

Empirical Determination of the Basic Modes of Cumulus Heating and Drying Profiles

CHICHUNG LIN AND AKIO ARAKAWA

Department of Atmospheric Sciences, University of California, Los Angeles, Los Angeles, California

(Manuscript received 28 January 1999, in final form 7 October 1999)

ABSTRACT

The constraint on the coupled vertical profiles of cumulus heating and drying, which can be used as a partial closure in cumulus parameterization, is examined using observational data from convectively active regions in the summertime. The data used in this study include those derived from Global Atmospheric Research Programme (GARP) Atlantic Tropical Experiment Phase III, Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment over the intensive flux array region, and four subsets of the European Centre for Medium-Range Weather Forecasts Re-Analysis data that cover areas ranging from tropical to midlatitude continents. The profiles of Q_1 and Q_2 calculated from those data are analyzed using a statistical method. The proposed method is a revised version of the rotated principal component analysis based on the Promax rotation (RPCA_{Promax}), which is believed suitable for identifying basic structures embedded within a given dataset. It is designed in such a way that the distortion of identified structures due to the use of a linear model is minimized. The revised RPCA_{Promax}, together with some selected statistical tools, are evaluated using synthetic datasets before they are applied to observations.

The analysis of the observational data shows that, for all the convectively active regions examined, most of the variance of observed Q_1 and Q_2 can be explained by retaining only two modes. Moreover, while these two modes have different amplitudes in time and space, the shapes of the Q_1 and Q_2 profiles associated with each mode are similar from one region to another. In this sense, they are analogous to the cloud types in the spectral cumulus ensemble model of the Arakawa–Schubert cumulus parameterization, in which the spectral distribution of cloud-base mass flux varies with large-scale conditions while the vertical profile of normalized mass flux is fixed for each cloud type. It is suggested that, as far as deep convection is concerned, the cloud model in cumulus parameterization probably can be constructed based on the empirically determined Q_1 and Q_2 profiles.

1. Introduction

The objective of cumulus parameterization is to formulate the collective effect of cumulus clouds without predicting individual clouds. It is a closure problem in which we seek a limited number of equations that govern the statistics of a system with huge dimensions (Arakawa 1993). The essence of cumulus parameterization problem is, therefore, in the choice of appropriate closure assumptions with which the cumulus effects on the grid-scale dynamics and thermodynamics can be formulated. A part of the necessary closures can be provided by the cloud model of cumulus parameterization (CMCP) itself, such as the spectral cumulus ensemble model in the Arakawa–Schubert (1974) parameterization. With clouds divided into subensembles (cloud types), the spectral model determines the vertical distributions of *normalized* mass flux associated with each cloud type for given large-scale thermodynamic vertical

profiles. The constraint on the coupling of cumulus heating and drying through this mass flux is defined by Arakawa and Chen (1987) as the type II closure. This type of closure also exists, though implicitly, in other cumulus parameterization schemes such as the moist-convective adjustment scheme (Manabe et al. 1965) and the Kuo scheme (Kuo 1965, 1974).

The type II closure has been explored in several studies using the observational data. Chen (1989; see also Arakawa and Chen 1987) treated the vertical profiles of $Q_1 - Q_R$ and Q_2 derived from the Global Atmospheric Research Programme (GARP) Atlantic Tropical Experiment (GATE) data as pairs of vectors and performed canonical correlation analysis on these pairs to extract the best correlated $Q_1 - Q_R$ and Q_2 profiles.¹ Here Q_1 and Q_2 are the apparent heat source and apparent moisture sink (Yanai et al. 1973), respectively, and Q_R is the radiative heating rate. Alexander et al. (1993) analyzed

Corresponding author address: Dr. Akio Arakawa, Department of Atmospheric Sciences, University of California, Los Angeles, Los Angeles, CA 90095.
E-mail: aar@atmos.ucla.edu

¹ The same analysis was made on an Asian dataset described by He et al. (1987). The radiation data, however, are not available in this dataset so that Q_1 and Q_2 are analyzed instead of $Q_1 - Q_R$ and Q_2 .

