An Objective Analysis of the Boundary-Layer Thermodynamic Structure During GATE. Part I: Method

J. B. JALICKEE AND C. F. ROPELEWSKI

Environmental Data and Information Service, NOAA, Center for Environmental Assessment Services, Washington, DC 20235

(Manuscript received 8 August 1978, in revised form 25 October 1978)

ABSTRACT

A method for the analysis and objective classification of meteorological data is presented and illustrated. The first step in this approach is the application of an efficient method to describe the data both qualitatively and quantitatively. This approach, based on the singular decomposition theorem, extracts the important features from the data and is similar to principal component or empirical orthogonal function analysis. The extracted features are then combined in the classification step to form new features, each of which describes a subset of the data. The mathematical details are given and computational considerations are discussed.

The method is illustrated with a set of boundary-layer potential temperature profiles derived from structure sonde data taken during the GARP Atlantic Tropical Experiment (GATE). This set of profiles, taken at 3 h intervals over a 20-day period, was objectively classified into two distinct groups. Time series of surface divergence and radar-estimated precipitation rates are used to demonstrate that one group of profiles is associated with a convectively modified boundary layer, while the other group of profiles is more typical of undisturbed situations.

Other uses and limitations of the objective classification method are discussed.

1. Introduction

Summarization is one of the first, and sometimes the most important, step in the process of analyzing large sets of data from observational systems. Generally, it is this step which often provides a first impression and overview of important features, and which may open up new topics of investigation, in addition to those planned. Methods suitable for summarization range from elementary to sophisticated, but the utility of any one method must be judged in terms of its overall effectiveness in describing essential features insofar as these features can be assimilated and understood by users. A sophisticated approach yielding very complex results can be no more useful than a very elementary approach that yields almost no pertinent information.

The most popular method of summarization is that of averaging or forming means over suitably chosen intervals of space and/or time. Just how effective this method proves to be, however, is critically dependent on the choice of averaging interval, since the implicit assumption in this method is that the data in some way are realizations of the same random process. For example, when the data sets consist of atmospheric observations taken from different weather regimes, a single average might give a misleading picture. However, if the regimes are first segregated and averages then taken, this composite summarization might prove very useful in establishing overall meaningful summaries.

In forming composites, means or other "typical" representations of data for two or more different classes or regions that are meant to describe and distinguish basic physical processes, we may use as guides mathematical relationships implicit in known physical descriptions of those processes, numerical or conceptual models, or simply characteristic patterns that can be extracted from and are completely defined by the data themselves. In some instances the bases for forming composites or means are not clearly indicated and are based on one or more subjective criteria, biasing the results. In other instances, the complexity of the underlying physical processes and the large numbers of observations weaken the effectiveness of standard approaches.

The analysis and objective classification scheme described here has the property that the data set itself completely determines those features which are implicitly characteristic of the underlying physical process manifested through the observations, and which form the basis for classification and definition. The only free parameters available in this scheme are the number of classes required and the features used.

In this paper, the technique is described and il-

lustrated with a set of atmospheric profiles, followed in Part II (Ropelewski and Jalickee, 1979) by a comprehensive analysis based on the technique.

Section 2 describes singular decomposition, which is used to extract essential features from a data array. The results form the basis for the classification, also known as empirical orthogonal functions or principal components. The second step of the method (objective classification) is presented in Section 4. It consists of a generalization of the method of "factor rotation," commonly used in factor analysis, and was described with application to classifying oceanographic data by Jalickee and Hamilton (1977). In this step the extracted features are recombined to form mutually exclusive classes. The new aspect of this procedure, in contrast to traditional practices, is that any number of groups can be formed from any number of features.

2. Singular decomposition

The singular decomposition theorem (see e.g., Noble and Daniel, 1969) is the foundation of the first phase of the analysis—the extraction of the essential features of a given array of data. This theorem formally relates, to any matrix, a series of terms which are ordered in such a way that the important characteristics of the matrix are contained in the leading terms with successive terms describing less of the main features and more of the specific details. The leading terms are used to approximate the original array and this approximation can be shown to be optimum in the least-squares sense. An efficient compaction of the data set is thus achieved and, as a result, the meaningful information is separated from the redundancy and "noise" inherent in the data.

The method of empirical orthogonal functions that Lorenz (1956) introduced to study characteristic patterns in meteorological fields is similar to this approach as are techniques of principal components and factor analysis. The distinction is mainly that the singular decomposition 1) puts the rows and columns of the data array, space and time indices for example, on an equal footing and 2) lends itself to computational procedures which are numerically more efficient and accurate. This latter distinction becomes more critical as larger data sets are analyzed. The mathematical similarities will be discussed briefly at the end of this section.

