Comments on "The Number of Significant Proper Functions of Two-Dimensional Fields"

HARRY R. GLAHN

Techniques Development Laboratory, NWS/NOAA, Silver Spring, MD 20910

19 September 1978

There is one point concerning Buell's (1978) "self-consistency" method of determining the number of statistically significant proper functions on which I wish to comment. Suppose a sample of data is drawn from a population and a 40×40 correlation matrix computed. Also, suppose another sample is drawn from the same population and another 40×40 correlation matrix computed. All 40 eigenvalues and corresponding eigenvectors are then found for each matrix. (There could be less than 40, but seldom would this happen with real data unless the sample size were rather small.) The eigenvectors for each matrix are then ordered according to the corresponding eigenvalues. Each set of values and vectors will be equal to the population values and vectors plus a random component. Because of this random component, the two ordered sets may not "match" exactly. For instance, if the seventh and eighth eigenvalues are (nearly) equal, the seventh eigenvector calculated from data set one may "match" the eighth eigenvector calculated from data set two, and vice versa. I believe this same thing will likely happen with Buell's quadrature techniques.

In fact, evidence for this is found in his Figs. 5 and 6. Each of the four curves shows a tendency for two successive points to be located low on the graph (small correlation). The points connected by solid lines in Fig. 6 demonstrate this most vividly. Note the pairs composed of points 7 and 8, 13 and 14, 22 and 23, 29 and 30, and 37 and 38. Each point in these pairs has nearly the same correlation value as the other point (~0.1 for points 7 and 8) as would be expected for a pairwise mismatch, and these correlations are much lower than the correlations for adjacent points.

Although I do not necessarily disagree with Buell's conclusion concerning the data he presented, I feel a pairwise mismatch such as I have described does not invalidate the corresponding vectors. That is, this mismatch alone should not cause them to fail a significance test.

In practice, a pairwise mismatch (three or more points *could* be involved in the shuffle) will usually occur after the first few functions, so Buell's technique may still furnish useful results. One *could* reorder one set of vectors so that the mismatched pairs are interchanged and then calculate the correlation. The appearance of the curves would probably change considerably.

I find it puzzling that the general level of the correlation curves does not decrease with index number. For instance, one might expect function 40 to be mostly noise and have a low correlation with function 40 calculated in any other way. However, the opposite seems to be true; in Fig. 6 (solid lines), functions 31–40 (except for a pairwise mismatch) all have high (above 0.8) correlations, while functions 9–21 (including a pairwise mismatch) have lower correlations. Perhaps the author has an explanation for this.

REFERENCE

Buell, C. E., 1978: The number of significant proper functions of two-dimensional fields. *J. Appl. Meteor.*, **17**, 717–722.