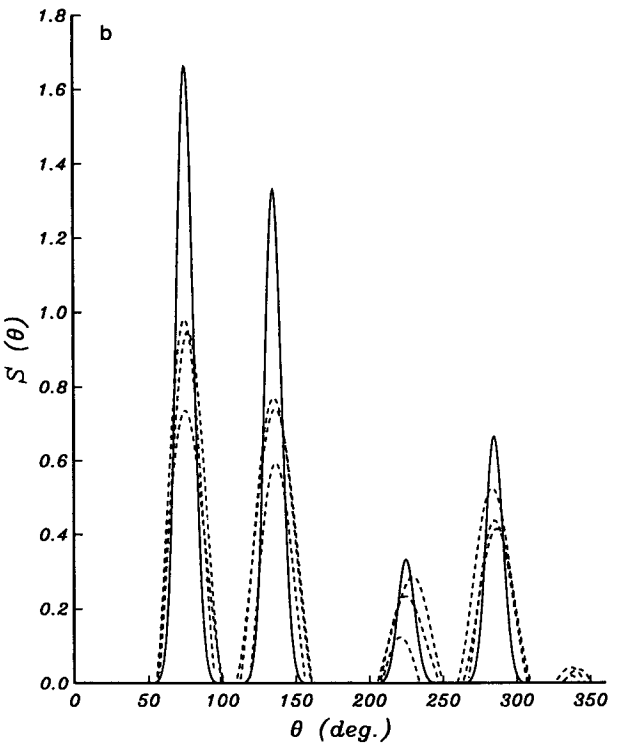


FIG. 14. Response of the Cape Canaveral array (Fig. 13) to a directional distribution of the form $S(\theta) \sim \cos^{400}[(\theta-75°)/2] + 0.8\cos^{400}[(\theta-135°)/2] + 0.2\cos^{400}[(\theta-225°)/2] + 0.4\cos^{400}[(\theta-285°)/2]$; $v = 144$. (a) wind waves $f = 0.2$ Hz, (b) swell $f = 0.1$ Hz.

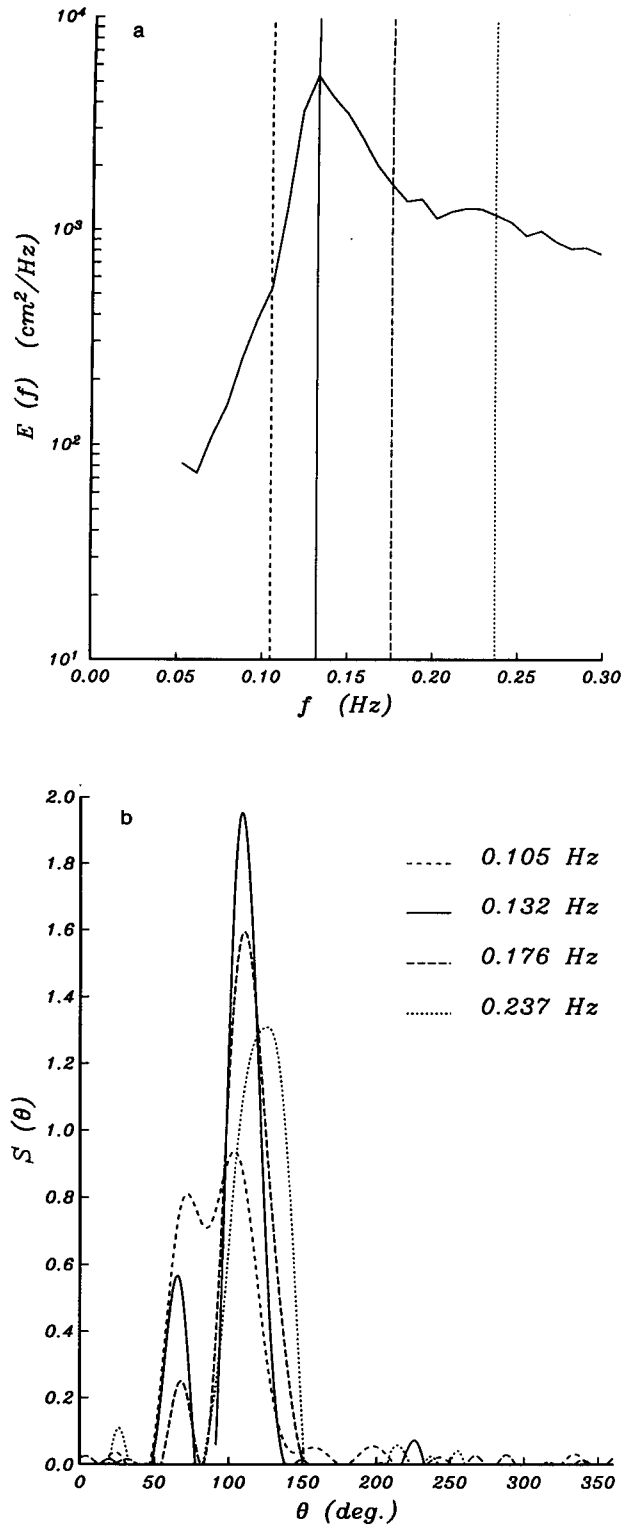


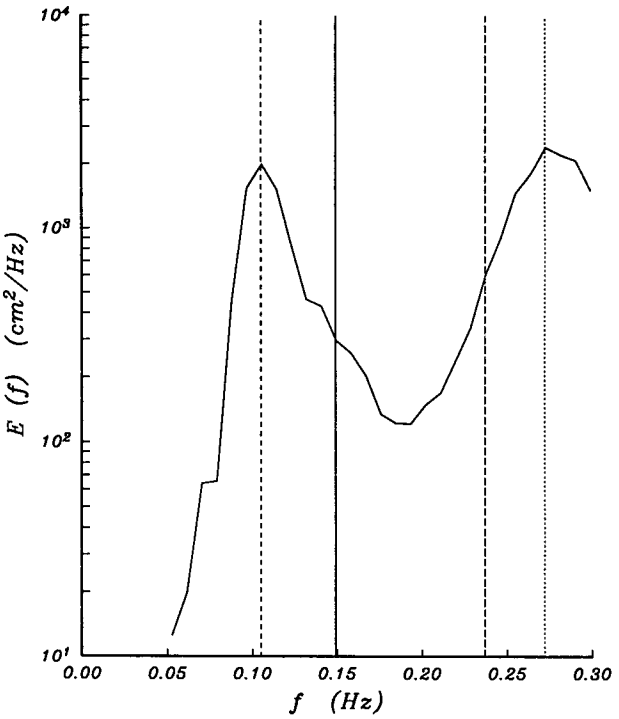
FIG. 15. Directional wave spectrum estimates obtained from data collected between 4 July 22.33 h and 5 July 00.50 h, 1988 with the array shown in Fig. 13. (a) the frequency spectrum $\tilde{E}$, total variance = 368 cm^2. (b) directional distributions $\hat{S}(\theta)$ at four frequencies (indicated by vertical lines in $\tilde{E}$).

