

A technique for objective analysis and design of oceanographic experiments applied to MODE-73*

FRANCIS P. BRETHERTON[†], RUSS E. DAVIS[‡] and C. B. FANDRY[§]

(Received 21 April 1975; in revised form 3 November 1975; accepted 3 November 1975)

Abstract—A technique for the objective analysis of oceanic data has been developed and used on simulated data. The technique is based on a standard statistical result—the Gauss–Markov Theorem—which gives an expression for the least square error linear estimate of some physical variable (velocity, stream function, temperature, etc.) given measurements at a limited number of data points, the statistics of the field being estimated in the form of space–time spectra, and the measurement errors. An expression for the r.m.s. error expected in this estimate is also derived and illustrated in the form of ‘error maps’.

Efficient sampling arrays can be designed through trial-and-error adjustment of array configurations until a suitable balance of mapping coverage and accuracy, as measured by the error maps, is achieved. Examples of the mapping ability of some simple arrays are given.

Using statistics inferred from the preliminary Mid Ocean Dynamics Experiments various realizations of likely flow fields were simulated. The 16 element MODE-I array was tested by comparison of the simulated fields and the objective maps based on inferred ‘measurements’ at the array points. The reliability of statistics inferred from observations was estimated by comparing correlations derived from limited observations of the simulated fields with the known statistics. Correlations derived from two realizations differed significantly but most calculations reproduced the known statistics moderately well.

An intercomparison of Eulerian measurements (current meters) and Lagrangian measurements (neutrally buoyant drifters) was also carried out using the objective interpolation method.

1. INTRODUCTION

IN RECENT years there has been an increasing effort to determine the spatial structure of time-dependent velocity and density fields in the deep ocean. In addition to the ongoing programme of tracking meanders in the Gulf Stream, extensive networks of hydrographic stations have been analysed by a number of workers. For example, SAUNDERS (1972) investigated the hydrography in a 5° square in the Mediterranean, and ELLIOTT, HOWE and TAIT (1974) analysed stations in the northwest Atlantic. During the POLYGON experiment (BREKHOVSKIKH, FEDOROV, FOMIN, KOSHLYAKOV and YAMPOLSKY, 1971), scientists from the Soviet Union deployed a large array of current meters in the subtropical Atlantic near 18°N 55°W. A total of 17 moorings spaced along a cross with arms 210 km long with current meters at various levels

in the upper 1500 m was maintained continuously for 7 months, during which period at least one major closed eddy was observed to move westwards through the area.

The Mid Ocean Dynamics Experiment (MODE) was a combined programme of observation and theory involving scientists from many different institutions in the U.S.A. and Great Britain. The objective is an intensive study of the mesoscale motion on space scales of 50 to 500 km and time scales of 1 week to a few months.

*Mid Ocean Dynamics Experiment (MODE) Contribution No. 38.

[†]National Center for Atmospheric Research, Boulder, Colorado, U.S.A.

[‡]Scripps Institution of Oceanography, La Jolla, California, U.S.A.

[§]Department of Mathematics, Monash University, Clayton, Victoria, Australia, 3168.

MODE-I involved deployment of more than 100 current meters on 23 moorings within a circle of radius 200 km centred on 28°N 69°40'W in the Sargasso Sea for a period of about 100 days. A hydrographic programme comprising some 700 CTD and STD soundings was carried out at the same time together with the tracking of 18 long-range SOFAR floats and some 30 short-range floats. Airdrop probes measuring the vertical average of the horizontal currents, moored temperature sensors, bottom-mounted pressure gauges, inverted echo sounders, and electromagnetic current meters were also involved, as well as free-fall instruments measuring detailed current profiles.

Measurement programmes of this magnitude require substantial resources and it is desirable to consider carefully the design of the instrument arrays and the procedure for data analysis. The study reported here arose as part of the planning process for MODE-I, but it is believed to be of sufficient general interest to warrant separate description. The conclusions provided some of the criteria according to which the main field programme was designed. The latter will not be discussed here in detail. The emphasis is rather on a theoretical exercise in techniques of data analysis and array design. The choice of examples is coloured by the programme requirements but the principles are of wider application.

The original problem addressed was how to interpolate horizontally between data points from an irregular two-dimensional array of some 16 current meters spaced 20 to 100 km apart, in order to construct a synoptic map of the velocity field. Given also the desire to ascertain the overall pattern of the flow over a wide area, and the logistic difficulties of deploying additional meters, the array spacing has to be as large as possible consistent with reasonable accuracy in the interpolation and some redundancy to cover instrument failure. These conflicting requirements imply that the data available must be expected to be barely adequate for the task at hand, and considerable effort is justified to maximize the information obtained from it.

The primary interest is in the low-frequency

currents—operationally defined as the output of a Gaussian filter applied to the time series of velocity components $[u(t), v(t)]$ at fixed points, yielding a 3-day running average velocity which is then resampled daily. This pre-processing greatly reduces the inertial oscillations and internal waves in the record and increases the horizontal scale over which the velocities are correlated. A possible analysis procedure would be to draw, by hand, contour maps for each component $u(x, y), v(x, y)$ separately at any given time t , smoothing to whatever degree appeared appropriate to the internal consistency of the data. However, theoretical considerations strongly suggest that the low-frequency currents should be in approximate geostrophic balance, and horizontally non-divergent to an accuracy of a few per cent, i.e. there exists a stream function $\psi(x, y, t)$ such that

$$u = -\frac{\partial\psi}{\partial y}, \quad v = \frac{\partial\psi}{\partial x}. \quad (1)$$

It would be very desirable to measure the divergence $\partial u/\partial x + \partial v/\partial y$ and compare it to the magnitude of the vorticity $\partial v/\partial x - \partial u/\partial y$, but unless the theoretical expectations are grossly in error, observational errors were clearly too large for a decisive test. Accordingly, unless there is a clear discrepancy, it seems preferable to raise equation (1) to the level of an axiom of the analysis, and to estimate one scalar field $\psi(x, y)$ from 32 pieces of information, rather than attempting to patch together two somewhat inconsistent estimates for u, v , each obtained separately from only 16 pieces of information.

A few preliminary trials showed clearly the advantage of using a computer to aid in this task. A human analyst finds it difficult to take cognizance simultaneously of both the direction and the magnitude of the gradient of ψ , and a more objective technique is desirable. There is, of course, no unique way of interpolating between a limited number of data points, and every analysis implies assumptions about the character or degree of smoothness of the field being analysed.

The method suggested here has the advantage that it is based upon certain explicitly stated (and

testable) statistical assumptions. Granted these assumptions, it yields at every point (x, y) an estimate $\hat{\psi}$ which is optimal, in the sense that of all estimates, which depend linearly on the data supplied, on the average this one has the least error. Furthermore, this minimum mean square error can be calculated *a priori* for different array designs, and the latter may be adjusted so that the expected interpolation errors do not exceed required tolerances. At the same time, estimates for the velocity components $(-\partial\hat{\psi}/\partial y, \partial\hat{\psi}/\partial x)$ are also optimized in the same sense, and are automatically nondivergent.

The second problem given detailed consideration concerns space-time interpolation of the depths of a given isotherm, measured non-synoptically from a ship criss-crossing the area in an irregular manner. This is a scalar field from a scalar measurement, unrestricted by considerations of nondivergence, yet complicated by the time dependence of the pattern being mapped. Reliable updated extrapolations of such fields can be of value operationally during a cruise when investigating special features of a flow pattern or when deciding where to make supplementary measurements, provided they are available in time and the probable error bounds are known. A modification of the same method was used on a trial basis during the MODE-I field programme.

The approach is based on the Gauss-Markov theorem—a standard result in statistical estimation theory (LIEBELT, 1967). It has been applied in a manner similar to that described here by meteorologists in the Soviet Union (GANDIN, 1965) for the objective analysis of pressure and wind fields in the atmosphere, and is routinely used in the preparation of numerical weather prediction. In the oceanic context it should still be considered experimental—a supplement rather than a replacement of careful hand analysis which, using the experience of the analyst, probes all of the aspects of the data.

The approach was used to design sampling arrays for MODE-I. This is possible because the method is based on minimization of the expected interpolation error and results in a map of this error which depends on the statistics of the field

and not the measurements themselves. Through trial-and-error adjustments of array configurations it is possible to seek out an array which strikes an appropriate balance between coverage and accuracy. An essential aspect of this design procedure is a precise definition of the quantity to be mapped. It is obvious that the array design depends on the scale of the phenomena one is hoping to measure. Examples of the ability of simple arrays to map various quantities were estimated from the associated error maps.

A primary objective of this study was to test the above analysis procedures and the MODE-I array design on simulated data for which the precise statistics are known. These statistics were chosen to be consistent with preliminary observations in the MODE area (Arrays I and III) in the manner described in Section 4. A number of realizations of randomly generated Gaussian fields with the appropriate space-time spectra were used to construct time series of velocity vectors at the projected locations, random noise was added to simulate instrumental and sampling errors, and the resulting 'data' were fed to the analysis routines. The reconstructed stream function was then compared with the original from which the velocities were derived. 'Float tracks' obtained by integrating the Lagrangian equations

$$\frac{dx}{dt} = u(x, y, t); \quad \frac{dy}{dt} = v(x, y, t) \quad (2)$$

$$x = x_0, y = y_0 \text{ at } t = 0,$$

using both the original and objectively interpolated velocity fields (u, v) , were compared. While the potential performance of a particular array can be evaluated using maps of the r.m.s. error it may also be useful to compare directly the simulated and reconstructed fields.

An important element of the method is that good estimates of the statistics of the field must be known. Variations between estimated space-time correlations obtained from limited sampling of the simulated fields provided a reminder of the dangers of drawing statistical inferences from too

little data. Improved methods of estimating the basic statistics will be discussed elsewhere, but no refinement of analysis will eliminate the need for caution tempered by judgement in the application of these methods to the designs of major programmes.

One final point of interest concerns the use of objective maps to test dynamical hypotheses. It is interesting to learn that, for the special case of linear dynamics, the verification of models can be carried out equally well through comparison of objective maps of the data and model 'hindcasts' or through comparison of statistics estimated from data and model simulations.