The singular decomposition theorem states that every $N \times M$, $N \leq M$, matrix $\mathbf{A} = \{A_{i,j}\}$ can be represented by a series of the form

$$A_{i,j} = \lambda_1 x_i^{(1)} y_j^{(1)} + \lambda_2 x_i^{(2)} y_j^{(2)} + \ldots + \lambda_N x_i^{(N)} y_j^{(N)}, \quad (1)$$

where the $\{x_i^{(k)}\}_{i=1,N}$ is known as the kth left singular vector, $\{y_j^{(k)}\}_{j=1,M}$ as the right singular vector, and the λ's are the singular values with

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0. \quad (2)$$

These vectors are mutually orthonormal, that is

$$\sum_{i=1}^{N} x_i^{(m)} x_i^{(n)} = \sum_{j=1}^{M} y_j^{(m)} y_j^{(n)} = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases} \quad (3)$$

The leading terms of the singular expansion can be effectively used to approximate $\mathbf{A}$. If we let

$$A_{i,j} \approx A_{i,j}^{(L)} = \sum_{k=1}^{L} \lambda_k x_i^{(k)} y_j^{(k)}, \quad (4)$$

then it can be shown that $\mathbf{A}^{(L)}$ is an optimum approximation to $\mathbf{A}$ in the least-squares sense, i.e.,

$$\sum_{i=1}^{N} \sum_{j=1}^{M} [A_{i,j} - A_{i,j}^{(L)}]^2 = \text{minimum} \quad (5)$$

with respect to all choices of $\{\mathbf{X},\mathbf{Y},\boldsymbol{\lambda}\}$.

Since the vectors are orthonormal, it can also be seen that the variance explained by each term is in proportion to the magnitude of the respective singular value since

$$\sum_{i=1}^{N} \sum_{j=1}^{M} [A_{i,j} - A_{i,j}^{(L)}]^2 = \sum_{i=1}^{N} \sum_{j=1}^{M} A_{i,j}^2 - \sum_{k=1}^{L} \lambda_k^2. \quad (6)$$

Good (1969) has discussed this theorem as well as several different general applications. Gabriel (1972) and Brier and Meltesen (1976) have used the singular decomposition theorem for analysis of meteorological data and prediction of time series. Golub and Reinsch (1970) have discussed numerical aspects of the mechanics of the decomposition and developed an accurate and efficient calculation procedure. The numerical technique which we call asymptotic singular decomposition (ASD) is specifically designed for calculating the leading terms of this expansion when the array is very large. The ASD method has been described by Jalickee and Klepczynski (1977).

The similarity between the singular decomposition and the methods of empirical orthogonal functions (EOF) and principal components follows from the fact that the right singular vectors $\mathbf{Y}$ are just the eigenvectors of the product $\mathbf{A}^T\mathbf{A}$ and for $\mathbf{A}\mathbf{A}^T$ the eigenvectors are the left singular vectors $\mathbf{X}$, both with corresponding eigenvalues given by the respective λ^2. The EOF's or principal components are calculated, as a matter of course, from eigenanalysis of either $\mathbf{A}^T\mathbf{A}$ or $\mathbf{A}\mathbf{A}^T$ whichever has the smaller dimension. The companion vectors can then be calculated from pre- or post-multiplication of the eigenvectors with the data matrix. However, it may be shown that this procedure is quite vulnerable to round-off error problems when the rows (or columns) of the data matrix are highly correlated (see Golub and Reinsch, 1970). This consideration is especially critical when the arrays are large in one or both dimensions.

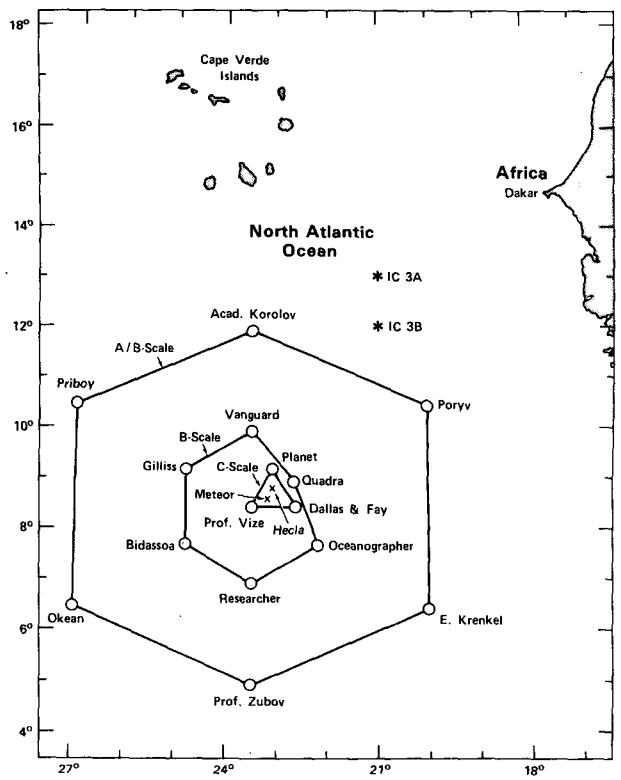

FIG. 1. C-array GATE Phase III.

3. Data description and analysis

The potential temperature profiles discussed below were derived from structure sonde data gathered during Phase III, 30 August to 19 September 1974, of the Global Atmospheric Research Program's (GARP) Atlantic Tropical Experiment (GATE). The structure sondes have an ascent rate of 3 m s^{-1} and a sampling rate 0.5 hz to provide high-resolution measurements of pressure, wet-bulb temperature and dry-bulb temperature from the surface to 700 mb. (WMO, 1975). These sondes were scheduled to be launched at 3 h intervals starting at 0100 h from each of the ships that formed the C-array, the *Fay, Meteor* and *Planet* (Fig. 1).

The data, obtained from the Boundary Layer Subprogram Data Center, Hamburg, FRG, were interpolated to 5 mb intervals starting at 1010 mb. We have restricted our analysis to the 24 levels from the surface to 895 mb. Profiles were excluded from the analysis if the highest pressure for a sounding was less than 1010 mb. A total of 122 soundings were included in the analysis.