each case $\tilde{E}$, the mean of the surface elevation frequency spectra (computed for each sensor using Table 1), and $\hat{S}(\theta)$ at four selected frequencies are shown for a 136.5 minutes data run. As in the model tests of Fig. 14, $v = 144$. The frequency spectrum was unimodal on 4/5 July (Fig. 15a) and bimodal on 16/17 July (Fig. 16a). In these examples (Figs. 15b, 16b), and in other cases (not shown), $\hat{S}(\theta)$ is smooth with either unimodal or bimodal shapes.

On 4/5 July (Fig. 15b) both peaks of the bimodal $\hat{S}(\theta)$ are, in general, increasingly closer to normal incidence ($\theta = 90°$) with decreasing frequency. This is qualitatively consistent with the refraction of a spectrum of waves having common peak directions in deep water. The estimated directional spectrum on 16/17 July is more complicated (Fig. 16b). At the swell peak frequency (0.105 Hz) $\hat{S}(\theta)$ is unimodal with the peak direction ($\theta = 87°$) close to normal incidence. At $f = 0.149$ Hz the estimate $\hat{S}(\theta)$ is bimodal with two peaks of nearly equal magnitude separated by only 24° (peak directions 73° and 97°). It is possible that similar bimodal structure is also present at $f = 0.105$ Hz with even smaller peak separation, as longer wavelength 0.105 Hz swell from the same two generation sources would be more strongly refracted to normal incidence at the array. The array simply lacks resolving power to separate strongly columnated swell peaks, and a smooth unimodal estimate of $S(\theta)$ can be obtained for $f = 0.105$ Hz that is consistent with the data. Note that the directional peaks at 0.105 Hz, which were resolved on 4/5 July, were separated by a relatively larger 32°. The two directional distributions at higher frequencies shown in Fig. 16b are very different. $\hat{S}(\theta)$ at the wind wave peak frequency $f = 0.272$ Hz is narrow with peak direction $\theta = 56°$ while $\hat{S}(\theta)$ at $f = 0.237$ Hz is broad with peak direction $\theta = 95°$.

The key feature of $\hat{S}(\theta)$ on both days is the very low energy levels for both waves reflected from the shoreline ($|\theta - 90°| > 90°$) and waves propagating shoreward with more than 60° deviation from normal incidence ($60° < |\theta - 90°| < 90°$). On plane parallel contours, 0.1 Hz waves are in theory refracted to within 60° of normal incidence, even for grazing incidence in deep water. The estimates are thus consistent both with refraction and weak reflection, and illustrate the validity of topographical constraints used in the model tests (Figs. 10–12).

The fraction of $\hat{S}(\theta)$ directed offshore (reflected) varies between 1% and 3% for the eight estimates shown in Figs. 15b and 16b. Although integrating $\hat{S}(\theta)$ is not an optimal method for estimating bulk reflection coefficients, it seems very unlikely that the ratio of reflected/incident wave energy on this mildly sloping beach exceeded 10%. The assumption of no reflected energy from natural beaches has been used in the past to remove the directional ambiguity of linear arrays (e.g., Pawka 1983). Low levels of reflected energy have been



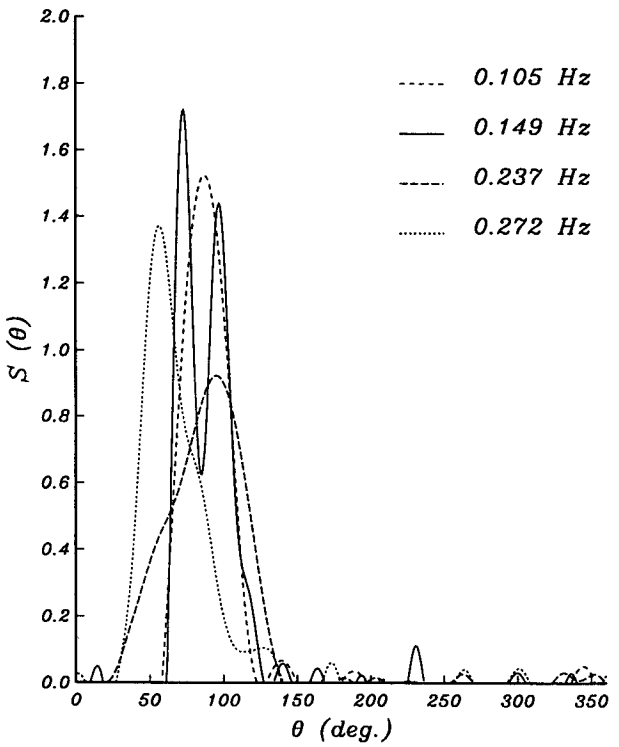


FIG. 16. Directional wave spectrum estimates obtained from data collected between 16 July 23.36 h and 17 July 01.53 h, 1988 with the array shown in Fig. 13 (same format as Fig. 15). Total variance = 199 cm^2.

reported in many laboratory studies using monochromatic, unidirectional waves. However, the spatially extensive two-dimensional arrays needed to measure reflected waves in a directionally spread natural wave field have rarely been deployed close to shore. Freilich and Guza (1984) report average (across the wind wave and swell frequency bands) reflected energies of about 15% from a similar mildly sloping California beach. They conclude (based on model tests) that this figure is biased high and the true amount of seaward propagating energy is probably very small (0–10% of the total energy), consistent with the present observations.