2. BASIC THEORY

2.1 Simple scalar field

For simplicity we present first the objective analysis procedure for a rather special case—estimating the value $\theta_{\mathbf{x}}$ of a scalar variable $\theta(x, y)$ at a general point $\mathbf{x} = (x, y)$ from measurements φ_r at a limited number of data points \mathbf{x}_r , ($r = 1, \dots, N$). The central assumption is that the field θ is one realization out of a homogeneous statistical ensemble of zero mean and known covariance function

$$\overline{\theta_{\mathbf{x}}} = 0 \tag{3}$$

$$\overline{\theta_{\mathbf{x}} \theta_{\mathbf{x} + \xi}} = F(\xi) = \int \mathbf{f}(\boldsymbol{\kappa}) e^{i\boldsymbol{\kappa} \cdot \xi} d\boldsymbol{\kappa}, \tag{4}$$

where $\mathbf{f}(\boldsymbol{\kappa})$ is the wave number spectrum. The requirement that the mean be zero is not critical as will be shown later. In practice the two point covariance function and the mean are estimated from a finite sample of realization of θ combined with some *a priori* prejudices. We will return later to how this may be done. We suppose also that a measured value φ_r is the true point value plus some random noise (which may be local sampling error or instrumental uncertainty)

$$\varphi_r = \theta(\mathbf{x}_r) + \varepsilon_r, \quad (r = 1, \dots, N), \tag{5}$$

where the errors ε_r are uncorrelated with one

another and with the field θ but have known variance, E :

$$\begin{aligned} \overline{\varepsilon_r \varphi_s} &\equiv 0 \\ \overline{\varepsilon_r \varepsilon_s} &= E \delta_{rs} \quad (r, s = 1, \dots, N). \end{aligned} \tag{6}$$

Systematic or calibration errors are thus not permitted here.

Then the Gauss–Markov theorem states that the least squares optimum linear estimator for $\theta(\mathbf{x}, y)$ is

$$\hat{\theta}_{\mathbf{x}} = \sum_{r=1}^N C_{\mathbf{x}r} \left(\sum_{s=1}^N A_{rs}^{-1} \varphi_s \right), \tag{7}$$

where

$$A_{rs} = \overline{\varphi_r \varphi_s} = F(\mathbf{x}_r - \mathbf{x}_s) + E \delta_{rs}, \tag{8}$$

is the matrix of covariance between all pairs of observations, and A^{-1} is the inverse matrix of A .

$$C_{\mathbf{x}r} = \overline{\theta_{\mathbf{x}} \varphi_r} = F(\mathbf{x} - \mathbf{x}_r) \tag{9}$$

is the covariance between the quantity $\theta_{\mathbf{x}}$ to be estimated and the r th measurement. For given positions of the observation points, the matrices A_{rs}^{-1} , $C_{\mathbf{x}r}$ are constants. Thus for different realizations of the field $\theta(\mathbf{x})$, the estimate $\hat{\theta}_{\mathbf{x}}$ depends linearly on the observations φ_s . Hence $\hat{\theta}_{\mathbf{x}}$ is a linear estimator.

Confidence levels in this estimate are also available. The variance of the error in $\hat{\theta}_{\mathbf{x}}$ is

$$(\theta_{\mathbf{x}} - \hat{\theta}_{\mathbf{x}})^2 = C_{\mathbf{x} \mathbf{x}} - \sum_{r, s=1}^N C_{\mathbf{x}r} C_{\mathbf{x}s} A_{rs}^{-1}. \tag{10}$$

Equations (7) and (10) will be derived later in this section. The first term in (10) is the natural variation in θ in the absence of any data. The second term measures the information provided by the data. An important tool in the evaluation of a particular array design is the construction of maps of probable errors in the estimates to be made. Equation (10) shows that such a calculation requires only the location of the data points and a

knowledge of the covariance function $F(\xi)$ and noise level E . Individual sets of data are not needed. Some elegant matrix algebra also provided formulae for the decreased error associated with inserting extra data points, but for this the reader should consult LIEBELT (1967) and GANDIN (1965). Roughly speaking, if the variance given by equation (10) at a given point is already small, the insertion of an extra measurement there will not improve the overall analysis very much.

Given the spectrum $\mathcal{J}(\kappa)$ and the noise level E the procedure is quite straightforward for computing an optimal map $\hat{\theta}_x$ from a limited number of synoptic data points. The matrix A is positive definite and may readily be inverted and stored. For each set of observations $\{\varphi_r\}$ the weights

$$\eta_r = \sum_{s=1}^N A_{rs}^{-1} \varphi_s \quad (11)$$

follow immediately. For each map point x and observation r the correlation C_{xr} is found by interpolation in a precomputed table from a functional representation of $F(\xi)$, and the estimate is formed by summation of

$$\hat{\theta}_x = \sum C_{xr} \eta_r \quad (12)$$

When the number of data becomes too large the inversion of the full A matrix is impractical and it becomes necessary to map θ over segments of x using subsets of the data which are most pertinent. This involves a difficult sorting process based on the covariance C_{xr} and is to be avoided if possible.

To derive equation (7), we consider a quite general linear estimator

$$\hat{\theta}_x = \sum_{s=1}^N \alpha_{xs} \varphi_s \quad (13)$$

where the coefficients α_{xs} are still to be determined. Considering a large number of different realizations of the field $\theta(x, y)$ and observations φ_s , the error variance for $\hat{\theta}$ is

$$\begin{aligned} \overline{(\theta_x - \hat{\theta}_x)^2} &= \overline{(\theta_x - \sum_{s=1}^N \alpha_{xs} \varphi_s)^2} \\ &= C_{xx} - 2 \sum_s \alpha_{xs} C_{xs} + \sum_{r,s} \alpha_{xr} \alpha_{xs} A_{rs} \\ &= \sum_{r,s} (\alpha_{xr} - \sum_{r'} C_{xr'} A_{r'r}^{-1}) (\alpha_{xs} - \sum_{s'} C_{xs'} A_{s's}^{-1}) A_{rs} \\ &\quad + C_{xx} - \sum_{r,s=1} C_{xr} C_{xs} A_{rs}^{-1}. \end{aligned} \quad (14)$$

The values of the coefficients α_{xr} which minimize this variance are precisely those used in equation (7), and the minimum which remains is given by expression (10). The positions of the observations x_r and the analysis point x are, of course, held constant with this process. Given the statistics of the field being measured and the noise levels involved, no other analysis procedure could perform better. This is the Gauss-Markov theorem, a very simple and powerful result.

The remainder of Section 2 is concerned with various details involved in adapting the central algorithm to specific cases of interest. The discussion is oriented toward the use of synoptic data but if such data are not available, so that interpolation in both space and time is required, the formulation remains the same except that the position variable x must be interpreted as a pseudo-vector determining position in both space and time. The correlation $F(\xi)$ is, in this case, a space-time covariance. This is the procedure used for the second problem described in the Introduction.

2.2 Gross errors

In general, the optimal estimate at a data point does not precisely correspond to the measurement there. We have

$$\begin{aligned} \hat{\theta}_r &= \sum_{p=1}^N C_{rp} \left\{ \sum_{s=1}^N A_{ps}^{-1} \varphi_s \right\} \\ &= \sum_{p=1}^N (A_{rp} - E\delta_{rp}) \left\{ \sum_{s=1}^N A_{ps}^{-1} \varphi_s \right\} \\ &= \varphi_r - E\eta_r. \end{aligned} \quad (15)$$

The discrepancy $-E\eta_r$ is an estimate of the random error on this particular observation φ_r , obtained by comparison with surrounding observations and the statistics for this data point. The standard deviation of this discrepancy is

$$\begin{aligned} \overline{E\{\eta_r^2\}}^{1/2} &= E\left\{\sum_{s,s'} A_{rs}^{-1} A_{rs'}^{-1} \overline{\varphi_s \varphi_{s'}}\right\}^{1/2} \\ &= E\{A_{rr}^{-1}\}^{1/2}. \end{aligned} \tag{16}$$

The ratio of the discrepancy, $E\eta_r$, and the standard deviation of this discrepancy is

$$\lambda_r = \eta_r / \{A_{rr}^{-1}\}^{1/2}, \tag{17}$$

and is a convenient indicator whether a particular observation φ_r is grossly incorrect, e.g. because of an instrument malfunction or transcription error not included in the original tolerances. If it exceeds about 3, there is a *prima facie* case for rejecting the observation from the analysis. Of course, this approach will work only if the frequency of gross errors is small, and if the statistical tolerances have indeed been accurately estimated.

2.3 Estimated mean

For some fields the spectrum $\mathcal{F}(\kappa)$ is so red that the integral over κ diverges near $\kappa = 0$. In that case the covariance function $F(\xi)$ is not well determined, though the structure function

$$G(\xi) = \overline{(\theta_{x+\xi} - \theta_x)^2} \tag{18}$$

is. In effect the spatial mean $\bar{\theta}$ in the vicinity of the data points is unknown, so that it cannot serve as a base for the covariance function

$$\overline{(\theta_x - \bar{\theta})(\theta_{x+\xi} - \bar{\theta})}.$$

Thus a large unknown constant H has been added to $F(\xi)$, where the latter is now restricted to separations no larger than the array size. The structure function

$$G(\xi) = 2[F(0) - F(\xi)] \tag{19}$$

does not depend on H . The above theory may be modified to allow for this contingency by formally including the constant H , and rearranging the expressions so that they are finite in the limit $H \rightarrow \infty$, and adopting that limit (see Appendix). The result is that an estimated mean $\hat{\theta}$ should be subtracted from each observation φ_r at the outset, and added back to the estimate $\hat{\theta}_x$ at the end

$$\hat{\theta}_x = \hat{\theta} + \sum_r C_{xr} \left\{ \sum_s A_{rs}^{-1} (\varphi_s - \hat{\theta}) \right\}, \tag{20}$$

where $\hat{\theta}$ is calculated in such a way that the sum of the weights is exactly zero:

$$\sum_r \eta_r = \sum_{r,s} A_{rs}^{-1} (\varphi_s - \hat{\theta}) = 0. \tag{21}$$

Note that the estimated mean $\hat{\theta}$ is not the arithmetic average of the observations. Two widely spaced data points have more influence on $\hat{\theta}$ than two close together. This procedure gives the minimum error variance [equation (24) below] in the final estimate which is slightly larger than if the mean were in fact known *a priori*. It may be shown that the results for $\hat{\theta}_x$ or $\hat{\theta}$ are unaffected by subtracting any constant from the function $F(\xi)$ used to calculate C_{xr} and A_{rs} . Thus instead of $F(\xi)$ we may use $-1/2 G(\xi)$, which is unambiguous.

It is also worth noting that the result above can be achieved by insisting that the estimate

$$\hat{\theta}_x = \sum_s \hat{\alpha}_{xs} \varphi_s \tag{22}$$

have a zero mean error, or zero bias. Since $\hat{\theta}_x = \bar{\varphi}$, this is accomplished if

$$\sum_{s=1}^N \hat{\alpha}_{xs} = 1. \tag{23}$$

The modified estimator (20) is obtained by minimizing the right-hand side of equation (14) subject to equation (23) as a constraint. The

minimum which remains is the variance of the error and is easily shown to be:

$$\begin{aligned} \overline{(\theta_x - \hat{\theta}_x)^2} &= C_{xx} - \sum_{r,s} C_{xr} A_{rs}^{-1} C_{sx} \\ &+ \frac{(1 - \sum_{r,s} C_{xs} A_{sr}^{-1})^2}{\sum_{r,s} A_{rs}^{-1}}. \end{aligned} \quad (24)$$

The last term is the increase associated with the uncertainties of the estimated mean.