To illustrate the method, we discuss the analysis and objective classification of potential temperature profiles derived from structure-sonde data taken by the *Fay*. A detailed analysis of potential temperature profiles from the other ships as well as an analysis of specific humidity and moist static energy profiles are found in Ropelewski and Jalickee (1979).

The 24 by 122 array of potential temperature data was analyzed by the ASD technique. With the first two terms of the singular expansion, 99.8% of the variance was accounted for, i.e., if the 122 profiles are approximated by those reconstructed from the two leading terms, only 0.2% of the original variance remains to be accounted for. The left vector of the first term $\mathbf{x}^{(1)}$ from Eq. (1), which is shown in Fig. 2, can be thought of as the characteristic profile since to first approximation each profile of the data set is proportional to it, the constant of proportionality being $[\lambda_1 y_j^{(1)}]$ for the jth profile. This characteristic profile is generally quite similar to the mean profile determined by averaging the data

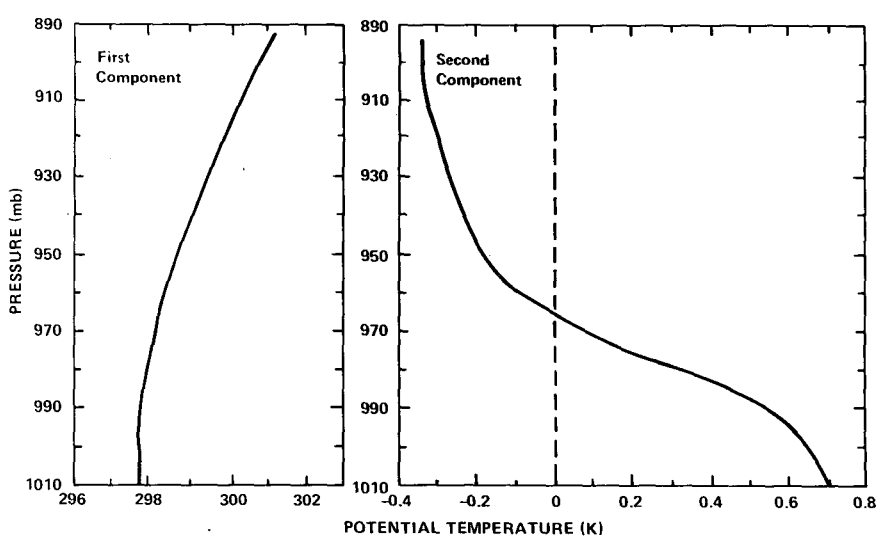

FIG. 2. Leading ASD components of the potential temperature data set scaled to show variations in degrees.

level by level. On the other hand, the $y^{(1)}$'s describe the characteristic sounding dependence or time dependence of the data set since the profile data are ordered in time. It is the combination of the two characteristics that provides the best first approximation to the entire data set.

The second characteristic profile $\mathbf{x}^{(2)}$, which is also shown in Fig. 2 and in effect is added to or subtracted from the first, represents the major variations about the first; similarly with $\mathbf{y}^{(2)}$. Thus, these two terms provide a great deal of information about the entire data set in a compact and efficient way.

Generally, when further terms are included in the approximation, the similar interpretations become cumbersome, but they can still provide a great deal of information about the characteristics of the data. In the classification that follows, the information contained in the approximation is presented in a different and perhaps in a more meaningful way.

4. Classification

The second step in the analysis is to rearrange the leading terms of the approximation to form characteristic shape functions. These functions are intended to be characteristic of mutually exclusive subgroups of the data, in contrast to the original components, which are ordered in importance with respect to the entire data set. The rearrangement is made to preserve the total approximation and the new shape functions are generally not independent or orthogonal. The classification is first derived heuristically.

The vectors of the approximation defined previously by Eq. (4) can be rotated so that the approximation is preserved. Let

$$A_{i,j}^{(L)} = \sum_{k=1}^{L} \lambda_k x_i^{(k)} y_j^{(k)} = \sum_{n=1}^{L} \hat{a}_i^{(n)} \hat{y}_j^{(n)}, \tag{7}$$

where $\hat{\mathbf{a}}^{(k)}$ will be called the kth characteristic shape function, and the jth element of new right vector $\hat{y}_j^{(k)}$ is the coefficient of this function, j referring to the jth data profile of the matrix $\mathbf{A}$. The new right vectors will remain normalized but not necessarily orthogonal.

In developing the particular transformation required for producing the best possible classification, we consider for convenience, only two groups ($L = 2$), and later generalize to arbitrary numbers of groups.

Suppose that $\hat{\mathbf{a}}^{(1)}$ and $\hat{\mathbf{a}}^{(2)}$ are two shape functions, each best describing a different subset of the data to the exclusion of the other. Then, when the jth profile, belonging to group 1, say, is being approximated by $\mathbf{A}^{(2)}$, shape function 2 can be neglected and

$$A_{i,j}^{(2)} \propto \hat{a}_i^{(1)}, \quad j \in \text{group 1}.$$

On the other hand, if j' should denote a member of group 2, the approximation should be

$$A_{i,j}^{(2)} \propto \hat{a}_i^{(2)}, \quad j' \in \text{group 2}.$$

This is precisely what happens when the coefficients have the properties

$$|\hat{y}_j^{(1)}| \gg |\hat{y}_j^{(2)}|, \quad j \in \text{group 1},$$
$$|\hat{y}_j^{(2)}| \gg |\hat{y}_j^{(1)}|, \quad j \in \text{group 2}.$$

A mathematical statement of this requirement is that the quantity

$$\sum_{j=1}^{M} [\hat{y}_j^{(1)} \hat{y}_j^{(2)}]^2 = \text{minimum},$$

subject to the constraint that the $\hat{y}$'s are normalized. This is the criterion we use to find the particular transformation.