A priori assumption of weak reflectivity of the Cape Canaveral beach can be utilized to drastically reduce the number of instruments in the array of Fig. 13 without degrading the directional spectra estimates. This is demonstrated in Fig. 17 where the calculations of Figs. 15 and 16 were repeated using only the seven sensors on the longshore axis. A geographic constraint of no reflection: $\hat{S}(\theta) = 0$ on $180° < \theta < 360°$ was imposed to replace the cross-shore axis of the array. The results are only marginally different from the full array/no geographic constraint estimates. Evidently, a two-dimensional array is not needed for collecting wave data at this site. A linear array aligned with the coastline together with a no reflection constraint suffices.

6. Discussion

A reviewer suggested a comparison of the estimates we obtained from simulated data to existing techniques. Davis and Regier (1977), Pawka (1983), Oltman-Shay and Guza (1984), Lygre and Krogstad (1986), Marsden and Juszko (1987), and others examined the fidelity of directional spectrum estimation techniques by intercomparing the response to simulated cross spectra of arbitrarily selected target directional distributions $S(\theta)$. Such intercomparisons generally showed the shortcomings of the earlier methods that could produce estimates that were not physically plausible (i.e., negative energy) or were inconsistent with the observations (i.e., improbably large misfits between the estimate and the data). However, the more recent methods all yield nonnegative estimates that are consistent with the data. The ability of these methods to resolve an arbitrary $S(\theta)$ is therefore not a particularly meaningful basis for comparison. Consider, for example, a test with perfect data (i.e., infinite degrees of freedom and error-free sensors) in which two different estimates, $S_1(\theta)$ and $S_2(\theta)$ are obtained that both fit the cross-spectra H_{nm} exactly. Suppose that $S_1(\theta)$ more closely resembles the "true" $S(\theta)$ than $S_2(\theta)$. It might be concluded that the method that produced $S_1(\theta)$ has "superior resolving power." However, since both $S_2(\theta)$ and $S(\theta)$ fit the H_{nm} exactly, replacing the initial input distribution $S(\theta)$ with $S_2(\theta)$ yields the same cross-spectra H_{nm}. Naturally $S_2(\theta)$ now resolves $S(\theta)$ [equal

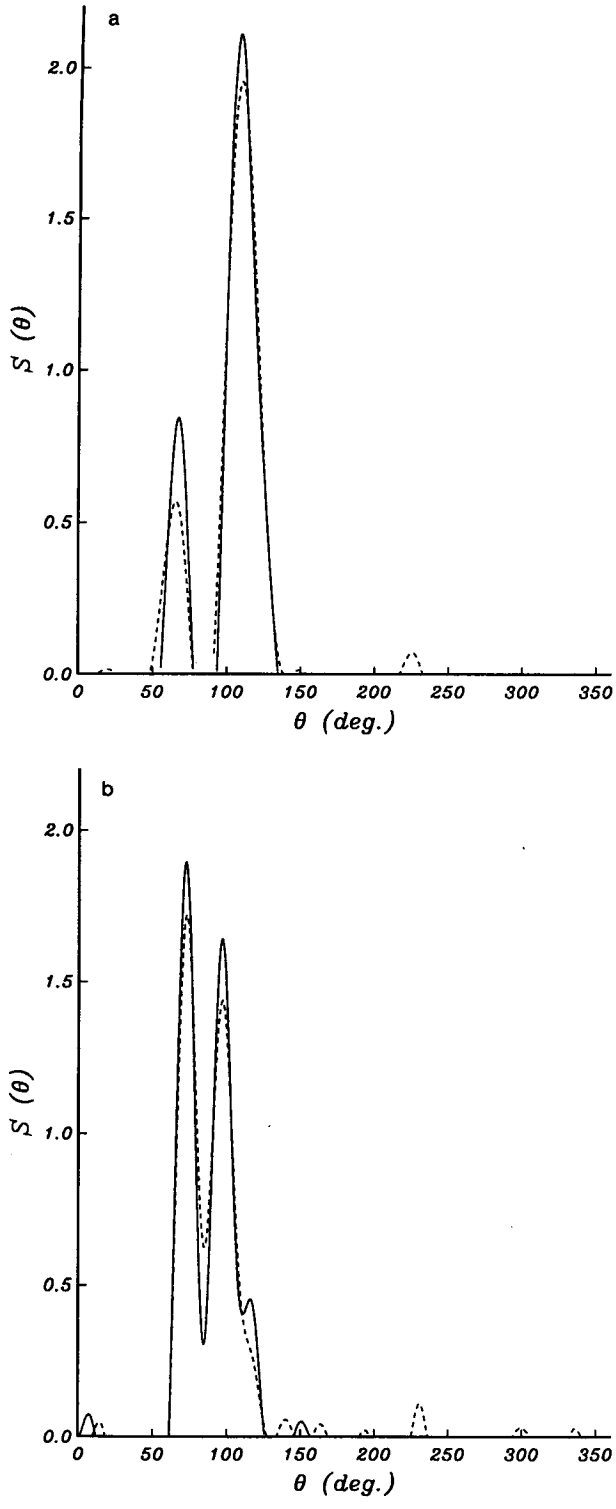


FIG. 17. A comparison of estimates from the 7 pressure sensors on the longshore axis of the Cape Canaveral array (Fig. 13) using a "no reflection constraint" on $\hat{S}(\theta)$(solid lines), and estimates from the full array without a geographic constraint on $\hat{S}(\theta)$(dashed lines, same estimates as in Figs. 15 and 16). Only two frequency bands are shown, other frequency bands show comparable agreement. (a) 4/5 July, $f = 0.132$ Hz, (b) 16/17 July $f = 0.149$ Hz.

to $S_2(\theta)$] perfectly and might be considered a greatly superior estimate. The intercomparisons will always be biased by the choice of the target distribution $S(\theta)$. This subjectivity is particularly obvious for the model fitting technique of Long and Hasselmann (1979) that solves for the estimate that is statistically consistent with the data and closest to an a priori defined preferred model $S_m(\theta)$. Naturally this estimator will perform very well if a target $S(\theta)$ is chosen that resembles $S_m(\theta)$. Intercomparisons of the resolving power of directional spectrum estimation methods are not meaningful if the methods that are compared are optimal in the sense that all available information is absorbed in the estimate. A large discrepancy between an $S(\theta)$ and a physically plausible estimate that is consistent with the data should not be interpreted as a weakness of the estimation technique but as a weakness of the array data. If another method is applied that yields an estimate in closer agreement with $S(\theta)$, it does not mean that this method is superior but rather a false sense of resolution is created by the particular choice of $S(\theta)$.

