much greater, new idealizations (such as spatial statistical homogeneity) are needed and so-called theories appear to require adhoc leaps. Researchers may prefer to give up and stick with the "old ways," compensating coarseness of numerical models with ad hoc eddy viscosities. This is mentioned in the panel's report as the "simplest and most commonly used approach." It is not purported to be correct; indeed the panel notes observed eddy fluxes that are countergradient (negative mixing). Regrettably it is profoundly unclear what to do about the "eddy problem." For the most part, modellers await bigger computers which may permit smaller eddy viscosities.

Theoretical studies from statistical mechanics suggest that "old ways" are systematically incorrect. The problem with eddy viscosities, diffusivities and "drags" in general is that they tend to draw mean fields toward a state of no relative motion. From theory, this appears most improbable. More relevant is to consider overall entropy of the ocean as a dynamical system, anticipating generalized forces acting on that system in response to gradients of entropy with respect to coordinates of the ocean state. Are these only big words? In fact one can characterize, if only "roughly," the higher entropy configuration of ocean states; then one can anticipate forces which should arise to drive oceans toward those higher entropy states. Very importantly, such generalized forces are not like eddy viscous drag; they may rather be the propelling force behind aspects of observed ocean circulations.

Despite controversy, a possibility for significant improvement on prognostic ocean models has emerged. The implication also carries over to inverse or data-assimilative models. Given partial observation of some aspects of oceans, one employs dynamics to infer other aspects. But then corrupt dynamics lead to corrupt inferences.

Given the uncertain, preliminary and controversial nature of these comments, it is well that the panel's report omits such matters. They appear here as aside remarks. However, as more attention is given to statistics as descriptors of oceans, the more we are moved to consider what dynamics underlie those descriptors. Tenets of the "old ways," even the presumed equations of motion, come up for fresh review.

Ideas mentioned in these comments are not sufficiently mature to warrant extensive literature. Reviews of the basic ideas, as applied in geophysical fluid mechanics, may be read in Salmon (1982), Holloway (1986) or Lesieur (1990). Examples of more recent investigations can be seen in Griffa and Castellari (1991) or Cummins (1992).

Comment

Andrew R. Solow

By necessity, this report, like the ocean itself, is a good deal broader than it is deep. For this reason, the report is unlikely to stir up much controversy. Let me give it a shot anyhow.

While the report provides an admirably panoramic view of data-rich areas within the field of physical oceanography, it does seem a little short on statistics. This is unfortunate, because there is no reason to believe that it will be harder to teach statisticians what they need to know about physical oceanography than to teach physical oceanographers what they need to know about sound statistical practice. In particular, the need to think carefully about a *statistical* model for data is often lost on oceanographers (and other scientists) in their search for methods. Methods are, of course, a dime a dozen. The trick lies in understanding when and why they work and when and why they do not.

A good example of this is the application of principal component analysis to spatial time series. Briefly, consider a random field $Y(x,t)$ where x is location within some region R and t is time. In a typical oceanographic example, $Y(x,t)$ might represent mean annual sea surface temperature at location x. The field is observed over time at a set of locations $x_1, \ldots, x_p$. To reveal spatially coherent temporal variations in the field, it is common practice to extract the first few components from the spatial covariance (or correlation) matrix of the p stations estimated from replications over time and to map the individual station loadings. The oceanographic and meteorological literature is full of this kind of application. One example is given by Jolliffe (1986, page 58).

Andrew R. Solow is Associate Scientist at Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543.