For the general case, we require that the new coefficients satisfy the condition

$$\sum_{\substack{m,n=1 \\ m \neq n}}^{L} \sum_{j=1}^{M} [\hat{y}_j^{(m)} \hat{y}_j^{(n)}]^2 = \text{minimum}. \tag{8}$$

Note that when exclusive groups do exist within the approximation, only one coefficient is nonzero for each j and the minimum is identically zero. Once these coefficients have been calculated, the characteristic shape functions are uniquely determined by Eq. (7).

Having discussed the motivation for choosing the above function to define our new coefficient, we next describe the quantities in mathematical form, and then develop the numerical technique to solve for the new coefficients and characteristic shape functions.

The rotation of the y's is defined in terms of the matrix $\boldsymbol{\alpha}$,

$$\hat{y}_j^{(m)} = \sum_{k=1}^{L} \alpha_{m,k} y_j^{(k)}, \tag{9}$$

and the rotation of the $\mathbf{x}$'s in the terms of the matrix $\boldsymbol{\beta}$,

$$\hat{a}_i^{(n)} = \sum_{k=1}^{L} \lambda_k x_i^{(k)} \beta_{n,k}. \tag{10}$$

By substituting these equations into Eq. (7) and performing the necessary operations, we see that to preserve the approximation $\boldsymbol{\beta}$ must be $\boldsymbol{\alpha}^{-1}$, i.e.,

$$\sum_{k=1}^{L} \beta_{n,k} \alpha_{m,k} = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases} \tag{11}$$

The minimum function and the condition that the new coefficients remain normalized can be combined into one equation with the introduction of a set of Lagrange multipliers $\{\mu_m\}$. The function becomes

$$\sum_{\substack{m,n=1 \\ m\neq n}}^{L}\sum_{j=1}^{M}[\hat{y}_j^{(m)}\hat{y}_j^{(n)}]^2$$

$$+\sum_{m=1}^{L}\mu_m[1-\sum_{j=1}^{M}\hat{y}_j^{(m)}\hat{y}_j^{(m)}] = \text{minimum }\{\alpha\}. \quad (12)$$

The solution to this equation is not straightforward and in practice we find that numerical instabilities can arise when some general purpose computational procedures are used, often because of the nature of the data and typically high correlations between different profiles. The method described here is efficient for solving this particular problem and is relatively immune to numerical instabilities. It is an iteration procedure based on the observation that this minimum function is quadratic in each of the $\hat{y}$'s taken one at a time, and thus the local minimum (with respect to one particular $\hat{y}$) can be found by solving an eigenvalue problem, i.e., any particular $\hat{y}$ can be calculated if we know the others. We begin the procedure by making initial estimates for $\hat{y}^{(2)}$, $\hat{y}^{(3)}$, . . . , $\hat{y}^{(L)}$ and solving for $\hat{y}^{(1)}$; then for each $\hat{y}^{(m)}$, m = 2, 3, . . . , L, using our available estimates of $\hat{y}^{(n)}$, $n \neq m$. This gives us our first set of calculated $\hat{y}^{(m)}$'s, $m = 1, 2 \ldots, L$. This process is repeated using each set of previous estimates of the $\hat{y}$'s to calculate new $\hat{y}$'s until convergence of these coefficients is attained.

The procedure is formalized by deriving the necessary equations for the minimum function, given by Eq. (12), in terms of the individual components. Let m denote the particular $\hat{y}^{(m)}$, for which we wish to solve, assuming we know $\hat{y}^{(n)}$, $n \neq m$. By substituting Eq. (9) into Eq. (12) and rearranging terms, we find that the function to be minimized locally, i.e., with respect to the α_m's with the α_n's, $n \neq m$, held fixed, can be written in the form

$$\sum_{k,k'=1}^{L}\alpha_{m,k}\Lambda_{k,k'}^{(m)}\alpha_{m,k'} + \mu_m(1-\sum_{k=1}^{L}\alpha_{m,k}^2) = \text{minimum }\{\alpha_{m,k}\}_{k=1,L}, \quad (13)$$

where

$$\Lambda_{k,k'}^{(m)} = \sum_{\substack{n=1 \\ n\neq m}}^{L}\sum_{j=1}^{M} y_j^{(k)}[\hat{y}_j^{(n)}\hat{y}_j^{(n)}]y_j^{(k')}.$$

The required elements satisfy the set of L eigenvalue equations

$$\sum_{k'=1}^{L}\Lambda_{k,k'}^{(m)}\alpha_{m,k'} = \mu_m\alpha_{m,k}, \quad k = 1, L, \quad (14)$$

and are simply the elements of the normalized eigenvector of $\Lambda^{(m)}$ corresponding to the smallest eigenvalue. This may be seen by differentiating Eq. (13) with respect to $\alpha_{m,k}$, $k = 1, L$. The smallest is required since the quadratic form is to be minimized. The singular decomposition routine is used to calculate these elements.