The present method, based on smoothness optimization, can yield significantly smoother and often markedly different estimates than Long and Hasselmann (1979). This is illustrated by applying both estimators to test arrays and target spectra described in Long and Hasselmann (1979). Two slightly different array geometries were used (Fig. 18): a five-element bottom pressure transducer array actually deployed in the North Sea and a hypothetical six-element array. The input directional distribution has a $\cos^{20}(\theta/2)$ shape "typical of a wind sea near the spectral peak" (Long and Hasselmann 1979). Figure 19 shows the model test results of Long and Hasselmann (1979) for both arrays, with simulated cross-spectra having 30 and 180 degrees of freedom. All estimates are based on an isotropic preferred model $S_m(\theta) = 1/2\pi$. Long and Hasselmann's estimates for $\nu = 180$ (Fig. 19c,d) have a somewhat rough wind sea peak, with spurious side lobes for the five-element array. For noisy ($\nu = 30$)

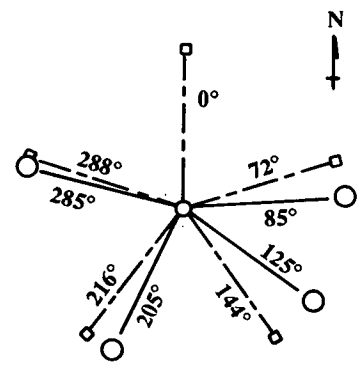


FIG. 18. Geometries of five-element (circles) and six-element (squares) wave gauge arrays used in the model tests of Figs. 19 and 20. The radius of both arrays is 90 m. (from Long and Hasselmann 1979).

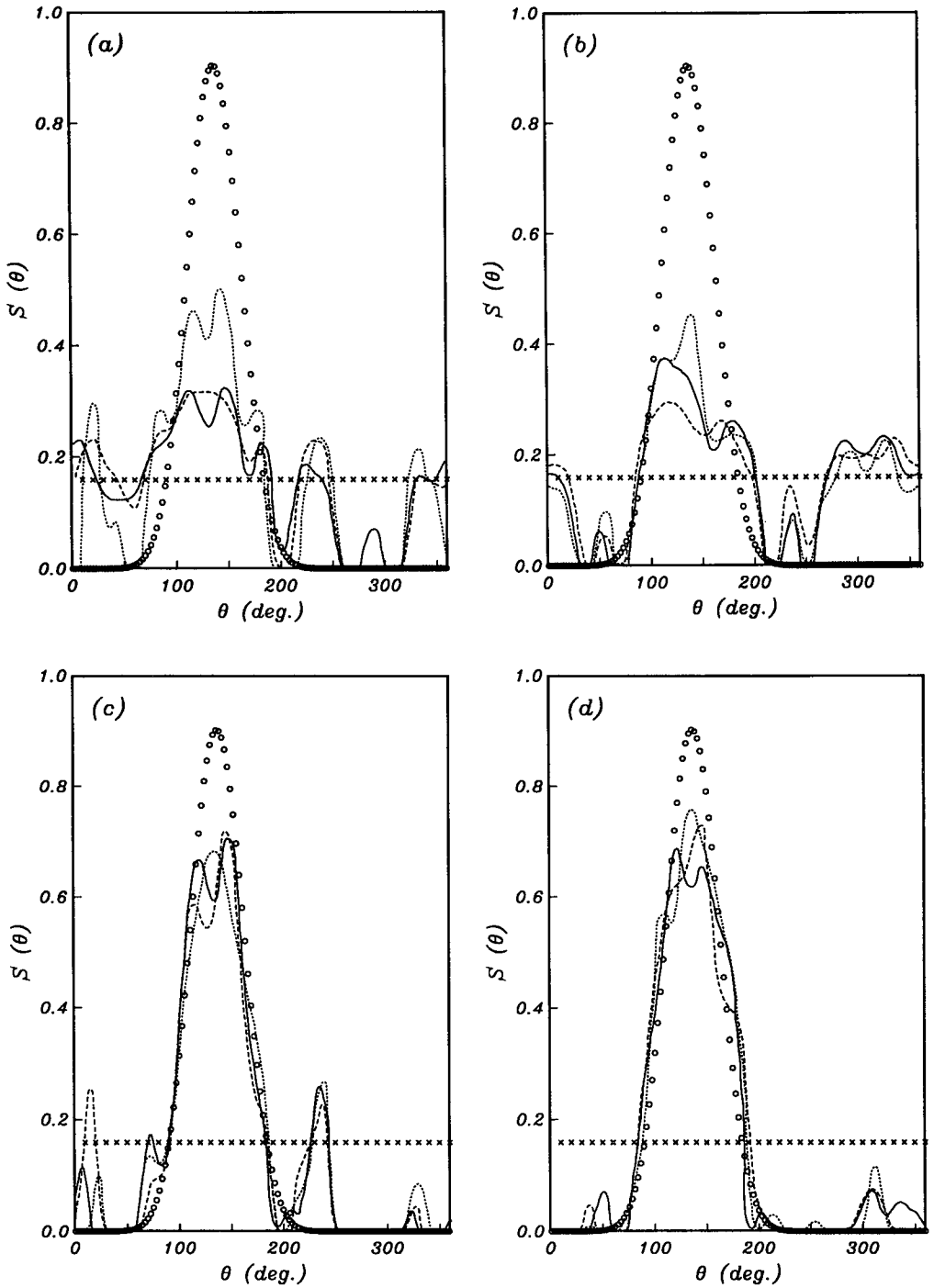


FIG. 19. Estimates of $S(\theta)$ obtained by Long and Hasselmann (1979) in model tests with the arrays of Fig. 18 and a $\cos^{20}[(\theta-137°)/2]$ input directional distribution (O symbols) using an isotropic preferred model (× symbols). The wave length is 100 m and θ is measured clockwise from north. Estimates (solid, dashed, and dotted curves) for three realizations are shown in each panel: (a) five-element array, $\upsilon = 30$; (b) six-element array, $\upsilon = 30$; (c) five-element array, $\upsilon = 180$; (d) six-element array, $\upsilon = 180$. (from Long and Hasselmann 1979)

array data, the fit to the preferred isotropic model more strongly controls the estimates, and the wind sea peak level is reduced at the expense of large side lobes. The present smoothness optimization technique yields estimates (Fig. 20) that are much smoother than Long and Hasselmann's and also do not exhibit the spurious