2.4 Nonstationary mean

Unfortunately, it is often the case that the ensemble mean cannot be estimated from data, particularly when this mean value is not stationary: that is when it depends on x . Imperfect knowledge of the mean value influences application of the basic algorithm in two related ways. First, the fundamental estimator (7) is the optimal linear estimator only for variables with zero mean. Second, if the mean is unknown it is not possible to compute the covariance $F(\xi)$ from observations. The method discussed in the previous section provides an objective method of dealing with both these difficulties in the case that the mean value is independent of x . But many mapping tasks do not fall in this category and for them some alternative method must be employed. In particular, mapping the MODE-I isotherms depth posed the problem of dealing with unknown and apparently non-stationary mean values.

In assessing the importance of imperfect knowledge of the relevant mean values it is important to note that the proof of the central algorithm, as given in (13) and (14), does not depend on these mean values vanishing. If they are not zero it is still true that the basic estimator (7) is that linear predictor of the form (13) which minimizes the mean square error between θ_x and $\hat{\theta}_x$. But this mean square error is not necessarily the error variance and the statistical quantities A_{rs} and C_{xs} defined by

$$A_{rs} = \overline{\varphi_r \varphi_s} \quad C_{xr} = \overline{\theta_x \varphi_r}$$

are mean products and not necessarily covariances. The difficulty when the mean values are not zero is that an estimate of the form (13), which is homogeneous with respect to the data φ_s , is not the best possible linear estimator and is outperformed by the more general form

$$\hat{\theta}_x = \sum_r \alpha_{xr} \varphi_r + a_x.$$

Unfortunately, there is no direct method of finding a_x without knowledge of the mean values of θ_x and φ . The method of Section 2.4 is not this optimum estimate. It is, rather, the best estimate of the form (13) which is not influenced by the value of the unknown mean if that mean value is independent of x .

In order to minimize the effect of unknown mean values one must deal with a very fundamental problem. The basic method is based on mean values defined as the average over an ensemble of identically prepared realizations. Only if these mean values are independent of x can they be determined from spatial and temporal averages. If this is not the case no progress can be made without appeal to some *a priori* prejudice.

One practical approach is to consider the field to be composed of two components according to

$$\theta_x = \theta_x' + \langle \theta(x) \rangle, \quad \varphi_s = \varphi_s' + \langle \theta(x_s) \rangle.$$

The primed quantities represent small-scale components of the field which are assumed to have small mean values. The bracketed quantity is a large-scale component of the field which includes the overall mean of θ but is not an approximation of it. The arbitrary separation of these components is determined by the operational rule used to find the value of $\langle \theta \rangle$. The objective used in choosing this operation is to make the mean value of θ' as small as possible; in this case 'mean' refers to the average of θ' taken over an ensemble of realizations in which $\langle \theta \rangle$ is chosen by the same operational rule but has different values.

The choice of $\langle \theta \rangle$ must be based on an *a priori* judgement. But this choice is not crucial to

accurate analysis since the basic estimator (7) will produce a good estimate of $\theta_{\mathbf{x}'}$, and hence $\theta_{\mathbf{x}}$, even if the mean value $\bar{\theta}_{\mathbf{x}'}$ is not zero. In the MODE-I isotherm analysis $\langle\theta\rangle$ was defined to be

$$\langle\theta(\mathbf{x})\rangle = \sum_n a_n f_n(\mathbf{x}),$$

where the f_n were chosen to be a constant and three functions which depend linearly on the position (x, y, t) . It is doubtful that any method of choosing the a_n could insure that the mean of θ' would vanish. Fortunately, this is not essential since the basic algorithm allows θ' to be well mapped with any reasonable choice. For example, if the measurement positions \mathbf{x}_r are uniformly distributed the usual squares fitting technique of minimizing

$$\sum_r [\varphi_r - \sum_n a_n f_n(\mathbf{x}_r)]^2 \quad (25)$$

might be employed. Despite the arbitrary nature of this procedure, it defines $\langle\theta\rangle$ in such a way that the mean of θ' is probably not large.

A somewhat more refined method of defining $\langle\theta\rangle$, which is more appropriate to nonuniformly spaced measurements, makes use of an extension of the objective estimation technique. Clearly, minimizing the sum (25) yields results which depend not only on the field θ but also on the positions at which the measurements are made. A less arbitrary definition of $\langle\theta\rangle$ can be obtained by minimizing

$$\int_S [\hat{\theta}_{\mathbf{x}} - \sum_n a_n f_n(\mathbf{x})]^2 d\mathbf{x},$$

where S is that portion of \mathbf{x} over which θ is to be mapped and $\hat{\theta}$ is the best estimate over that region. If the trial functions are, for convenience, made orthogonal over this region, then determining a_n involves evaluation of the integral

$$\int_S \hat{\theta}_{\mathbf{x}} f_n(\mathbf{x}) d\mathbf{x}. \quad (26)$$

Techniques for accomplishing this integral are given in Section 2.6 where it is shown that this weighted integral of the optimal estimate $\hat{\theta}$ is equal to the optimal estimate of the same weighted integral of θ itself. It must be pointed out that this more refined definition of $\langle\theta\rangle$ involves an iterative process associated with the cycle of evaluating $\langle\theta\rangle$, calculating the mean products of θ' and φ' , and then re-evaluating $\langle\theta\rangle$.

2.5 Vector fields

It is not necessary that the variable to be estimated be a scalar. For example, given observations of velocity at N points \mathbf{x}_r , it is a simple matter to adapt the basic technique to making optimal estimates of the continuous velocity field. One procedure is to estimate separately the components $u_i(\mathbf{x})$ using as data the measured values of the components of velocity at \mathbf{x}_r . Thus for a two-dimensional field there are $2N$ data which can be denoted as

$$\varphi_r = \tilde{u}_1(\mathbf{x}_r) \quad \varphi_{r+N} = \tilde{u}_2(\mathbf{x}_r),$$

where $r = 1, \dots, N$ and the tilde denotes a measured value. The matrix A has $2N \times 2N$ elements

$$A_{r,s} = \overline{u_1(\mathbf{x}_r) u_1(\mathbf{x}_s)} + E \delta_{rs},$$

$$A_{r,s+N} = A_{s+N,r} = \overline{u_1(\mathbf{x}_r) u_2(\mathbf{x}_s)},$$

$$A_{N+r,N+s} = \overline{u_2(\mathbf{x}_r) u_2(\mathbf{x}_s)} + E \delta_{rs}.$$

According to the fundamental result (7), the best estimate of the velocity components is

$$\hat{u}_i(\mathbf{x}) = \sum_{r=1}^{2N} C_{\mathbf{x}r}^{(i)} \eta_r = \sum_{r=1}^{2N} C_{\mathbf{x}r}^{(i)} \left\{ \sum_{s=1}^{2N} A_{rs}^{-1} \varphi_s \right\}, \quad (27)$$

where, from (9),

$$C_{\mathbf{x},r}^{(i)} = \overline{u_i(\mathbf{x}) u_1(\mathbf{x}_r)}, \quad C_{\mathbf{x},r+N}^{(i)} = \overline{u_i(\mathbf{x}) u_2(\mathbf{x}_r)}.$$

If the velocity field is homogeneous and isotropic the velocity covariances can be derived (cf. BATCHELOR, 1960) from two scalar functions

$R(\rho)$ and $S(\rho)$ according to

$$\overline{u_i(\mathbf{x}) u_j(\mathbf{x}+\mathbf{r})} = \chi_i \chi_j [R(\rho) - S(\rho)] + \delta_{ij} S(\rho),$$

where $\rho = |\mathbf{r}|$, R and S are known, respectively, the longitudinal and transverse velocity covariances, and χ_i is the cosine of the angle between \mathbf{r} and the i axis. If the field is known to be non-divergent then the velocity is derivable from a stream function $\psi(\mathbf{x})$ according to

$$u_1 = -\frac{\partial\psi}{\partial x_2}, \quad u_2 = \frac{\partial\psi}{\partial x_1}, \quad (28a, b)$$

and the two covariance functions can be derived from a single function. Taking the covariance of the stream function to be

$$\overline{\psi(\mathbf{x}+\mathbf{r}) \psi(\mathbf{x})} = F(\rho), \quad (28c)$$

then

$$R(\rho) = -\frac{1}{\rho} \frac{dF}{d\rho}, \quad S(\rho) = -\frac{d^2F}{d\rho^2}. \quad (28d, e)$$

The r.m.s. errors associated with the estimate $\hat{\mathbf{u}}(\mathbf{x})$ are most usefully plotted as maps of the normalized standard deviation of the total velocity

$$\frac{\{\overline{|\hat{\mathbf{u}}-\mathbf{u}|^2}\}^{1/2}}{\{\overline{|\mathbf{u}|^2}\}^{1/2}} = \left\{ 1 - \frac{\sum_i \sum_{r,s} C_{xr}^{(i)} A_{rs}^{-1} C_{xs}^{(i)}}{\sum_i C_{xx}^{(i)}} \right\}^{1/2}. \quad (29)$$

It is important to note that if $\hat{\mathbf{u}}$ is determined using (27) and the covariances $C_{xr}^{(i)}$ are derived from a covariance function appropriate to a non-divergent field then the divergence of the estimated field will vanish identically. This is so because

$$\begin{aligned} \frac{\partial \hat{u}_1}{\partial x_1} + \frac{\partial \hat{u}_2}{\partial x_2} &= \sum_r \eta_r \left\{ \frac{\partial C_{xr}^{(1)}}{\partial x_1} + \frac{\partial C_{xr}^{(2)}}{\partial x_2} \right\} \\ &= \overline{\left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \right) \varphi_r} \end{aligned}$$

vanishes when the covariances are taken from a nondivergent field, regardless of the weights η_r or the data φ_r upon which they depend. In a later section it will be shown that this is related to a general result of considerable significance when objective maps are used to test dynamical hypotheses.

It is generally found that when the velocity field is nondivergent the most useful presentation of the estimated field $\hat{\mathbf{u}}$ is in the form of a map of the stream function ψ . There are two seemingly different approaches possible here but, happily, they produce the same result. The first approach is to abandon the velocity estimate in (27) and form the best estimate of the true stream function. It is seen that this is accomplished using an expression similar to (7),

$$\hat{\psi}(\mathbf{x}) = \sum_{r=1}^{2N} P_{xr} \left\{ \sum_{s=1}^{2N} A_{rs}^{-1} \varphi_s \right\} \quad (30)$$

$$P_{x,r} = \overline{\psi(\mathbf{x}) u_1(\mathbf{x}_1)}, \quad P_{x,r+N} = \overline{\psi(\mathbf{x}) u_2(\mathbf{x}_r)}.$$

The alternative is to derive a stream function from the velocity estimates $\hat{\mathbf{u}}$. This is accomplished by computing $\hat{\psi}$ from

$$\hat{\psi}(\mathbf{x}) = \int_{\mathbf{x}_0}^{\mathbf{x}} u_2(\mathbf{r}) d\mathbf{r}_1 - \int_{\mathbf{x}_0}^{\mathbf{x}} u_1(\mathbf{r}) d\mathbf{r}_2 + \psi(\mathbf{x}_0), \quad (31)$$

which, it should be noted, does not depend on the integration path if the divergence of $\hat{\mathbf{u}}$ vanishes. Inspection of (27), from which $\hat{\mathbf{u}}$ is computed, shows that this result is identical with that obtained by estimating the stream function directly. This is the result of an important aspect of the basic estimation algorithm, which is discussed in the following section.