The singular decomposition routine can also be used to compute the inverse of a matrix $\boldsymbol{\beta}$; in this case, for the calculation of the typical profiles defined in Eq. (10). If the singular decomposition of $\boldsymbol{\alpha}$ is

$$\alpha_{m,k} = \sum_{n=1}^{L}\kappa_n\gamma_m^{(n)}\delta_k^{(n)}, \quad (15)$$

then

$$\beta_{m,n} = \sum_{n=1}^{L}\kappa_n^{-1}\gamma_m^{(n)}\delta_k^{(n)}. \quad (16)$$

These equations essentially form the basis for our computational procedure.

In practice, the application of this classification procedure sometimes produces two identical characteristic shape functions. One way this can happen is that, although the terms of the original approximation adequately describe the data array, the distribution of coefficient values for one of the groups is random and, therefore, a new group cannot be determined. Consider the hypothetical case where the entire data array is completely determined by two terms of the singular expansion but the values of $y_j^{(1)}$ and $y_j^{(2)}$ are randomly distributed, i.e., in the two-dimensional space defined by ($\mathbf{y}^{(1)}$, $\mathbf{y}^{(2)}$) the points are scattered. No groupings are possible, and the application of the classification procedure seemingly fails, since we have assumed in describing this classification procedure that there are as many groups as there are terms of the approximation. But, when this is not the case, a simple modification allows us to proceed with the determination of the groups that are different. When two groups are identical, then the two sets of transformation elements corresponding to these groups will be identical. In this case, the transformation is not of full rank, the determinant of $\boldsymbol{\alpha}$ is zero, the inverse does not exist, and therefore the characteristic shape functions cannot be determined. This situation is resolved by introducing the pseudo-inverse of $\boldsymbol{\alpha}$.

When $\boldsymbol{\alpha}$ is of rank $L' < L$, i.e., there are only L' different meaningful groups, then in Eq. (15)

$$\kappa_n = 0, n = L' + 1, L' + 2, \ldots, L.$$

The pseudo-inverse of $\boldsymbol{\alpha}$ then is simply given by

$$\beta_{m,k} = \sum_{n=1}^{L'}\kappa_n^{-1}\gamma_m^{(n)}\delta_k^{(n)}. \quad (17)$$

In this situation, the approximation is not preserved because of the $L - L'$ missing groups, but the classification results can still be meaningful.

Although the case of not full rank occurs infrequently, it can be treated, almost trivially, in this way. To our knowledge, this has not previously been considered in the literature. The implication of this result is that a classification into N groups using M

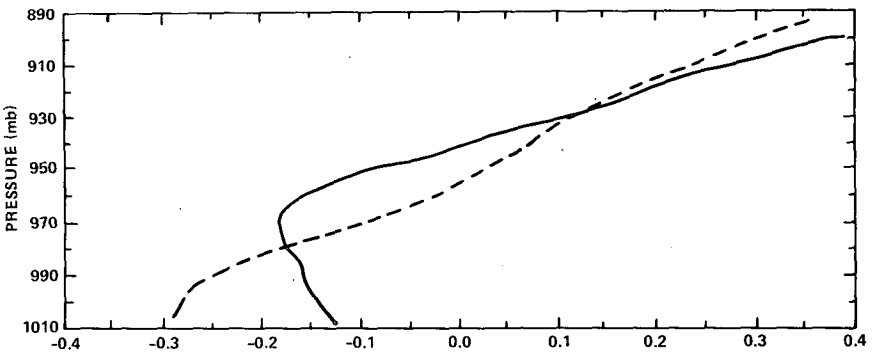

FIG. 3. Characteristic shape functions resulting from the classification for the potential temperature profiles: solid line, undisturbed; dashed line, disturbed.

singular terms can be made with no restrictions on N or M.

5. Classification example

The power of the objective classification scheme is that it identifies distinct profile shapes, while the ASD analysis provides the independent components (e.g., Fig. 2) that combine to describe the data set as a whole. An analysis of potential temperature profiles was chosen to illustrate the objective classification because we expect clear differences in the shapes of these profiles as a function of convective activity. Betts (1976), Seguin and Garstang (1976), Houze (1977), Zipser (1977), as well as others, have reported a well-mixed layer of constant potential temperature below cloud base for undisturbed conditions, i.e., no convective activity, and stable, monotonically increasing, potential temperature profiles associated with disturbed, i.e., convectively active, conditions. Thus we can restrict the example to an analysis of two objectively defined classes to keep the interpretation simple.

The 122 potential temperature profiles were analyzed according to the method presented in the previous section. The characteristic shape functions resulting from the classification are shown in Fig. 3. The time series of coefficient pairs $\hat{y}_j^{(1)}$ and $\hat{y}_j^{(2)}$ for the period 1630 GMT 1 September–2200 GMT 18 September are shown in Fig. 4. It is the combination of the characteristic shape function and coefficient pairs that reproduce a particular profile. There is a coefficient pair corresponding to each of the potential temperature profiles. If a profile is missing for any 3 h period, the plots are interpolated. In general, we see that when one of the coefficients for a particular sounding are near zero, the other coefficient tends to be large. The profiles, then, for the most part, are quite accurately described by the mean profile corresponding to the larger coefficient. The mean profiles (Fig. 5) are formed by averaging profiles that belong to the same class, i.e., all profiles for which $y_j^{(1)} > y_j^{(2)}$ are averaged to form the mean profile for class 1 and likewise for class 2. We note in Fig. 6 that the individual members of a class have very similar shapes to the mean profile for that class. These data profiles were taken from the region indicated on Fig. 4. Note that in this sample the profiles go from a disturbed, through a transition, to an undisturbed state. If the data consisted of two mutually exclusive groups, then the coefficient of one or the other would always be zero in Fig. 4 and the mean profile corresponding to the other class would be the data profile (to within a scale factor related to the nonzero coefficient). In regions where the magnitude of the coefficients are approximately equal, the profiles do not clearly fall