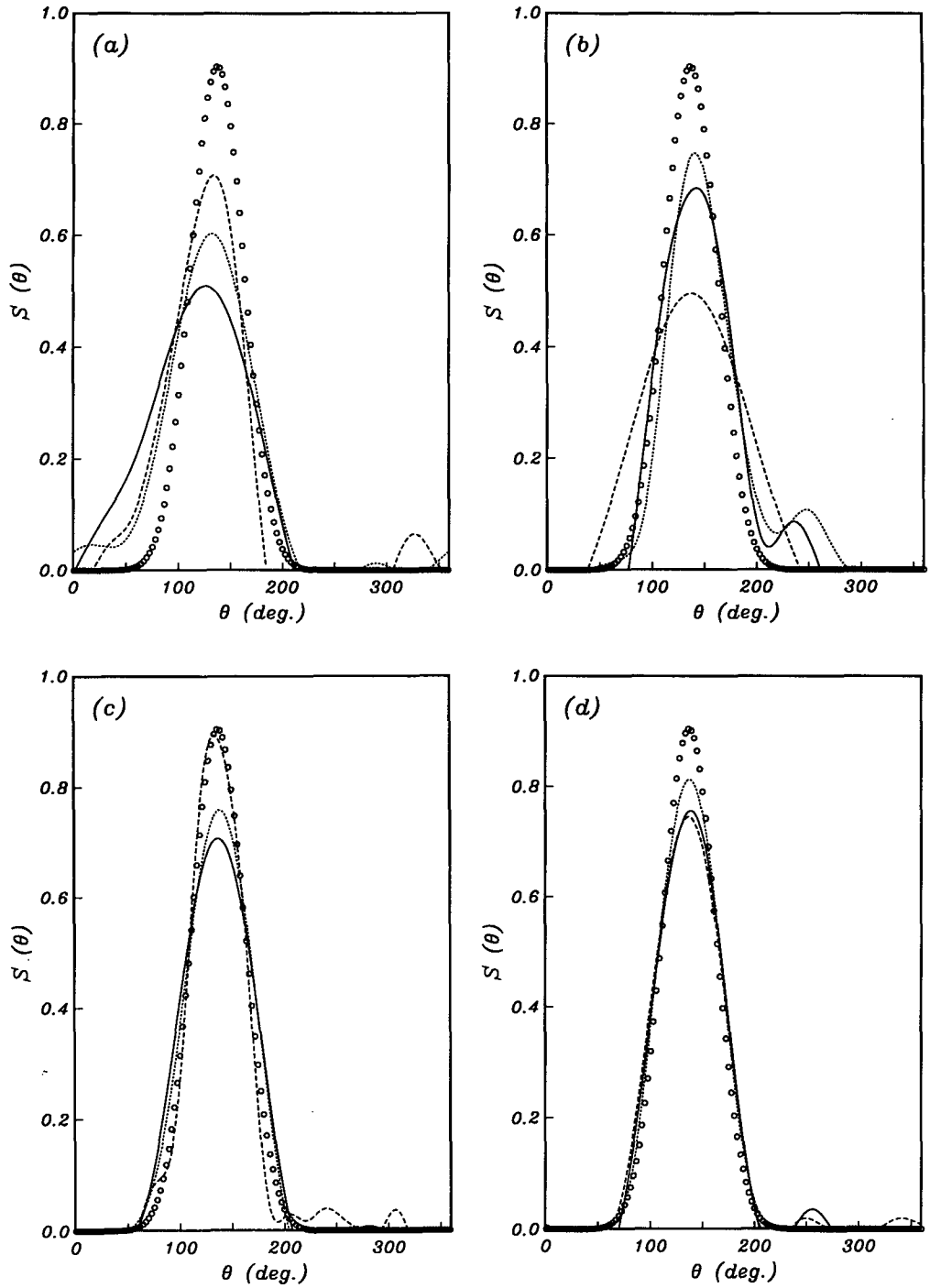


FIG. 20. Estimates of $S(\theta)$ obtained with the present method for the same model tests as Fig. 19 (see caption of Fig. 19 for the details).

side lobes evident in Fig. 19. The smoothness optimization scheme presented here is indeed very effective. Striking too is the fact that the estimates obtained in the present study apparently resolve the input distribution better than those by Long and Hasselmann. For $v = 30$ Long and Hasselmann's estimates show wildly oscillating structure while the present estimates still seem to resolve $S(\theta)$ fairly well. This is a rather misleading artifact of the subjectivity of the test. Even though for $v = 30$ the resolving power of the array data

is very strongly degraded by statistical uncertainty, relatively good agreement is obtained (Fig. 20a,b) because the smooth target $S(\theta)$ resembles the smoothest solution to the inverse problem. However, Fig. 19a,b shows that widely different solutions exist that are also consistent with the noisy data. In the absence of additional information there is simply no way to decide which of these rather different estimates (Fig. 19 or 20) is "better."