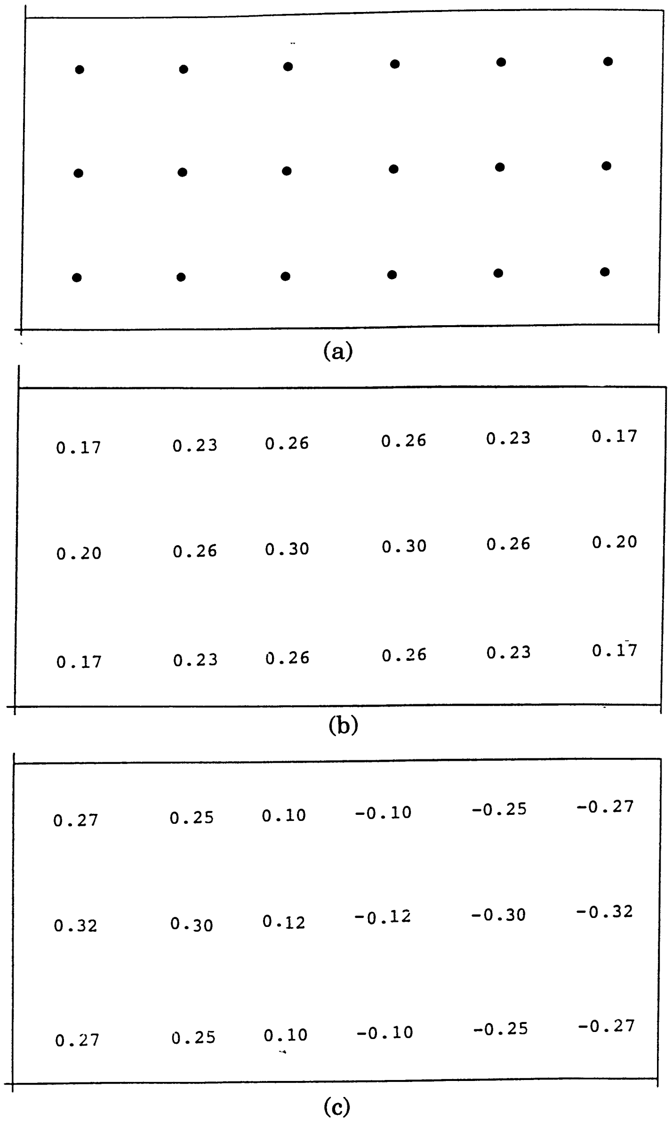

FIG. 1. (a) *Location of observing locations within the region R*; (b) *station loadings in the first component*; (c) *station loadings in the second component.*

To see what can happen, suppose that the random field is stationary with spatial covariance $C(h) = \exp(-0.69h)$. Here, $C(h)$ is the covariance of the random field at locations separated by distance h. The form of $C(h)$ gives a pretty good idea about the spatial variation of the field: it is continuous, but not mean square differentiable, with positive correlation to a range of about 4. Suppose that the observing stations lie on a regular 3-by-6 grid with unit grid spacing within a rectangular region R (Figure 1a). The station loadings for the first component (which accounts for 29% of the variance) are shown in Figure 1b. These loadings suggest a regionwide coherent mode of variation (i.e., sea surface temperatures tend to fluctuate together across the basin). This is qualitatively consistent with what one would expect from the form of $C(h)$. However, Figure 1c shows the station loadings for the second component (which accounts for a further 15% of the variance). These loadings suggest the existence of an oscillation along the long dimension of R. The effect of El Niño, perhaps, or sunspots. There is, of course nothing *mathematically* incorrect with decomposing the spatial variation of the field in this way, but the inevitable *physical* interpretation of the components can be very misleading.

No sampling was involved in the example outlined above. In practice, of course, $C(h)$ is unknown and must be estimated from data. In that case, a decision must be made about how many of the p components to retain for interpretation and analysis. This is another problem where a lack of careful statistical thinking has led to the adoption of questionable methods. This problem is commonly addressed using the following method, which is also described in Jolliffe (1986, page 102). Suppose that observations at the p observing stations are taken in n years. The distribution of the ordered eigenvalues of the sample covariance matrix of p independent white noise series of length n is estimated by simulation. Starting with the first observed component, components are retained provided the corresponding eigenvalues exceed, say, the 0.95-quantiles of their marginal sampling distributions under independence. The process terminates when a component is not retained.

To see why this method is *fundamentally* incorrect, suppose that the first component is determined to be significant. This is equivalent to a determination that the station records are not independent. However, the significance of the second component is determined by comparing it to its sampling distribution *under independence*. In effect, this is a second test of independence. It is, however, an incorrect test, because it does not condition on the value of the first eigenvalue. Thus, this rule consists of a sequence of incorrect tests of the wrong hypothesis.