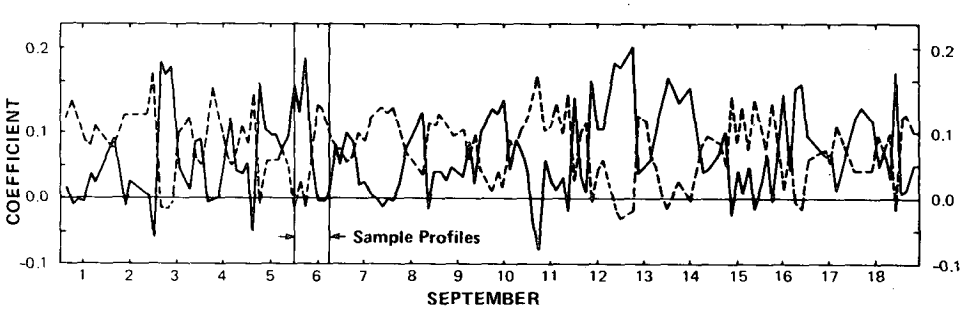

FIG. 4. Time series of classification coefficient pairs $\hat{y}^1$ and $\hat{y}^2$: dashed line, undisturbed; solid line, disturbed.

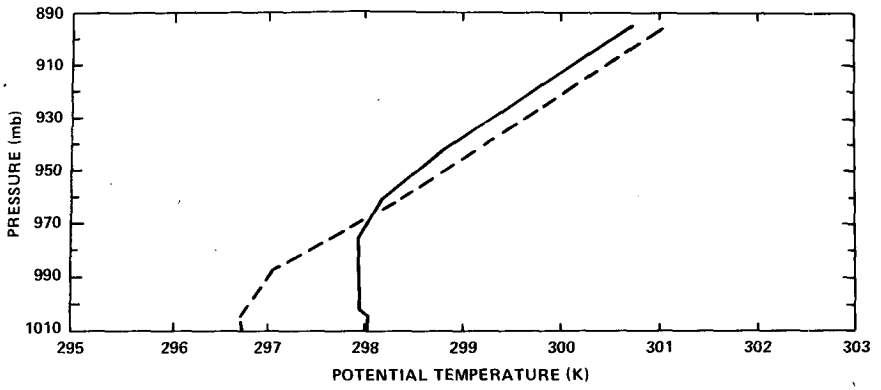

FIG. 5. Mean potential temperature profiles: solid lines, undistrubed; dashed lines, disturbed.

into either class. These regions may correspond to periods of transition or they may be indicative of the existence of a third class of profiles. If the transition periods constitute a large fraction of the whole analysis period, then analysis with more classes is indicated. The actual number of classes chosen for the analysis of each data set must be, of course, constrained by the physical processes being analyzed. It is clear from the time series of potential temperature profile coefficients that for the data presented here two classes are adequate most of the time. Furthermore, the mean potential temperature profiles have shapes that are consistent with those found for undisturbed and disturbed conditions in the previous studies discussed above.

An analysis of potential temperature profiles derived from tethered sonde data (not shown) and data from the two other C-array ships yield similar potential temperature profile shapes (Ropelewski and Jalickee, 1979). Thus, these objectively derived profiles are not only consistent with profiles from other studies, but they are also repeatable with other data sets and not an artifact of one particular data set.

We now demonstrate that the variation in the coefficients, $\hat{y}_j^{(1)}$, $\hat{y}_j^{(2)}$, are related to independent estimates of the strength of convective activity and therefore can be used as a basis to classify and composite other data. The mean profile corresponding to the class 2 coefficients has a shape identified, for instance, by Seguin and Garstang (1976), and with GATE data by Gaynor and Ropelewski (1979), with a disturbed atmosphere in the wake of convective activity. One independent estimate of the strength of the convective activity over the C-array is the radar estimated precipitation rate over the

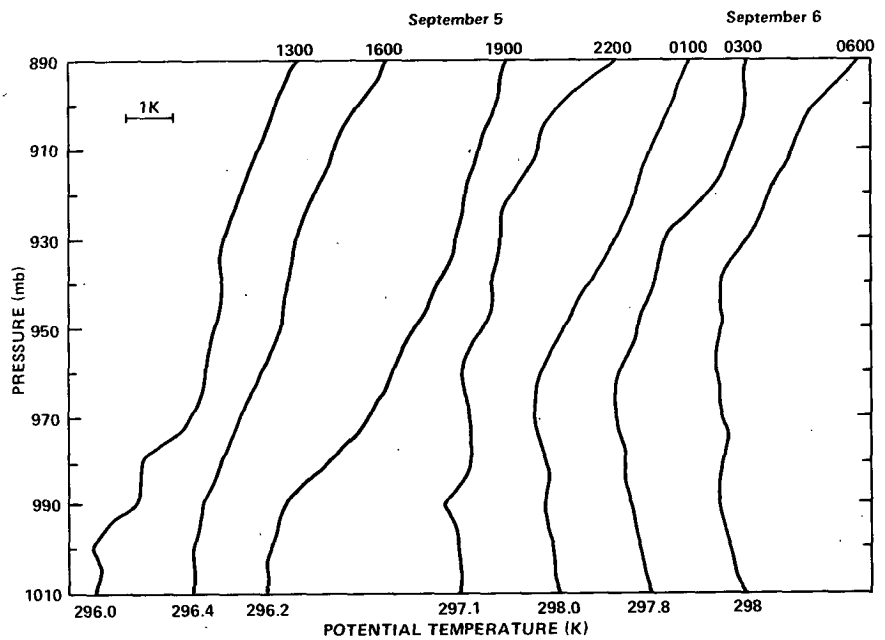

FIG. 6. A sample of potential temperature profiles illustrating a transition from a disturbed to an undisturbed state.