The fidelity of physical plausible directional spectrum estimates that are consistent with the observations is not limited by the estimation technique but only by the resolving power of array data. In practice this resolving power is often severely limited because in situ arrays are necessarily sparse, and the statistical uncertainty in the data is often significant because of the limited duration of stationary conditions. The roughness minimization proposed in this study does not improve nor degrade the resolving power even though an intercomparison test with simulated data (e.g., Figs. 19 and 20) may suggest this. Only extra physical (e.g., topographic, Figs. 1, 11) constraints can truly improve the estimates. However, the smoothness optimization principle, even though it is somewhat arbitrary, has the important advantage that it simplifies both the interpretation and the numerical evaluation (relative to other extrema) of the solution. The interpretation of directional spectrum estimates is complicated by the nonuniqueness of solutions to the inverse problem. Methods that do not penalize roughness often produce solutions with spurious peaks. This is particularly evident in the estimates Long and Hasselmann obtained for less than ideal array data (e.g., Fig. 19). It is difficult to discriminate between truly resolved and spurious features in the estimates. Long and Hasselmann (1979, their Fig. 2) inspect different solutions to the inverse problem [i.e., fits to different models $S_m(\theta)$] for common features, to discriminate between true features in North Sea data and spurious ones resulting from the basic limitations of the array data and the choice of $S_m(\theta)$. This very cumbersome procedure can be avoided by seeking the simplest or smoothest solution to the inverse problem. Model tests (e.g., Fig. 20) indicate that the present smoothness optimization technique very effectively rejects peaks not required by the data and that any remaining structure in the estimates is generally shared by the input directional distribution. The numerical evaluation of optimal (non-negative and consistent with the data) estimates of the directional spectrum is far from trivial because of the nonlinear nature of the inverse problem. An extremal solution can generally be obtained only through an iterative procedure, and numerical convergence to this solution is difficult if it is a wildly fluctuating function. The present method is mathematically complex but by virtue of minimizing roughness has the important advantage that smooth well behaved solutions are obtained that are not prohibitively difficult to evaluate numerically. Extensive model tests have confirmed the robustness of the present algorithm even for extreme cases of noisy data (e.g., Fig. 20a,b).

7. Summary and conclusions

An improved method is presented for estimating directional spectra of surface gravity waves from multicomponent data. The method is general in the sense that is can be applied to any array of instruments that measure surface elevation, pressure, velocity and/or other variables at arbitrary but fixed locations in uniform or slowly varying water depth. The measurements are related to surface elevation at a chosen array origin with linear wave theory, using ray theory to account for refraction between the instrument locations if the array is in variable depth water. On a frequency by frequency band basis we seek an estimate of the directional distribution $S(\theta)$ that is consistent with the array observations and any available a priori data-independent information.

Following Long and Hasselmann (1979) we impose the constraint that the misfit between the estimate $\hat{S}(\theta)$ and the array data must be smaller than a chosen confidence level for statistical uncertainty in cross spectra of finite length data records. A rigid constraint that always must be satisfied is the condition that $\hat{S}(\theta)$ is a nonnegative and periodic $[\hat{S}(0) = \hat{S}(2\pi)]$ function with unit integral. In addition we introduce the optional constraint that $\hat{S}(\theta)$ vanishes on an arbitrary number of intervals. This condition can be used to incorporate data-independent geographic information about the experiment site in the inverse model. For example, in the proximity of weakly reflective land masses $S(\theta)$ may be very small at angles directed offshore.

These constraints do not uniquely define $S(\theta)$ but rather a class of possible $S(\theta)$. We seek the "smoothest" member of that class by minimizing a roughness measure of the form $\int d\theta[d^2S(\theta)/d\theta^2]^2$. This choice was motivated by our desire to find the simplest solution to the inverse problem, having no peaks that are not required to explain the observations. Smooth solutions are not only easier to interpret but also have the advantage that they are numerically more readily evaluated than wildly fluctuating ones.

The solution to the inverse problem is derived through a variational formulation with Lagrange multipliers. The resulting nonlinear system of equations is solved iteratively. Extensive model tests have shown that the method indeed yields smooth estimates of $S(\theta)$ without spurious features. The non-negativity constraint renders the inverse problem nonlinear so that numerically complex and time consuming iterations are needed to evaluate the solution. Relaxing this constraint would greatly simplify the computations. However, model tests indicate that it can be a very powerful constraint on $S(\theta)$ and its omission may considerably degrade the estimate.

Geographic constraints on $S(\theta)$ can be used very effectively in conjunction with array data. If instruments fail in a field deployment, the missing lags in the array may seriously degrade its performance. In this case estimates of $S(\theta)$ often have spurious peaks, or worse, spurious peaks only when there is directional ambiguity because of crucial missing lags. Independent geographic constraints may supply the information missing in the degraded array and improve the estimates. Of course, if the geographic constraints can be used with a high degree of confidence, then less information has to be extracted from the array data so that a smaller number of instruments can be deployed in the first place.

Estimates of $S(\theta)$ were obtained from an extensive array of pressure transducers deployed in shallow water offshore of a mildly sloping beach. The preliminary analysis of this experiment yielded smooth unimodal or bimodal directional distributions which are very small at both large angles of incidence and directions corresponding to reflection from the shoreline. These results support the use of topographic constraints in similar coastal environments.

To illustrate the differences between smoothness optimization and model fitting, the present method was applied to test arrays and spectra reported in Long and Hasselmann (1979). The comparison shows that the present method yields smoother estimates without the spurious peaks observed in Long and Hasselmann's (fit to an isotropic directional distribution) estimates. The examples also illustrate the subjectivity of intercomparing estimation techniques. The apparent superiority of the present method to resolve the smooth input $S(\theta)$ merely reflects a bias in the selection of that $S(\theta)$.

Acknowledgments. This work is a result of research sponsored by the Office of Naval Research, under Grant N00014-86-K-0789. The staff of the Center for Coastal Studies (Scripps Institution of Oceanography) installed and maintained the array at Cape Canaveral. J. T. Kirby and his group from the University of Florida, Gainesville obtained the bathymetry and generously assisted us in surveying the instrument locations. We greatly appreciate the efforts of all who contributed to the field experiment. We thank R. B. Long and the reviewers for helpful comments and suggestions. Joan Semler expertly typed the manuscript.