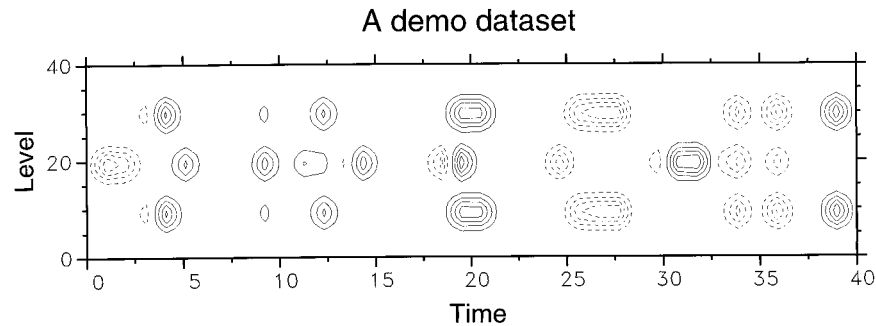


FIG. 1. The time–height cross section of synthetic Q_1 anomalies. Solid and dashed contours show positive and negative values, respectively, with an interval of 0.5 unit.

the Q_1 and Q_2 that were derived from both disturbed and undisturbed periods of the Australian Monsoon Experiment (AMEX).² In their analysis, the derived Q_1 and Q_2 profiles were combined to form a vector (Q_1, Q_2) . A rotated principal component analysis (RPCA) was then applied to the combined vectors to identify the coupled modes of Q_1 and Q_2 profiles. Similar analysis was also performed by Liu (1995; see also Arakawa 1993) using the same data that were analyzed by Chen (1989).

A comparison of the results obtained from these studies reveal some significant differences among the extracted modes of Q_1 and Q_2 profiles, which could be due to regional differences in the coupling between Q_1 and Q_2 or the different methodologies used. This gives rise to the following questions: Are there basic coupled modes of Q_1 and Q_2 in nature? What is the tool to be used for identifying such basic modes if they exist at all? These questions are important particularly from the viewpoint of constructing an empirical cumulus parameterization in which the cumulus heating and drying profiles predicted by a CMCP are replaced by empirical results. If we compare the basic modes of Q_1 and Q_2 to the heating and drying profiles associated with various cloud types in the spectral cumulus ensemble model (Arakawa and Schubert 1974), we would expect that they are realizable by their own at one time or another in different parts of the world and, with the possibility of coexistence, they should be able to explain the observed Q_1 and Q_2 profiles in any regions. An analysis method is thus considered a useful tool if it can be used to identify the coupled modes of Q_1 and Q_2 with the aforementioned two characteristics.

To address the questions raised above, we first review in section 2 the potentially useful eigenmodels with special attention given to RPCA, which we feel suitable for identifying basic modes from a given dataset as illustrated by an example presented in section 3. In section 4, a revised RPCA is proposed, and its performance, together

with some other methods, are assessed using synthesized datasets. The revised RPCA is then applied to GATE, Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE), and European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis data to extract basic coupled modes of Q_1 and Q_2 in section 5, and discussion of the analysis results is given in section 6.

2. Eigenmodels

A wide range of eigenvectorial techniques (eigenmodels) have been used in analyzing meteorological data with increasing frequency (Richman 1986). Among the well-documented techniques are principal component analysis (PCA) (e.g., Jolliffe 1986; Preisendorfer 1988), canonical correlation analysis (CCA) (e.g., Cooley and Lohnes 1971), and common factor analysis (CFA) (e.g., Mulaik 1972; Harman 1976). PCA is also frequently referred to as empirical orthogonal functions (EOFs). With EOF/PCA, the variance of a dataset can be described using minimal number of extracted components that are spatially and temporally orthogonal to each other. EOF/PCA was first applied to the meteorological research by Lorenz (1956) and has since become the most extensively used eigenmodel for analyzing spatial or temporal data variability.

While EOF/PCA is primarily used to describe the variability of a field, CCA is employed to explore the covariability between two variables. In CCA, two variables are independently analyzed into two sets of components through PCA. Linear transformations are then performed on these two sets of components in such a way that each pair of the transformed components have locally maximum correlation, thus representing the coupled modes between these two variables.