array (Hudlow and Patterson, 1978). If we plot the $\hat{y}_j^{(2)}$ coefficients on the same graph as the radar estimated precipitation (Fig. 7) we see that the occurrences of precipitation over the C-array correspond very closely to the increases in the value of $\hat{y}_j^{(2)}$ coefficient. This indicates that the type-2 profiles are associated with precipitation and presumably convective activity.

Increases in low-level convergence have been linked to increases in convective activity by several investigators (e.g., Ropelewski and Reeves, 1977; Ulanski and Garstang 1978). We have computed the surface divergence over the C-array using the ship mast winds from the *Fay, Meteor,* and *Planet*. We see that increases in the $\hat{y}_j^{(2)}$ coefficients in (Fig. 7) correspond closely to the increases in the magnitude of the divergence giving an independent verification that this class of profile is associated with increased convective activity. The variations in the class 2 coefficients appear to have as good a relation to the variations in the divergence as the divergence has to the radar estimated precipitation.

Finally, we note that the objective classification shows 43 profiles or 35% of the profiles in the disturbed category. If the disturbances occur at random with respect to the profile times, then we can infer that the lower atmosphere is disturbed 35% of the time in the GATE C-array. This agrees well with an estimate by Gaynor and Mandics (1978), who show, using acoustic sounder data, that the boundary layer is disturbed 31% of the time during GATE in the vicinity of the *Oceanographer*.

6. Summary and conclusions

A new method for classification and analysis of meteorological data has been described and illustrated. Asymptotic singular decomposition (ASD), the first step of the method, has been described mathematically and then applied to structure sonde data. The second step, classification, effected by the rearrangement of the leading terms of the singular expansion to form exclusive subgroups and typical profiles, has been developed heuristically, then specified formally. Practical aspects of the necessary calculations have been discussed, and the classification procedure has been illustrated with a structure sonde data set.

As with any mathematical tool, several practical considerations must be dealt with to use the

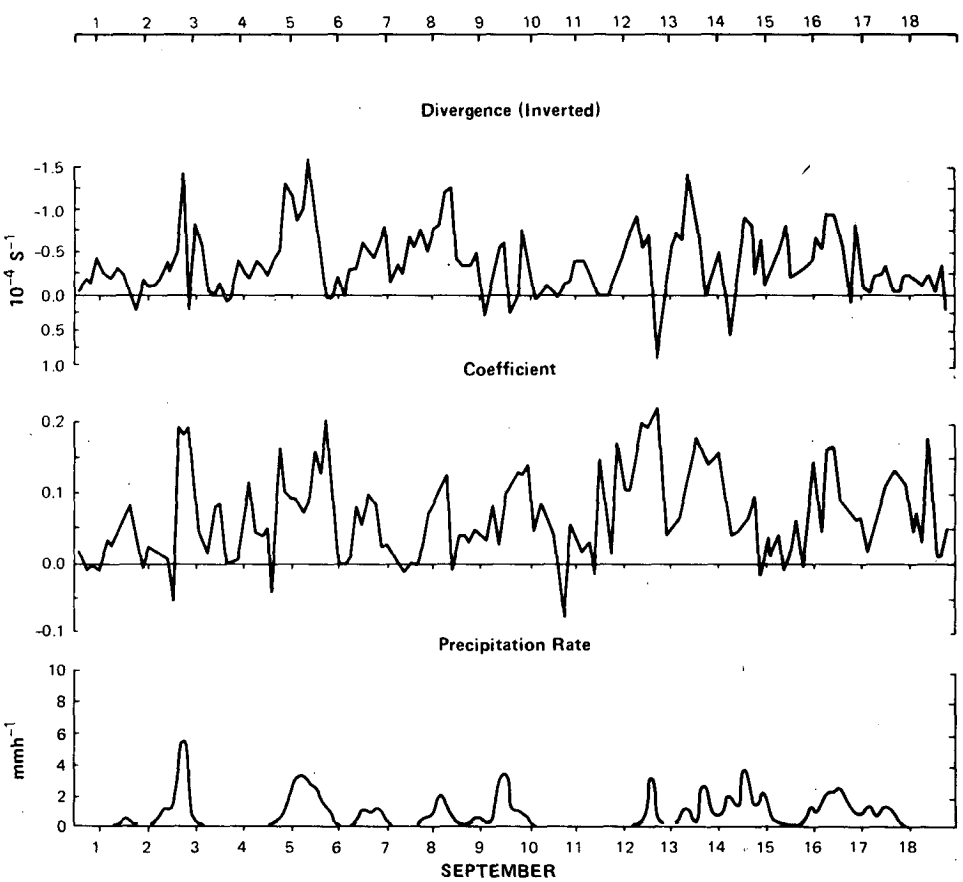

FIG. 7. Time series of C-array divergence (note the inverted scale), $\hat{y}^2$ coefficients, and radar-estimated precipitation rate.

classification scheme effectively. Among these is the question of how to deal with missing data since the ASD and classification scheme will give meaningful results only with full data sets. It may be practical to simply remove from the data set those members (e.g., profiles as was done here) that have missing elements, or interpolation schemes may have to be used. In any case, the method chosen to handle missing data must be taken into account when interpreting the results of the analysis.