APPENDIX A

Wave Propagation in the Array Vicinity

In deep or uniform water depth the spatial evolution of wave direction, amplitude, and phase (Eq. 6) are given by

$$\theta(\mathbf{x}, \theta_0) = \theta_0 \tag{A1a}$$

$$A(\mathbf{x}, \theta_0) = 1 \tag{A1b}$$

$$B(\mathbf{x}, \theta_0) = k(x\cos\theta_0 + y\sin\theta_0). \tag{A1c}$$

In shallow or intermediate variable depth water the functions θ, A and B can, in principle, be evaluated with a numerical wave propagation model (e.g., Izumiya and Horikawa 1987). Often, however, directional wave data is collected on beaches with small depth variations in the longshore (x) direction. If we neglect longshore depth variations ($h = h(y)$) then ray theory yields simple expressions for θ, A, and B:

$$\theta(\mathbf{x}, \theta_0) = \begin{cases} \arccos[f(y, \theta_0)] & \text{for } 0 \leqslant \theta_0 < \pi \\ 2\pi - \arccos[f(y, \theta_0)] & \text{for } \pi \leqslant \theta_0 < 2\pi \end{cases} \tag{A2a}$$

$$A(\mathbf{x}, \theta_0) = \left[\frac{C_{g0}}{C_g(y)}\right]^{1/2}\left[\frac{\sin^2\theta_0}{1 - f^2(y, \theta_0)}\right]^{1/4} \tag{A2b}$$

$$B(\mathbf{x}, \theta_0) = k_0 x\cos\theta_0 + \sin\theta_0 \int_0^y dy' k(y')\left[\frac{1 - f^2(y', \theta_0)}{\sin^2\theta_0}\right]^{1/2} \tag{A2c}$$

where $k(y)$ is given by Eq. (4), $C_g(y)$ is the group speed $\sigma/(2k(y)) + \sigma h(y)/\sinh[2k(y)h(y)]$ and $f(y, \theta_0) = k_0\cos\theta_0/k(y)$. All variables with subscript 0 are evaluated at x = 0. A necessary condition for the validity of Eq. (A2) is that there are no wave components propagating at very large angles relative to the beach normal. We do not consider the pathological case when $|f(y, \theta_0)| = 1$ (i.e., $\theta(\mathbf{x}, \theta_0) = 0$ or π) and a caustic crosses the array causing a singularity in $A(\mathbf{x}, \theta_0)$. We assume that $S(\theta_0)$ is negligibly small at angles θ_0 with appreciable variations of $A(\mathbf{x}, \theta_0)$ on wavelength scales and replace the integral in Eq. (6) by

$$\int_{\theta_m}^{\pi-\theta_m} + \int_{\pi+\theta_m}^{2\pi-\theta_m} \tag{A3}$$

with θ_m an upper bound on the obliquity of wave approach.

APPENDIX B

Geographic Constraints

It is straightforward to add a data independent geographic constraint of the form Eq. (44) to the inverse problem. With this additional constraint, Eqs. (28) and (29) can be written

$$\hat{\boldsymbol{\epsilon}} = \tilde{\mathbf{d}} - \sum_{m=1}^{M} \int_{\phi_m{}^1}^{\phi_m{}^2} d\theta \mathbf{b}(\theta) F(\theta) H(\theta) \tag{B1}$$

$$\sum_{m=1}^{M} \int_{\phi_m{}^1}^{\phi_m{}^2} d\theta F(\theta) H(\theta) = 1. \tag{B2}$$

We impose the smoothness conditions

$$\frac{d^2}{d\theta^2}F(\theta) \quad \text{is bounded everywhere} \tag{B3}$$

$$\frac{d^4}{d\theta^4}F(\theta) \quad \text{is bounded on} \quad \phi_m{}^1 < \theta < \phi_m{}^2$$

$$m = 1, 2, \cdots, M. \tag{B4}$$

We minimize

$$\phi \equiv R + \alpha\hat{\epsilon}^{\mathrm{T}}\tilde{\mathbf{V}}^{-1}\hat{\epsilon} + \beta \sum_{m=1}^{M} \int_{\phi_m{}^1}^{\phi_m{}^2} d\theta F(\theta)H(\theta) \tag{B5}$$

where R is the roughness measure

$$R \equiv \sum_{m=1}^{M} \int_{\phi_m{}^1}^{\phi_m{}^2} d\theta \left[\frac{d^2}{d\theta^2}F(\theta)\right]^2 \tag{B6}$$

and α and β are the Lagrange multipliers for constraints (23) and (B2), respectively.

Setting $\delta\phi = 0$ yields the equation

$$\frac{d^4}{d\theta^4}F(\theta) = [\alpha\hat{\epsilon}^{\mathrm{T}}\tilde{\mathbf{V}}^{-1}\mathbf{b}(\theta) - \beta/2]H(\theta). \tag{B7}$$

The solution $F(\theta)$ can be obtained in the same way Eq. (41) was derived for the "data only" case. Tedious but straightforward integrations yield

$$F(\theta) = \boldsymbol{\lambda}^{\mathrm{T}}\{\mathbf{B}_4(\theta) - \mathbf{w}H_4(\theta)\} + \frac{H_4(\theta)}{\int_0^{2\pi} d\theta H_4(\theta)H(\theta)} \tag{B8}$$

$$[\mathbf{C}_2 + \alpha^{-1}\tilde{\mathbf{V}}]\boldsymbol{\lambda} = \mathbf{D} \tag{B9}$$

where $\mathbf{C}_2$ and $\mathbf{D}$ are now defined as

$$\mathbf{C}_2 \equiv \int_0^{2\pi} d\theta[\mathbf{B}_2(\theta) - \mathbf{w}H_2(\theta)][\mathbf{B}_2(\theta) - \mathbf{w}H_2(\theta)]^{\mathrm{T}} \tag{B10}$$

$$\mathbf{D} \equiv \tilde{\mathbf{d}} - \mathbf{w} \tag{B11}$$

with

$$\mathbf{w} \equiv \frac{\int_0^{2\pi} d\theta \mathbf{B}_4(\theta)H(\theta)}{\int_0^{2\pi} d\theta H_4(\theta)H(\theta)}. \tag{B12}$$