In the case of an isotropic nondivergent field the covariances P appearing in (30) can be found directly from (28):

$$\begin{aligned} \overline{\psi(\mathbf{x}) u_1(\mathbf{x}+\mathbf{r})} &= -\gamma_2 \frac{dF}{d\rho}, \\ \overline{\psi(\mathbf{x}) u_2(\mathbf{x}+\mathbf{r})} &= \gamma_1 \frac{dF}{d\rho}, \end{aligned}$$

where $\rho = |\mathbf{r}|$. When the field is not isotropic the relevant covariances must be derived from the more complicated velocity covariance matrix.

2.6 Operationally defined fields

The stream function, whose estimation was discussed in the previous section, is an example of a quantity which is defined by a linear operation on the field which is measured; the definition (30) relates the stream function to the components of velocity through an integral operator. It is often desirable to estimate such derived quantities, and it is a major asset of the objective analysis technique that this is readily accomplished. Examples of dynamically interesting derived fields are transport between two positions (which might be compared with dynamic height differences to test geostrophy), vorticity, divergence, heat content (the integral of temperature over a region), circulation, and weighted integrals such as (26). One additional example, which plays a central role in the evaluation of the synoptic mapping ability of arrays, is the estimation of smoothed or filtered fields defined by operations of the form

$$T_x = \int_r \theta(\mathbf{r}) W(\mathbf{r} - \mathbf{x}) d\mathbf{r}, \quad (32)$$

where the smoothing function W is centred around $\mathbf{r} - \mathbf{x} = 0$.

The determination of such derived fields is accomplished without significant extension of the basic technique. Let the derived quantity be defined by a linear differential or integral operator L acting on the field θ according to

$$T_x = L_x(\theta). \quad (33)$$

According to the basic algorithm (7), the least mean square error estimate of T_x is

$$\hat{T}_x = \sum_r Q_{xr} \eta_r = \sum_{r,s} Q_{xr} A_{rs}^{-1} \varphi_s, \quad (34)$$

where φ_s are imperfect measurements of θ , A_{rs} is given by (8) and, from (9)

$$Q_{xr} = \overline{\varphi_r T_x} = \overline{\varphi_r L_x(\theta)}.$$

The mean square error, given by a modification of (10), is

$$\overline{(T_x - \hat{T}_x)^2} = \overline{T_x^2} - \sum_{r,s} Q_{xr} Q_{xs} A_{rs}^{-1}. \quad (35)$$

Two features of the estimate \hat{T} are important to note. First, the weights η_r are the same regardless of what quantity is being estimated. Thus the weights used to estimate θ_x are the same as those used to estimate $L_x(\theta)$; it is only the covariance Q_{xr} which changes. The second observation follows directly from the first: the best estimate of $L_x(\theta)$, a linear operation on θ , is equal to that operation on the best estimate of θ , that is $L(\hat{\theta}_x)$. This is easily shown, since

$$L_x(\hat{\theta}) = L_x \sum_r \overline{\theta_x \varphi_r} \cdot \eta_r = \sum_r \overline{\varphi_r L_x(\theta)} \cdot \eta_r$$

is identical to (34).

The determination of the covariance Q_{xr} is usually accomplished from the more fundamental covariance

$$C_{xr} = \overline{\theta_x \varphi_r} = \overline{\theta_x \theta_{x_r}},$$

through application of the operation L . For example, the covariance associated with the filtering operation (31) is

$$Q_{xs} = \int_r \overline{\theta_{xs} \theta_r} W(\mathbf{r} - \mathbf{x}) d\mathbf{r}$$

where, if θ has stationary statistics, the integrand can be expressed in terms of the covariance F of (9).

2.7 Constraints and hypothesis testing

In the previous section it was shown that the optimal estimate of some linear operation on a field is equal to the linear operation on the optimal estimate of the field. Further, it was shown that the result of linear operations on the estimated field is strongly influenced, and in some cases

completely determined, by the statistics used to estimate that field. For example, in Section 2.5, estimation of components of the velocity vector was discussed and it was found that the inferred velocity field would be nondivergent if the statistics used to estimate it were consistent with a nondivergent field.

These properties of the objective analysis technique provide a useful tool for improving the accuracy of estimates by allowing the application of *a priori* constraints which are justified by reasons found outside the data base. For example, on theoretical rather than observational grounds it is desirable to constrain the horizontal velocity estimates associated with mesoscale motion to be nondivergent; enforcing this constraint is easily accomplished by using, in the estimation process, statistics which are consistent with it. Another example, which is used widely in meteorological objective analysis, arises in mapping pressure using both pressure measurements and observations of the wind, which is assumed to be geostrophically balanced. By employing pressure-wind statistics consistent with this dynamical constraint, it is possible to improve pressure maps by including wind observations.

On the other hand, the properties of the analysis technique can make difficult the process of testing dynamical hypotheses through the use of objective maps. This is so because it is difficult to avoid the inadvertent application of some dynamical constraints when selecting the statistics used to estimate the field. If these constraints bear on the dynamical hypothesis being tested it is possible to arrive at a conclusion based more on the method of analysis than the data themselves. For example, if velocity components were estimated using the constrained velocity statistics of (28), the conclusion that the field was nondivergent would be reached regardless of the nature of the measurements. It is evident from this example that one way of testing the hypothesis of nondivergent flow is to compare the statistics of the flow field to see if they are consistent with the hypothesis. The alternative is to map separately each component using only the observations of that component and then compute the divergence.

The difficulty of using objective maps to test certain dynamical hypotheses raises an interesting and fundamental question concerning the most appropriate way to answer dynamically motivated questions through the use of data. The question is essentially whether verification of a dynamical model of a phenomenon is best tested through comparison of synoptic observations and model 'hindcasts', or through comparison of measured and predicted statistics. This is not the place to discuss this in detail, but a simple example will serve to focus the question.

Suppose that the phenomenon under examination is to be modelled by the linear differential system

$$L_x \theta = g \quad \text{with} \quad \theta = b \text{ on } B \quad (36)$$

where b is both the necessary initial and boundary data. The actual dynamics of the system are assumed to be unknown and the question is whether or not they can be adequately described by the model. We can answer the question by using an incomplete set of imperfect measurements φ_s and γ_s of the unknown and the forcing function, respectively.

One approach is to use the measurements to make an estimate \hat{b} of the boundary data over all B and an estimate \hat{g} of the forcing function. Using these, the model could be used to hindcast θ according to

$$L_x \hat{\theta} = \hat{g} \quad \hat{\theta} = \hat{b} \text{ on } B. \quad (37)$$

Model verification could be approached through comparison of the hindcast $\hat{\theta}$ with an objective estimation of the true field. The best estimate of the true field would, according to (7), be

$$\hat{\theta}_x = \sum_{r,s} \overline{\theta_x \varphi_r} A_{rs}^{-1} \varphi_s + \sum_{r,s} \overline{\theta_x \gamma_r} G_{rs}^{-1} \gamma_s, \quad (38)$$

where $A_{rs} = \overline{\varphi_r \varphi_s}$ and $G_{rs} = \overline{\gamma_r \gamma_s}$. Determination of the prescription data for the model involves estimating b over B and g over all x interior to B . The best estimate of b would use measurements of both θ and g , and would lead to exactly the

formula (38). Optimal estimation of g would, similarly, make use of all data, and would lead to

$$\hat{g}_x = \sum_{r,s} \overline{g_x \theta_r} A_{rs}^{-1} \varphi_s + \sum_{r,s} \overline{g_x \gamma_r} B_{rs}^{-1} \gamma_s.$$

The point of interest is that the outcome of the comparison of the hindcast $\hat{\theta}$ with the estimated field $\hat{\theta}$ can be determined through examination of the statistics used to form the estimate. This is easily seen by noting that, according to (37) and (38),

$$\begin{aligned} L_x(\hat{\theta} - \hat{\theta}) &= \sum_{r,s} \overline{(g_x \varphi_r - L \theta_x \varphi_r)} A_{rs}^{-1} \varphi_s \\ &+ \sum_{r,s} \overline{(g_x \gamma_r - L \theta_x \gamma_r)} G_{rs}^{-1} \gamma_s, \end{aligned} \quad (39)$$

and, as has already been shown, that $\hat{\theta} - \hat{\theta} = 0$ on B . Thus the degree of similarity between hindcast and 'measured' fields is directly related to the statistics

$$L \overline{\theta_x \varphi_r} \quad \text{and} \quad L \overline{\theta_x \gamma_r}.$$

If θ is, in fact, described by the model (36) and there are no errors in determining the relevant statistics, then

$$L \overline{\theta_x \varphi_r} = \overline{g_x \varphi_r}, \quad L \overline{\theta_x \gamma_r} = \overline{g_x \gamma_r}, \quad (40)$$

respectively, and, as long as the model has a unique solution, the estimated and hindcast fields will agree exactly even though they are not necessarily the true field from which the data were drawn.

The above example demonstrates that the model can be tested either by comparison of synoptic measurements and hindcasts, that is $\hat{\theta}$ and $\hat{\theta}$, or through examination of the statistics from which $\hat{\theta}$ is estimated, that is by determining if (40) holds. Regardless of how the comparison is done, the result is greatly influenced by the statistics.

2.8 Estimated covariances

Throughout the foregoing it has been assumed that the covariances, upon which the technique is based, are known. But it must be obvious that the determination of these quantities is a difficult task unless there is a great amount of data available. In the discussion of the MODE-I current meter array in Section 4 it will be shown that serious errors can be made if the covariances are estimated from limited data. The purpose of this subsection is to present some simple ideas which can aid in the estimation of covariances for use in objective analysis. Estimation of two-point statistics from unequally spaced data is a difficult problem which will be the subject of a forthcoming paper and, therefore, the present discussion must be considered as only introductory.

The first, and absolutely essential, point is that the estimated covariances must be possible covariances, that is they must be non-negative definite forms. The authors have, through unfortunate experience, demonstrated that if the estimated covariances are not non-negative forms then the objective analysis algorithm can yield extremely bad results. By non-negative form we do not mean that the covariance can not sometimes have a negative algebraic value, but rather that every covariance matrix

$$A_{rs} = \overline{\theta_r \theta_s}$$

drawn from the covariance function must be a non-negative definite matrix in the sense that none of its eigenvalues are negative.

The reason why it is essential that the covariance matrices be non-negative can be seen by recasting the proof (given in Section 2.1) of the fundamental algorithm. The basic aim of the estimator (7) is to minimize the difference between the true value of θ and an estimate $\hat{\theta}$ which is of the linear form (13). This error is

$$\delta = \hat{\theta}_x - \theta_x = \sum_{r=1}^N \alpha_r \varphi_r - \theta_x = \sum_{r=1}^{N+1} \alpha_r \varphi_r,$$

where, for notational simplicity, the definitions $\alpha_{N+1} = -1$ and $\varphi_{N+1} = \theta_x$ have been introduced.