Part of the problem here is that insufficient thought has gone into specifying exactly what it means for a component to be significant. Under the approach proposed by Bartlett (1950), the last $p - q$ components should be discarded if the corresponding eigenvalues are not distinct. One justification for this approach is that the last $p - q$ components are not unique in this case. Unfortunately, the distinctiveness of eigenvalues does not capture the notion of significance in the minds of oceanographers. A different approach would be to discard the last $p - q$ components if the projection of the original data onto them is without temporal structure. While this idea can be formalized, the implicit recognition that the search is for temporal structure (and not variance) brings this whole use of principal component analysis into question. Other methods (e.g., Shapiro and Switzer, 1989) may be more appropriate.

A final comment. It was disappointing to find no real mention of Bayesian methods in this report. There is some irony here, since a number of commonly used methods (kriging, for example) have a strong Bayesian flavor. In any case, whatever their predilections, statisticians must recognize that there have been enormous advances in practical Bayesian methods. Some of us actually use them!

Comment

Hans von Storch

1. GENERAL

Being in the process of preparing a monograph on "statistical analysis in climate research," I was intrigued by the title of the report of the National Research Council on the present use and future need of statistics in physical oceanography. But after having gone through it I became rather disappointed—apparently these people had a "physical oceanography" in mind which had hardly any overlap with the type of problems which I meet in my own research. Relevant topics were not mentioned, such as the variability of the thermohaline circulation (note that the deep ocean was excluded in Figure 2.1 of the report) and its implications for the global climate. Influential names, such as Frankignoul, did not appear. Fundamental papers, such as that of Hasselmann (1976) on stochastic climate models, were not cited. The data assimilation issue related to preoperational predictions of the oceans were not sufficiently taken into account (see Derber and Rosati, 1989, or Mellor and Ezer, 1991). I could not even identify the members of the committee who supposedly represent the community of physical oceanographers.

The solution to this inconsistency is likely that neither the committee nor I—and my colleagues whom I have contacted in this matter—represent the full spectrum of statistical thinking in what is called physical oceanography. I have to admit that I am in touch with only a narrow window of the spectrum, namely, that part with relevance for the dynamics of climate and for the concept of climate change. In the following I will go through a number of examples of statistical thinking in our field. These examples have been encountered by the Climate Dynamics and Oceanography division of the Max-Planck-Institut für Meteorologie in Hamburg, in the past.

Hans von Storch is with the Max-Planck-Institut für Meteorologie, Hamburg, Germany.

2. THE IDENTIFICATION OF DYNAMICAL SUBSYSTEMS

The dynamics of the ocean operate on a large phase space with spatial and temporal scales spanning a wide range. The sheer amount of information, representing the state of the ocean during any well-documented interval of time, inhibits any complete description of the oceanic dynamics—independently if we work with observed or simulated data. Therefore it is advisable, or even required, to split the full phase space into a "signal" subspace and a "noise" subspace. The definition of the two subspaces depends, of course, strongly on the considered problem: The physically significant part of the dynamics span the "signal" subspace whereas the "noise" subspace comprises those processes which contribute to the dynamics only through their overall statistics and not through their details. The identification of such dynamical subsystems represents a major challenge for ocean sciences.

2.1 Stochastic Climate Models

In the "stochastic climate model" approach (Hasselmann, 1976) the separation into signal and noise subspaces is done by means of time scales. The "high-frequency" part is considered as "noise" whereas the "low-frequency" part is understood as being the dynamical response to the "noise." To keep the system stationary, negative feedback must prevail in the "signal" subspace.

This concept has been applied to modeling the dynamics of sea-ice variability (Lemke, 1977) and of sea-surface temperature variability (Frankignoul and Hasselmann, 1977; Ortiz Bévia and Ruiz de Elvira, 1985; Herterich and Hasselmann, 1987). More recently the concept has been used in a "stochastically forced" ocean general circulation model experiment (Mikolajewciz and Maier-Reimer, 1990). In this run the ocean model was forced with climatological conditions without any temporal vari-