The classification scheme works best if the magnitude of the numbers in the data is not too large compared with the variations. In the example discussed here, the actual analysis was performed after 273 K was subtracted from each of the potential temperatures in the set.

The plots of the coefficients (as, for example, in Fig. 3) indicate the class into which each profile belongs. If the data are ordered in time, as they have been here, the plot of the coefficients indicates a probable sequence of events as the profiles go from one class to another. For example, we see that the coefficients indicate a transition from class 2 to class 1 profiles. This suggests that the profiles taken during the transition may be of special interest for analysis. This two-group analysis can even be interpreted to effectively give three classes of profiles: class 2, transition and class 1. In addition, by pointing out periods of like profiles, these plots may indicate periods that would be of special interest for extensive case study analysis or the coefficients may be used as a basis to form composites.

The method can also be used to edit data. In a two-group classification, the coefficients of one group may dominate for most members of the data set. They may indicate the "normal" data, while coefficients of the second group may occur in only a very few cases to indicate the "nonstandard" data that will be removed from the set. The process will then be repeated.

This discussion has been limited to two-group classifications. By way of an experiment, a test of the classification scheme was run with up to six classes of potential temperature profiles. The results showed profiles that were, in general, variations about the two types derived in the original analysis. Although the method itself is capable of generating as many classes as there are members (e.g., profiles) in the original data set, the number of classes must be consistent with the physical profile being studied. The method discussed here is a powerful tool that can readily be used in the analysis of large and complex data sets. We have shown only one example of its use but its generality and computational efficiency make it attractive for a variety of other uses.

Acknowledgments. We wish to thank Dierk Schriever of the GATE Boundary Layer Subprogram Data Center, Hamburg, FRG, for providing us with the structure sonde data. Thomas Carpenter, of CEAS, provided invaluable help in the initial stages of the analysis. Eugene Rasmusson's suggestions and encouragement are appreciated. Jerry Sullivan and Frank Richards also provided many helpful suggestions and comments.

REFERENCES

Betts, A. K., 1976: The thermodynamic transformation of the tropical subcloud layer by precipitation and downdrafts. *J. Atmos. Sci.*, **33**, 1008–1020.

Brier, G. W., and G. T. Meltesen, 1976: Eigenvector analysis for prediction of time series. *J. Appl. Meteor.*, **15**, 1307–1312.

Gaynor, J. E., and P. A. Mandics, 1978: Analysis of the tropical marine boundary layer during GATE using acoustic sounder data. *Mon. Wea. Rev.*, **106**, 223–232.

——, and C. F. Ropelewski, 1979: Analysis of convectively modified GATE boundary layer using *in situ* and acoustic sounder data. (In preparation).

Gabriel, K. R., 1972: Analysis of meteorological data by means of canonical decomposition and biplots. *J. Appl. Meteor.*, **11**, 1071–1077.

Good, I. J., 1969: Some applications of the singular decomposition of a matrix. *Technometrics*, **11**, 823–831.

Golub, G. H., and Reinsch, C., 1970: Singular value decomposition and least squares solutions. *Num. Math.*, **14**, 403–420.

Houze, R. A., 1971: Structure and dynamics of a tropical squall-line system. *Mon. Wea. Rev.*, **105**, 1540–1567.

Hudlow, M. D., and V. L. Patterson, 1978: Gate radar rainfall atlas. NOAA Special Report, CEDDA, NOAA, Washington, D.C. (In preparation).

Jalickee, J. B., and W. J. Klepczynski, 1977: A method for compacting navigation tables. *Navigation*, **24**, 125–131.

——, and D. R. Hamilton, 1977: Objective analysis and classification of oceanographic data. *Tellus*, **29**, 545–560.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather predictions. Sci. Rep. No. 1, Contract AF 19(604)-1566, Dept. Meteor., MIT, 49 pp. [NTIS AD-110268].

Noble, B., and J. W. Daniel, 1969: *Applied Linear Algebra*, 2nd ed. Prentice-Hall, 477 pp.

Ropelewski, C. F., and R. W. Reeves, 1977: GATE Convection Subprogram Data Center: Comparison of ship-surface, rawinsonde, and tethered sonde wind measurements. NOAA Tech. Rep. EDS 21, 22 pp. [NTIS PG-268-848].

——, and J. B. Jalickee, 1979: An objective analysis of the boundary layer thermodynamic structure during GATE, II: Analysis (in preparation).

Seguin, W. R., and M. Garstang, 1976: Some evidence of the effects on convection on the structure of the tropical subcloud layer. *J. Atmos. Sci.*, **33**, 660–666.

Ulanski, S., and M. Garstang, 1978: The role of surface divergence and vorticity in the life cycle of convective rainfall. *J. Atmos. Sci.*, **35**, 1047–1069.

WMO, 1975: Report on the field phase. GATE Rep., No. 16, World Meteorological Organization, Geneva,

Zipser, E. J., 1977: Mesoscale and convective-scale downdrafts as distinct components of squall-line structure. *Mon. Wea. Rev.*, **105**, 1568–1589.