The Gauss–Markov theorem provides the method for minimizing the mean square error

$$\overline{\delta^2} = \sum_{r,s}^{N+1} \overline{\varphi_r \varphi_s} \alpha_r \alpha_s.$$

Clearly the true mean square error must be non-negative for any choice of α , and, as a consequence the true covariance $\overline{\varphi_r \varphi_s}$ must be a non-negative matrix. The true covariance is usually imperfectly known and this is not crucial. What is crucial is that the estimated covariance be a possible covariance, that is non-negative. If this is not the case then the ‘mean square error’ can be negative, and the process of finding the weights, which involves extremizing this error, not minimizing its magnitude, will lead to a very bad estimator. Additionally, it might be pointed out that the optimum weights can be found from (7) only if the covariance matrix is positive definite, but in practice this distinction is rarely of significance.

In addition to insisting that the estimated covariance be positive definite, it is desirable to make it approximate the true covariance. The easiest way of accomplishing these two objectives is to estimate the covariance function $F(\mathbf{x})$ at a number of discrete values of the lag vector \mathbf{x} and to choose some functional representation for F which approximates these estimates and guarantees that all covariance matrices derived from it are positive definite. If the covariances are stationary, so that they can be derived from some $F(\mathbf{x})$ which depends only on separations, then the positive definite character is insured if the associated spectrum $\int \overline{f}(\boldsymbol{\kappa})$ in (4) is, for every $\boldsymbol{\kappa}$, a positive number.

The process of estimating the value of the covariance function $F(\mathbf{x})$ at discrete values of \mathbf{x} can be aided by use of the objective estimation algorithm. In doing this the data φ_s are products of measurements of the field, $u(\mathbf{x})$ say, for which the covariance is to be estimated. Thus to estimate $F(\mathbf{x}_0)$ the data to be used are all possible products:

$$\varphi_s = u(\mathbf{r})u(\mathbf{r} + \mathbf{x}_0).$$

If the data are irregularly spaced it may be

desirable to collect together all products with separations which are close, but do not equal \mathbf{x}_0 . In any case, the estimate of $F(\mathbf{x}_0)$ is taken to be

$$\hat{F}(\mathbf{x}_0) = \sum \alpha_s \varphi_s,$$

and, to insure that there is no mean error, it is required that

$$\overline{\hat{F}(\mathbf{x}_0) - F(\mathbf{x}_0)} = \sum_s \alpha_s \overline{\varphi_s} - F(\mathbf{x}_0) = (\sum_s \alpha_s - 1) F(\mathbf{x}_0) = 0.$$

The mean square error is then

$$\overline{(\hat{F} - F)^2} = \sum_{r,s} \overline{\varphi_r \varphi_s} \alpha_r \alpha_s - F^2(\mathbf{x}_0); \quad (41)$$

the simplicity of this as compared with (14) results from the fact that F is, itself, a mean value. Minimization of the mean square error, subject to the zero bias constraint that $\sum_s \alpha_s = 1$, is easily accomplished by the usual technique if the covariance

$$\overline{\varphi_r \varphi_s} = \overline{u(\mathbf{x}_r) u(\mathbf{x}_r + \mathbf{x}_0) u(\mathbf{x}_s) u(\mathbf{x}_s + \mathbf{x}_0)}$$

is known.

It is rarely possible to estimate the fourth moment involved in $\overline{\varphi_r \varphi_s}$ directly from data. But, if it is warranted to assume that the statistics of u are approximately Gaussian and $\overline{u} = 0$, then this statistic can be estimated from two-point covariances according to

$$\begin{aligned} \overline{\varphi_r \varphi_s} &= \overline{u(\mathbf{x}_r) u(\mathbf{x}_r + \mathbf{x}_0)} \cdot \overline{u(\mathbf{x}_s) u(\mathbf{x}_s + \mathbf{x}_0)} \\ &+ \overline{u(\mathbf{x}_r) u(\mathbf{x}_s) u(\mathbf{x}_r + \mathbf{x}_0) u(\mathbf{x}_s + \mathbf{x}_0)} \\ &+ \overline{u(\mathbf{x}_r) u(\mathbf{x}_s + \mathbf{x}_0) u(\mathbf{x}_r + \mathbf{x}_0) u(\mathbf{x}_s)}. \end{aligned}$$

Relating this to the covariance function $F(\mathbf{x})$ and substituting into (41) leads to a new expression for the mean square error:

$$\begin{aligned} \overline{(\hat{F} - F)^2} &= \sum_{r,s} \{ F^2(\mathbf{x}_r - \mathbf{x}_s) \\ &+ F(\mathbf{x}_s + \mathbf{x}_0 - \mathbf{x}_r) F(\mathbf{x}_r + \mathbf{x}_0 - \mathbf{x}_s) \} \alpha_r \alpha_s, \end{aligned} \quad (42)$$

where use has been made of the constraint $\sum_s \alpha_s = 1$.

The procedure for finding a good estimate of $F(\mathbf{x})$ is (1) choose a first guess of the function $F(\mathbf{x})$, (2) use this to compute the bracketed term in (42), (3) find the α which minimize the error in (42) subject to the constraint $\sum_s \alpha_s = 1$, (4) use these to evaluate $F(\mathbf{x})$ at various values of \mathbf{x} , (5) from the discrete values adjust the function F , and repeat steps (2) to (5) until F converges.

3. SYNOPTIC ARRAY DESIGN

3.1 Approach

As part of the array design exercise for MODE-I, an examination of the synoptic mapping ability of various possible array designs was made. From this, a general picture of the resolving power of arrays was gained, and it is the purpose of this section to present some of these results. The approach here is to choose a particular simple array geometry and to determine that array's ability to map various different fields. In practice, the procedure of array design involves doing this for a number of different arrays, and choosing that array which strikes an appropriate balance between accuracy and spatial coverage.

For the present purposes a somewhat different approach will be given to 'noise' than was used in the previous discussion. In practice, one is rarely able to determine from the data the extent to which the observed noise is due to instrumental effects or to a small-scale signal present in the field being sampled. All that is known is that the signals from the most closely spaced measurements available are not perfectly correlated; it is impossible to determine the shape of the correlation function for non-zero separations smaller than this smallest observed separation. The approach adopted here is to assume that the measurements are not subject to instrumental error but that there is a small-scale signal present. In this case all covariances can be computed from a single covariance function which, for the present purposes, is taken to depend only on the distance between the points involved. Thus

$$\overline{\varphi_s \varphi_s} = \overline{\theta_x \theta_x} = \overline{\theta_x \theta_x} = F(|\mathbf{x}_r - \mathbf{x}_s|). \quad (43a)$$

The covariance function F will be taken to be

$$F(x) = \exp(-x^2/R_c^2) + E \exp(-x^2/R_n^2), \quad (43b)$$

where $x = |\mathbf{x}|$, R_c is a measure of the scale of the signal (which has unit variance), and R_n is the scale of the noise which has variance E . The error in estimating the unknown θ at the position \mathbf{x} will be expressed as the standard deviation

$$\epsilon_x = \overline{(\theta_x - \hat{\theta}_x)^2},$$

which, since the large-scale signal variance is unity, is the r.m.s. error relative to the standard deviation of the signal.

One question of interest is how the mapping error depends on the scale of the noise and on whether the noise is instrumental or due to small-scale variability. If the scale of the noise is allowed to approach zero, then the noise covariance in (43b) approaches a delta function of value E . This case is similar to, but not the same as, the situation in the previous discussion, where the noise was assumed to be instrumental. So far as mapping errors are concerned, the difference is confined to estimates at the array element positions; in the present case perfect observations are possible at each element but not in the immediate neighbourhood, whereas if the noise is instrumental approximately the same error is associated with estimates at and near to the elements. Aside from the error at the array elements there is no difference between instrumental noise and variability with an infinitesimal spatial scale.

Before turning to the mapping ability of arrays it is instructive to examine the errors associated with extrapolation from a single measurement. According to (10) the error from one datum is

$$\epsilon_x = \{1 - F^2(x)/F^2(0)\}^{1/2}.$$

It is alarming to note that the correlation between measurement and the signal to be estimated, that is $F(x)/F(0)$, must exceed 0.95 if the error standard deviation is to be less than 30% of the r.m.s.

signal. Thus to gain a 30% estimate in the immediate neighborhood of the measurement (or even at the measurement location if the noise is instrumental) the noise variance must be less than 5% of the signal variance. It is clear from this that accurate mapping requires either closely spaced and noise-free measurements or an estimation procedure which takes into account many data in order to reduce errors; the objective interpolation method is just such a procedure. Further, by choosing to map spatially filtered or smoothed fields it is possible to reduce errors below those associated with estimation of point values.

3.2 Noise and scales

The influence of the noise variance E and the signal and noise scales R_c and R_n on mapping errors will now be examined for the simple uniformly spaced 12-element array shown in Fig. 1. In the discussion the array spacing will be maintained at unit value, and the field statistics, as described by (43b), will be allowed to vary. The error associated with mapping a scalar variable e_x will be determined from (10) and portrayed as the r.m.s. error ϵ_x .

Figure 1 depicts the spatial distribution of the mapping error for various different field statistics. Because the array is symmetric about the x and y axes, there are four error maps shown, one in each quadrant, and each is labelled according to the parameters R_c , E and R_n , which describe different characteristics of the field statistics. The gross influence of these statistical characteristics can be seen by comparing quadrants (b), (c) and (d) with (a), which serves as the standard. The parameters associated with map (a) are $R_c = 0.8$, $E = 0.1$ and $R_n = 0$. Because the noise scale R_n is infinitesimal, the error jumps discontinuously to zero at each array position; if the noise were instrumental the error map would be identical except that the error at each element would be essentially the same as that found in the immediate neighborhood of the element, that is approximately 0.45.

The effect of a finite scale of noise can be seen in map (b) for which $R_n = 0.3$ while R_c and E are

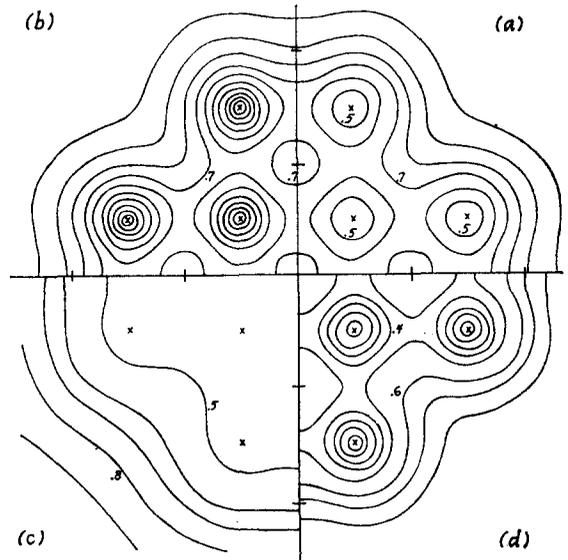


Fig. 1. Standard deviation of the mapping error for the symmetric array having 12 elements at the positions shown by \times . Each quadrant is one quarter of the error map for different parameters of the field statistics. The parameters (see text) are (a) $R_c = 0.8$, $E = 0.1$, $R_n = 0$; (b) $R_c = 0.8$, $E = 0.1$, $R_n = 0.3$; (c) $R_c = 1.2$, $E = 0.1$, $R_n = 0$; (d) $R_c = 0.8$, $E = 0$.

maintained at 0.8 and 0.1, respectively. The influence of varying the noise scale is confined to regions near each element in which the error decreases rapidly to zero. As discussed earlier, the hypothesis that the noise is of extremely small scale leads to conservative error estimates in the immediate areas surrounding each element.