On each interval $(\phi_m{}^1, \phi_m{}^2)$, the functions H_2, $\mathbf{B}_2$, H_4 and $\mathbf{B}_4$ are given by

$$\begin{bmatrix} H_2(\theta) \\ \mathbf{B}_2(\theta) \end{bmatrix} = \begin{bmatrix} h_2(\theta) \\ \mathbf{b}_2(\theta) \end{bmatrix} + \begin{bmatrix} p \\ \mathbf{q} \end{bmatrix}(\theta - \phi_m{}^1) + \begin{bmatrix} r \\ \mathbf{s} \end{bmatrix}$$

$$\begin{bmatrix} H_4(\theta) \\ \mathbf{B}_4(\theta) \end{bmatrix} = \begin{bmatrix} h_4(\theta) \\ \mathbf{b}_4(\theta) \end{bmatrix} + \begin{bmatrix} p \\ \mathbf{q} \end{bmatrix}\frac{(\theta - \phi_m{}^1)^3}{6} + \begin{bmatrix} r \\ \mathbf{s} \end{bmatrix}\frac{(\theta - \phi_m'{}^1)^2}{2} \tag{B13}$$

where $h_n(\theta)$, $\mathbf{b}_n(\theta)$ are the functions

$$\begin{bmatrix} h_n(\theta) \\ \mathbf{b}_n(\theta) \end{bmatrix} = \int_{\phi_m{}^1}^{\theta} d\theta^{(1)} \int_{\phi_m{}^1}^{\theta^{(1)}} d\theta^{(2)} \cdots \int_{\phi_m{}^1}^{\theta^{(n-1)}} d\theta^{(n)} H(\theta^{(n)}) \begin{bmatrix} 1 \\ \mathbf{b}(\theta^{(n)}) \end{bmatrix} \tag{B14}$$

and p, $\mathbf{q}$, r and $\mathbf{s}$ are the integration constants

$$\begin{bmatrix} p \\ \mathbf{q} \end{bmatrix} = \frac{-6}{(\phi_m{}^2 - \phi_m{}^1)^2}\begin{bmatrix} h_3(\phi_m{}^2) \\ \mathbf{b}_3(\phi_m{}^2) \end{bmatrix} + \frac{12}{(\phi_m{}^2 - \phi_m{}^1)^3}\begin{bmatrix} h_4(\phi_m{}^2) \\ \mathbf{b}_4(\phi_m{}^2) \end{bmatrix}$$

$$\begin{bmatrix} r \\ \mathbf{s} \end{bmatrix} = \frac{2}{(\phi_m{}^2 - \phi_m{}^1)}\begin{bmatrix} h_3(\phi_m{}^2) \\ \mathbf{b}_3(\phi_m{}^2) \end{bmatrix} - \frac{6}{(\phi_m{}^2 - \phi_m{}^1)^2}\begin{bmatrix} h_4(\phi_m{}^2) \\ \mathbf{b}_4(\phi_m{}^2) \end{bmatrix}. \tag{B15}$$

At angles θ that do not fall within any of the intervals $(\phi_m{}^1, \phi_m{}^2)$, $m = 1, 2, \cdots, M$, the functions $H_2(\theta)$, $\mathbf{B}_2(\theta)$, $H_4(\theta)$ and $\mathbf{B}_4(\theta)$ vanish.

REFERENCES

Backus, G. E., and J. F. Gilbert, 1967: Numerical applications of a formalism for geophysical inverse problems. *Geophys. J. Roy. Astr. Soc.*, **13**, 247–276.

Barber, N. F., 1963: The directional resolving power of an array of wave detectors. *Ocean Wave Spectra*, Prentice-Hall, 137–150.

Constable, S. C., R. L. Parker and C. G. Constable, 1987: Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**(3), 289–300.

Davis, R. E., and L. A. Regier, 1977: Methods for estimating directional wave spectra from multi-element arrays. *J. Mar. Res.*, **35**(3), 453–477.

Donelan, M. A., J. Hamilton and W. H. Hui, 1985: Directional spectra of wind-generated waves. *Phil. Trans. Roy. Soc. London*, **A315**, 509–562.

Freilich, M. H., and R. T. Guza, 1984: Nonlinear effects on shoaling surface gravity waves. *Phil. Trans. Roy. Soc. London*, **A311**, 1–41.

Hasselmann, D. E., M. Dunckel and J. A. Ewing, 1980: Directional wave spectra observed during JONSWAP 1973. *J. Phys. Oceanogr.*, **10**, 1264–1280.

Herbers, T. H. C., and R. T. Guza, 1989: Estimation of wave radiation stresses from slope array data. *J. Geophys. Res.*, **94**(C2), 2099–2104.

Izumiya, T., and K. Horikawa, 1987: On the transformation of directional random waves under combined refraction and diffraction. *Coastal Eng. in Japan,* **30**(1), 49–65.

Jenkins, G. M., and D. G. Watts, 1968: *Spectra Analysis and Its Applications,* Holden-Day, 525 pp.

Kirby, J. T., 1988: Survey data report: Cape Canaveral, March – July, 1988, UFL/COEL-88/011, 68 pp.

Lawson, L. M., and R. B. Long, 1983: Multimodal properties of the surface-wave field observed with pitch-roll buoys during GATE. *J. Phys. Oceanogr.,* **13**(3), 474–486.

Long, R. B., 1980: The statistical evaluation of directional spectrum estimates derived from pitch/roll buoy data. *J. Phys. Oceanogr.,* **10**(6), 944–952.

——, 1986: Inverse modeling in ocean wave studies. *Wave Dynamics and Radio Probing of the Ocean Surface,* O. M. Phillips, K. Hasselmann, Eds., Plenum, 571–593.

——, and K. Hasselmann, 1979: A variational technique for extracting directional spectra from multi-component wave data. *J. Phys. Oceanogr.,* **9**(2), 373–381.

Longuet-Higgins, M. S., D. E. Cartwright and N. D. Smith, 1963: Observations of the directional spectrum of sea waves using the motions of a floating buoy. *Ocean Wave Spectra,* Prentice-Hall, 111–136.

Lygre, A., and H. E. Krogstad, 1986: Maximum entropy estimation of the directional distribution in ocean wave spectra. *J. Phys. Oceanogr.,* **16**(12), 2052–2060.

Marsden, R., F., and B.-A. Juszko, 1987: An eigenvector method for the calculation of directional spectra from heave, pitch and roll buoy data. *J. Phys. Oceanogr.,* **17**(12), 2157–2167.

Oltman-Shay, J., and R. T. Guza, 1984: A data-adaptive ocean wave directional-spectrum estimator for pitch and roll type measurements. *J. Phys. Oceanogr.,* **14**(11), 1800–1810.

Parker, R. L., 1977: Understanding inverse theory. *Ann. Rev. Earth Planet Sci.,* **5,** 35–64.

Pawka, S. S., 1983: Island shadows in wave directional spectra. *J. Geophys. Res.,* **88**(C4), 2579–2591.

Provost, C., and R. Salmon, 1986: A variational method for inverting hydrographic data. *J. Mar. Res.,* **44,** 1–34.