The influence of the spatial scale of the signal can be seen through comparison of (a) and (c). In each the noise parameters are held at $E = 0.1$ and $R_n = 0$, but for (a) $R_c = 0.8$ while in (c) $R_c = 1.2$. As might be expected, the area of accurate mapping is greatly expanded when the signal scale exceeds the array spacing.

The penalty associated with small-scale noise can be assessed by comparison of maps (a) and (d). In both the signal scale $R_c = 0.8$; in (a) there is present noise of infinitesimal scale and variance $E = 0.1$ while in (d) there is no noise, that is $E = 0$. The extent of useful mapping is not greatly changed as the area within $\epsilon = 0.6$ is not significantly different between the maps. But the region

of accurate mapping is greatly increased when the noise vanishes.

A somewhat more complete picture of the relative importance of signal scale and noise variance can be seen in Fig. 2 which presents the mapping error at the center of the 12-element array for various values of R_c and E . These error estimates were made assuming $R_n = 0$ but, as was seen above, the error this far from the array elements is not affected by the value of R_n so long as it is less than about 0.3.

3.3 Smoothing

The discussion above pertains to mapping the pointwise value of the scalar variable θ . Typically the variable to be mapped is subject to some small-scale variability, associated with the noise parameters E and R_n , which can not be well resolved by an array designed to map the signal which has the larger scale R_c . If it is the large-scale signal which is of interest then the quantity which should be mapped is the filtered or smoothed variable

$$\theta_x = \int_r W(\mathbf{r}-\mathbf{x}) \theta_r \, dr,$$

where the filter function W is centred around $\mathbf{r}-\mathbf{x} = 0$ but extends over sufficient area to suppress the small-scale components of the field. Typically the errors associated with estimating the smoothed field are significantly less than those involved in estimating θ itself.

The filter function adopted for the present study is a radially symmetric Gaussian filter

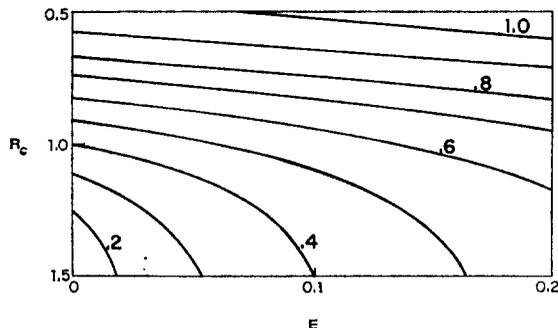


Fig. 2. Standard deviation of mapping error at the center of the 12-element array shown in Fig. 1. The noise scale is held at $R_n = 0$.

$$W(\mathbf{x}) = \frac{1}{\pi R_f^2} \exp(-|\mathbf{x}|^2/R_f^2),$$

and the measure of error is taken as the standard deviation of the estimation error normalized by the standard deviation of the smoothed field, that is

$$\tilde{\epsilon}_x = \overline{(\theta_x - \hat{\theta}_x)^2/\theta_x^2}^{1/2}.$$

Examination of the accuracy of estimating various smoothed fields is accomplished by holding the statistics of the field, itself, fixed and varying the filter width R_f .

Figure 3 depicts the errors associated with estimating various smoothed fields derived from fields with point-wise covariances given by (43) with $R_c = 0.8$, $E = 0.1$ and $R_n = 0.3$; the array is the same as that used in Section 3.2 and the parameter choice is the same as those used in error map (b) of Fig. 1. Map (a) is for $R_f = 0$ and reproduces, with a slight modification of normalization, the error estimates of Fig. 1(b); the error goes to zero at the array elements but increases to 0.7 in the interior of the array. Quadrants (b), (c) and (d) show the error associated

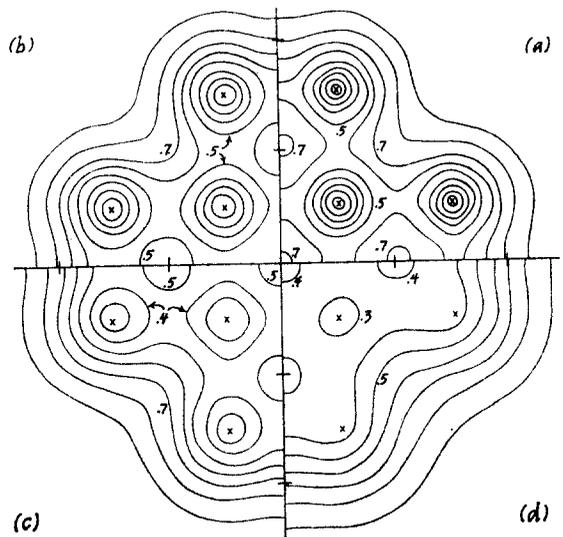


Fig. 3. Standard deviation of error in mapping filtered fields. The array is that of Fig. 1; the field statistics are described by $R_c = 0.8$, $E = 0.1$, $R_n = 0.3$ [as in Fig. 1(b)]. Quadrant (a) is for the unfiltered field ($R_f = 0$) while (b), (c) and (d) are for $R_f = 0.2, 0.4$ and 0.6 , respectively.

with increasingly severe smoothing; the filter parameter for these maps is $R_f = 0.2, 0.4$ and 0.6 , respectively. The trend is clear. With increasing filtering the error becomes more uniform over the array, with the error increasing near to the array element and decreasing in the interior.

It was shown in Section 2.6 that the operation of smoothing commutes with the estimation operation so that maps can be smoothed after they are drawn or smoothed maps may be generated directly. In either case, the error maps in Fig. 3 allow some idea of the accuracy of estimating different components of the field.

4. MODE-I CURRENT METER ARRAY

4.1 Simulated data

To test the objective interpolation algorithms, and to provide some insight into the expected performance of the proposed MODE-I current meter array, velocity fields were constructed to resemble (in a statistical sense) those previously observed in the MODE area. The most complete set of available data comes from Array-I (GOULD, SCHMITZ and WUNSCH, 1974) which was an array of 8 moorings set during fall of 1971. Continuous records were available for approximately 70 days from 6 current meters at 1500 m and 7 meters at 3000 m together with one extra record of doubtful quality. Because of the small size and limited duration of this experiment, compounded by apparent inconsistencies associated with different mooring designs, these data were insufficient to determine the complete properties of currents in the area, and even the statistical features are subject to considerable uncertainty. However, for a design study of this type it is preferable to start with a velocity field which is completely and precisely known, provided it is *qualitatively* similar to those which will be encountered in practice.

Accordingly, the first step was to generate a stream function $\psi(x, y)$ which was one realization of a definite statistical ensemble, using a random number algorithm on a computer. Such an ensemble is determined completely by the mean value $\bar{\psi}$ at every point (assumed zero) and the two-point covariances $\overline{\psi(x_1)\psi(x_2)}$ for all pairs x_1, x_2 . The latter were chosen so as not to be

inconsistent with the Array-I data, but were completely determined only by the application of considerable prejudice on the part of the authors. The velocity field was assumed to be

- (i) spatially homogeneous (ignoring persistent local topographic effects which may exist in realistic data),
- (ii) stationary in time (in a statistical sense only),
- (iii) isotropic (no preferred horizontal direction),
- (iv) nondivergent [achieved by first computing ψ and then the velocities according to equation (1)].

The stream function covariance is completely determined by the longitudinal velocity wavenumber spectrum $\hat{E}(k)$ which is the Fourier transform of the longitudinal covariance $R(\rho)$ of (28). Figure 4 depicts the spectrum adopted. It was chosen to be proportional to k^{-3} for $|k|$ greater than a critical value k_0 , and a predetermined slowly increasing function of k from $k = 0$ to $k = k_0$. This provides a finite mean square velocity with a dominant length scale, yet behaviour on smaller scales consistent with the theory of $2\frac{1}{2}$ -dimensional turbulence (CHARNEY, 1971). The

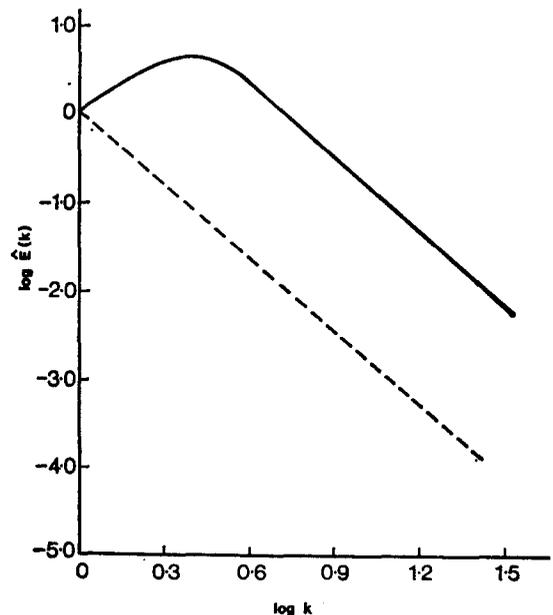


Fig. 4. The one-dimensional velocity spectrum of the simulated fields. The broken line marks the k^{-3} curve. $\log k = 0$ corresponds to a wavelength of 960 km.

value of k_0 and the r.m.s. velocity were adjusted to fit estimates from Array-I (see below).

This ensemble was implemented on a 48×48 square grid of points spaced 20 km apart, using Finite Fourier transforms from a two-dimensional wave number space with the same number of degrees of freedom. The different (quantized) Fourier components are all statistically independent of one another, normally distributed with zero mean and variance proportional to the spectral density at that wave number for the stream function. For any one realization a complete set of such components is generated as a pseudo-random sequence then transformed into physical space. A portion of the result is shown in Fig. 5a. With this procedure ψ is, of course, periodic in x and y with period 980 km, but as the area of interest does not exceed 500 km in diameter this is not a serious deficiency. For an independent realization a different starting value is supplied to the random number algorithm.

To simulate the time dependence of the stream function field, various alternatives were considered. To prescribe an arbitrary frequency spectrum in a similar manner would require a

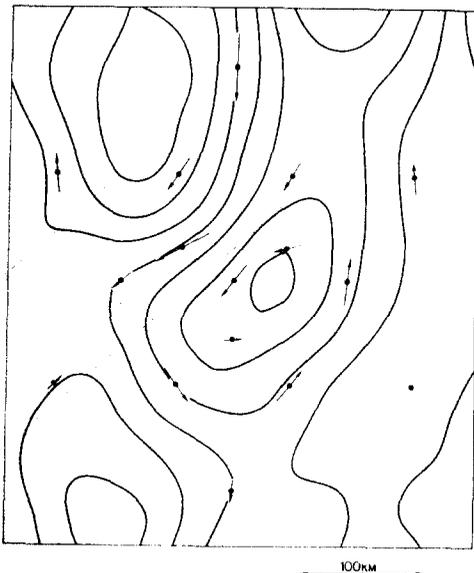


Fig. 5a. A typical simulated stream function field. Superimposed on the field is the MODE-I array with the velocities at the data points represented by the arrows.

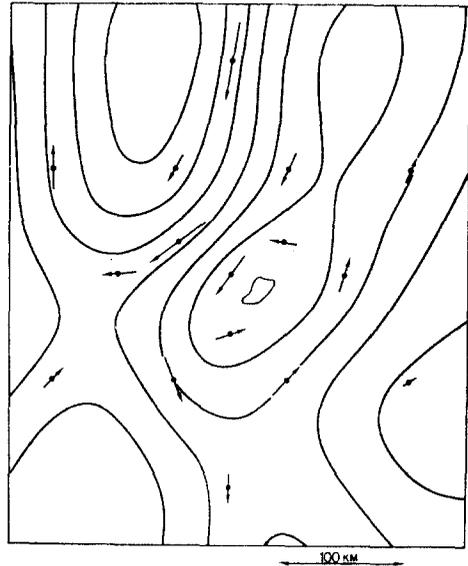


Fig. 5b. A reconstructed version of the stream function field shown in Fig. 5a using the objective analysis technique. An r.m.s. measurement noise of 1.8 km day^{-1} has been applied to the data at the MODE-I array points.

three-dimensional Fourier transform and considerably more storage than was readily available. The simplest scheme which gave a frequency spectrum similar to that in Array-I was to write each Fourier component as the sum of two propagating plane waves, each moving at a phase speed $C = 5 \text{ km day}^{-1}$ in opposite directions. The complete time development is thus specified by $2 \times 48 \times 48$ random real numbers in the initial conditions, and is subsequently fully deterministic. When additional small random adjustments at each time stop were made to the Fourier components (keeping the same r.m.s. level) the result was an overall frequency spectrum which contained more power at very low frequencies than seemed to be indicated by the data. It is recognized that this deterministic scheme (which simulates an intrinsic dynamics like that of the propagation of sound waves) is essentially arbitrary, but, taken with the assumed spatial spectrum, it does reproduce the marked tendency in the available data for the average velocity over 70 to 80 days at any one current meter to be much less than the typical values.

From this complete time history simulated

'data' were computed. The velocity components at a number of fixed points at intervals of 1 day were formed by finite differences. To each such 'measurement' was added a random Gaussian error of zero mean and known standard deviation (usually 1.3 km day^{-1} in each component) to represent the combined effects of instrument noise and imperfectly filtered internal waves. These time series became the current meter 'data'. In addition, 'float trajectories' were computed by integrating the equation

$$\frac{dx}{dt} = -\frac{\partial\psi}{\partial y}(x, y, t); \quad \frac{dy}{dt} = \frac{\partial\psi}{\partial x}(x, y, t)$$

forward in time from prescribed starting positions and recording daily fixes. In this case no random noise was added.

The value of this simulated data depends on the degree to which it resembles the area of the real ocean under study. The parameters which were chosen to fit the available data were the r.m.s. velocity, the cut-off wave number k_0 , the noise level E and the wave propagation speed C . Accordingly, we must now consider how this statistical comparison was made.

4.2 Comparison with observations

Full details of the Array-I data are given by GOULD, SCHMITZ and WUNSCH (1974). Crude estimates of the statistics are shown in Figs. 6 and 7. The longitudinal correlation coefficient for any pair of current meters is found by taking the low-frequency filtered components of velocity at each meter parallel to the line joining them, integrating their product over the shorter period for which the pair of records were available, and dividing by the product of the two r.m.s. values. For each pair the period of integration was between 50 and 70 days, beginning on 1st November 1971. The transverse coefficient involves the components perpendicular to the line of centres. The frequency spectrum is for the period 50 days beginning on the same day. Slightly different values are obtained by taking other base periods. As discussed by GOULD, SCHMITZ and WUNSCH (1974) the speeds (but not direction) of the meters on surface moorings are

probably systematic overestimates. Prior to analysis of this data, therefore, the speed for each meter was multiplied by a constant factor so that its r.m.s. value was equal to the r.m.s. value for the data set as a whole, i.e. 6.7 cm s^{-1} . This empirical adjustment results in a picture which is seemingly much more consistent. Its only effect on these statistics is to give each meter equal weight when computing the average frequency spectra. The correlation coefficients are unaffected.

Figure 6 enables rough estimates to be made of the spatial scale of the velocity field. Visual inspection of the data shows no obvious anisotropy, and this is confirmed by the correlation coefficients. Accordingly they are plotted as a function of meter separation only. Because of the brief duration of the experiment, the scatter is large except at zero separation where all coefficients are unity automatically. Nevertheless, the trend is clearly towards decreasing correlation at larger separations, with the lateral correlation passing through zero at a distance of 60 to 100 km. Also significant is the difference from unity for the smallest non-zero separation—where the coefficients appear to be around 0.92. This is a measure either of velocities of such small scale as to

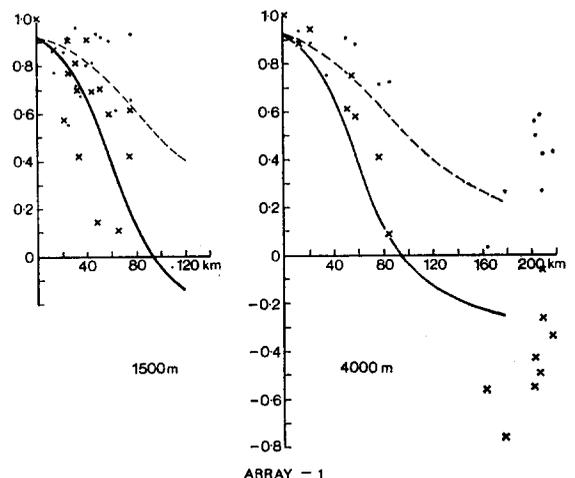


Fig. 6. The Array-I correlation coefficients at 1500 and 4000 m. Points marked \times represent the transverse coefficients and the solid line is an estimate of the transverse correlation function which was used in the simulations. Similarly for the longitudinal correlations represented by the dots and broken line.

be unresolved by this array, or noise due to imperfect sampling. But as was seen in Section 3, from a practical point of view it makes no difference to the present study. The interpolated velocities are smooth on a scale of at least 60 to 100 km, and any smaller features are averaged out or ignored.

Included on Fig. 6 are the correlation functions assumed for the simulation study. The longitudinal function is $R(\rho)$, of equation (28), which is the Fourier transform of the longitudinal velocity spectrum. The transverse function is

$$S(\rho) = \frac{d}{d\rho} [\rho R(\rho)].$$

The cut-off wave number k_0 was chosen to make S pass through zero at $\rho = 95$ km. These functions have been normalized by dividing by the r.m.s. value of 4.1 km day^{-1} for each component. The random noise $E = (1.8 \text{ km day}^{-1})^2$ was adjusted to give a correlation coefficient at small non-zero separation equal to 0.92, given the total observed r.m.s. of 5.8 km day^{-1} ($= 6.7 \text{ cm s}^{-1}$). These were the estimates made at the time, although the r.m.s. speed is probably too large, because of the tendency of current meters on surface moorings to indicate too high.

The frequency spectra in Fig. 7 are subject to considerable uncertainty because of the small number of degrees of freedom. However, they appear to indicate a dominant period of ~ 50 days with much less power in the mean. The original data also show what resembles one cycle of an approximately sinusoidal variation. This feature is reproduced in the simulations if the phase speed C of the propagating Fourier components is taken to be around 5 km day^{-1} .

The sample from which these parameters were estimated is a small one, and it is unclear how many independent degrees of freedom were included. Furthermore, the procedure for obtaining reliable error bounds was not well understood. Accordingly, the first application of the simulated data program was to construct a number of realizations of the space-time velocity field, compute time series for current meters at the same

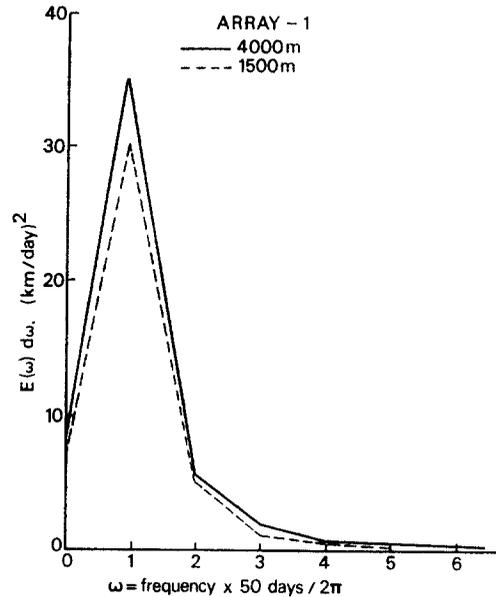


Fig. 7. The Array-I frequency spectra at 1500 and 4000 m determined from 6 and 4 current meters, respectively, over a time interval of 50 days. Approximately 80% of the total energy is contained in the 50-day period.

locations as in Array-I, and then analyse the latter in just the way described above. Now the true statistics are known exactly and the apparent deviations provide a measure of the uncertainty in the original inference.

The longitudinal and lateral correlation coefficients for realization 1 over a 70-day interval are shown in Fig. 8a, using the restricted 6-element array which was all that was available at 1500 m. The authors found these results disconcerting because they indicate fairly conclusively that the zero crossing r_C of the transverse correlation coefficient is less than 50 km, whereas its true value is $r_C = 95$ km. This parameter is critically important, because it influences the horizontal scale of the whole experiment and array design. However, realization 2 told a different story (Fig. 8b), and subsequent realizations did indeed indicate values between 80 and 100 km. Realization 1 is apparently an extreme case—but with a significant probability of occurrence. That it occurred first in a limited sequence is chance—but should be a salutary warning against drawing precise statistical conclusions from a single experiment. In practice,

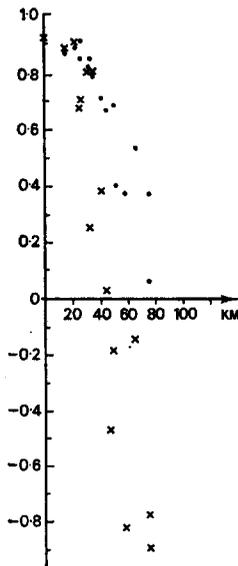


Fig. 8a. Realization 1 depicting the transverse (×) and longitudinal (•) correlation coefficients calculated from the simulated data using 6 data points as in the 1500 m Array-I. Note that the zero crossing of the transverse coefficient is at approximately 50 km.

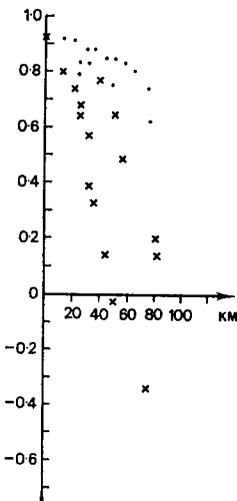


Fig. 8b. Realization 2 as in Fig. 6a. Zero crossing of transverse correlation at about 80 km.

this means that subjective judgement based on *all* available information or prejudice must remain the basis for many planning decisions—though studies of this type can be a major input into these judgements. Fortunately, prior to the deployment of MODE-I, further information from Array 3 became available and tended to confirm the larger

value $r_c = 100$ km which has been assumed for planning purposes. Analysis of the totality of data from arrays 1 and 3 plus the MODE-I data suggest that r_c is about 80 km, somewhat smaller than the value used in planning.

4.3 Application of the theory to the MODE-I array

The next question was how well would the interpolation scheme reproduce the flow field if applied to 'data' from the MODE-I current meter array. The answer presupposes that the estimates of the dominant length scale and noise level made in Section 4.2 are indeed accurate, as well as the more obvious requirement that all the instruments in the array should actually function correctly. Since the latter is improbable some further deterioration in performance is to be expected.

The first form for this answer requires only a knowledge of the array geometry, the spatial correlation functions and the noise level. Equation (10) gives the mean square error in the interpolated estimate for one component of velocity. A more intelligible measure is the r.m.s. error in the total velocity divided by the r.m.s. velocity itself, expressed as a function of position, which is given by equation (29). Figure 9 is a contour map of this

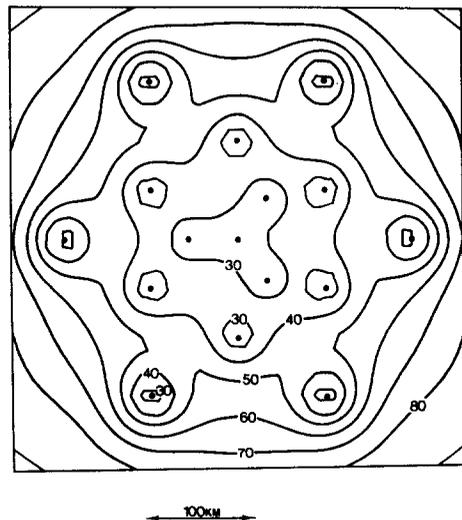


Fig. 9. The per cent r.m.s. error in velocity for the MODE-I array (configuration of dots). The correlation functions shown in Fig. 6 were used in this 'error map'. In the central area where the data coverage is good, the expected error is less than 30%.

quantity expressed as a percentage. Dots mark the positions of the data points and the correlation function $R(\rho)$ used is shown in Fig. 6 (as solid curve). In the central regions where the array is relatively dense, the maximum expected error is around 30%, whereas in the sparser pattern recognition areas it rises to 50%. At first sight these values are large, but it should be recalled that a random noise of 1.3 km day^{-1} in each component superimposed upon a current of r.m.s. speed 6.7 km day^{-1} represents an r.m.s. error of over 25%, so in the inner region a very much more dense array would be needed to give a significant improvement in accuracy. From Array-I this noise level appeared to be intrinsic either to the instrumentation plus mooring system, or to sampling in the presence of small-scale velocity features. It is hoped that these estimates will prove to be pessimistic, otherwise they provide a fairly fundamental limitation on the accuracy obtainable from a current meter array under these circumstances.

For many purposes great accuracy in the estimate of velocity at a point is not required. A test for geostrophic balance is best made by integrating over the horizontal distance between two points to give the stream function, and then comparing with the pressure difference between these points. The percentage error in a measurement of this type is reduced by averaging over several independent current meters. Figure 10 shows the percentage r.m.s. error in the stream function difference between a base point and the general point on the map. Over the central region this is less than 20%.

A more impressionistic method of testing the data array is provided by Fig. 5. Application of the formula given by equation (30) results in the interpolated stream function field shown in Fig. 5b. The data supplied for this reconstructed stream function field came from the original flow field shown in Fig. 5b at the sixteen data positions marked. This was the MODE-I array. The general features of the flow field are reproduced strikingly well, and pattern recognition is achieved everywhere. On the other hand closer investigation does reveal differences, particularly

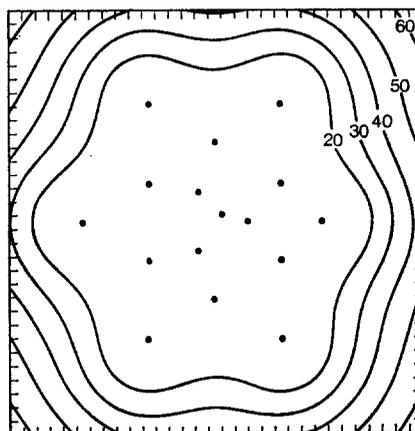


Fig. 10. The per cent r.m.s. error in transport of stream function difference between the central data point and the general point on the map. The expected error in the central region is less than 20%.

in the local gradients, which are symptoms of the errors quantified in Fig. 9.

Figure 11 describes a different test of the array. One objective of the experiment was an inter-

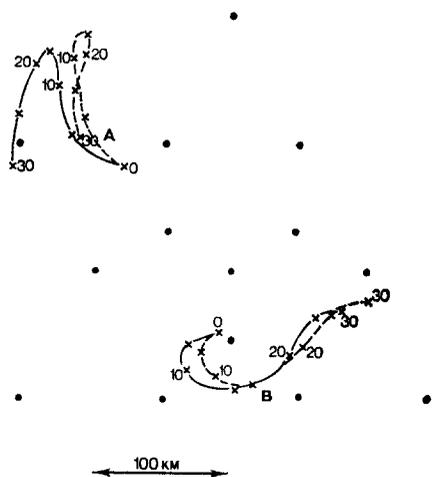


Fig. 11. The 30-day trajectories of two floats A and B marked by the solid curves are calculated from the simulated velocity field. The crosses mark the positions at 5-day intervals. Using the data array marked by the dots the velocity field was objectively reconstructed at 1-day intervals for 30 days and the broken curves representing the trajectories of floats A and B are redrawn from this reconstructed field.

comparison of the trajectories followed by freely moving floats with the velocities indicated by current meters. The true Lagrangian trajectory found by integrating in time the local exact velocity is compared to an estimated trajectory based on interpolations from the simulated data (including random noise) at the indicated current meter positions, starting at the same initial position. Again the general agreement is good, although the original and reconstructed trajectories diverge increasingly with time until after 30 days there was little resemblance. Float A was on the fringe of the array and moved outside it, at which time the most obvious discrepancies arose. Float B was at first within the central accurate mapping region. The differences in detail from the original do indicate significant errors in the reconstruction, which would be reduced only by a more dense coverage of current meters. This is a more sensitive test than comparing stream functions because it depends on point estimates of velocity rather than integrated values. However, it is felt that this approach is a reasonable way of intercomparing current meters and floats.

5. CONCLUSION

Given the statistics of a flow field we have shown how one can construct, objectively, a map of that flow field from a finite sample of data points. The accuracy of reproduction depends on how well the statistics are known, on the density of sampling points, and on the quantity to be mapped. It has been shown that with a knowledge of the exact statistics of a flow field (except for some simulated error) and a realistic density of data points, this technique will reproduce the pattern of flow with a high degree of accuracy. However, it has been pointed out that recovery of the known statistics of the simulated data with a limited sample of data points can lead to substantially different conclusions about the flow field. A continuing effort needs to be made to improve methods of determining flow statistics in the ocean, and only when such methods are available can this technique be applied confidently.

REFERENCES

BATCHELOR G. K. (1960) *Theory of homogeneous turbulence*, Cambridge University Press.
 BREKHOVSKIKH L. M., K. N. FEDOROV, L. M. FOMIN, M. N. KOSHYLYAKOV and A. D. YAMPOLSKY (1971) Large scale multi-buoy experiment in the tropical Atlantic. *Deep-Sea Research*, **18**, 1189-1206.
 CHARNEY J. G. (1971) Geostrophic turbulence. *Journal of Atmospheric Science*, **28**, 1087-1095.
 ELLIOTT A. J., M. R. HOWE and R. I. TAIT (1974) The lateral coherence of a system of thermo-haline layers in the deep ocean. *Deep-Sea Research*, **21**, 95-107.
 GANDIN L. S. (1965) *Objective analysis of meteorological fields*, Israel Program for Scientific Translations.
 GOULD W. J., W. J. SCHMITZ and C. WUNSCH (1974) Preliminary field results for a Mid-Ocean Dynamics Experiment (MODE-0). *Deep-Sea Research*, **21**, 911-931.
 LIEBELT P. B. (1967) *An introduction to optimal estimation*, Addison-Wesley.
 SAUNDERS P. M. (1972) Space and time variability of temperature in the upper ocean. *Deep-Sea Research*, **19**, 467-480.

APPENDIX

Equation (11) can be written as

$$\phi_r = \sum_{r=1}^N \eta_r A_{rr} \tag{A1}$$

Replacing A_{rr} by $A_{rr} + H$, (A1) becomes

$$\begin{aligned} \phi_r &= \sum_{r=1}^N \eta_r (A_{rr} + H), \\ \sum_{r=1}^N \eta_r A_{rr} &= \phi_r - H \sum_{r=1}^N \eta_r, \\ \eta_r &= \sum_{s=1}^N (A^{-1})_{rs} (\phi_s - \bar{\theta}), \end{aligned} \tag{A2}$$

where $\bar{\theta} = H \sum_{r=1}^N \eta_r$ is at present unknown.

Replacing C_{xr} by $C_{xr} + H$, equation (12) becomes

$$\begin{aligned} \hat{\theta}_x &= \sum_r (C_{xr} + H) \eta_r \\ &= \bar{\theta} + \sum_r C_{xr} \left(\sum_s A_{rs}^{-1} (\phi_s - \bar{\theta}) \right), \end{aligned} \tag{A3}$$

which is the same as equation (20).

In order to determine $\bar{\theta}$ we now sum (A2) for $r = 1, \dots, N$.

$$\sum_r \eta_r = \sum_{r,s} (A_{rs}^{-1})_{rs} (\phi_s - \bar{\theta}),$$

i.e.
$$\bar{\theta} = \frac{\sum_{r,s} A_{rs}^{-1} \phi_s}{\sum_{r,s} A_{rs}^{-1} + \frac{1}{H}}$$

Thus

$$\bar{\theta} = \frac{\sum_{r,s} A_{rs}^{-1} \phi_s}{\sum_{r,s} A_{rs}^{-1}} \quad (\text{A4})$$

Considering the formal limit $H \rightarrow \infty$ gives us an expression for $\bar{\theta}$ independent of H .

Note that with $\bar{\theta}$ given by (A4) $\sum_r \eta_r = 0$.