Compared with EOF/PCA or CCA, CFA is much less frequently used in meteorological applications. CFA excludes the variance unique to each individual variable, dealing only with the common variance, which is to be explained by hypothesized factors. A typical approach followed by many factor analysts is to extract preliminary factors through PCA and then obtain the final

² The dataset derived from the AMEX has been described by McBride et al. (1989).

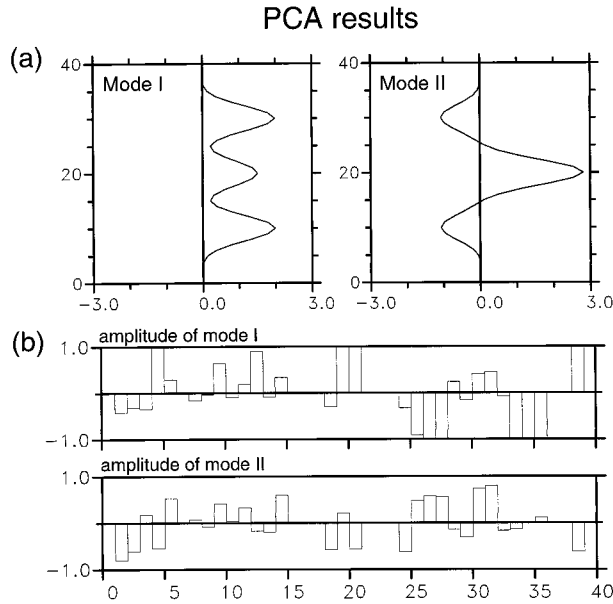


FIG. 2. (a) The vertical profiles of Q_1 anomalies associated with the two modes extracted by PCA. (b) The amplitudes for each of the two modes.

solution through rotations of the preliminary factors (Mulaik 1972; Harman 1976). Such an approach is referred to as RPCA in meteorological applications (Richman 1986).

A question that immediately arises is, what are the advantages of RPCA over the standard EOF/PCA or CCA? The EOF/PCA are well-established eigenmodels ideal for pure data reduction into uncorrelated components. However, if the objective of the analysis is to identify the basic structures that are physically realizable and, therefore, interpretable, EOF/PCA may not be useful because such basic structures could be interrelated. The extracted components from EOF/PCA are inherently orthogonal to each other. CCA also shares the same problem with EOF/PCA in this regard. RPCA solutions, on the other hand, often agree better with the physically

Results after the rotation based on the simple structure criterion

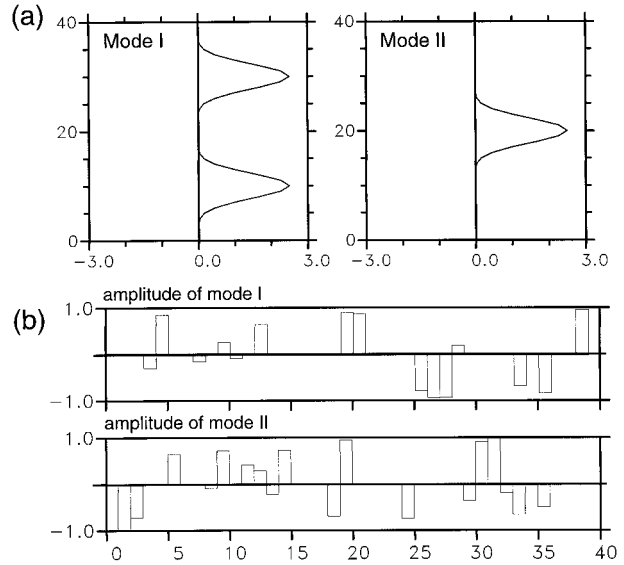


FIG. 4. (a) The vertical profiles of Q_1 anomalies associated with the two modes obtained from simple structure rotation. (b) The amplitudes for each of the two rotated modes.

realizable structures of the input datasets than EOF/PCA solutions. Such examples can be found in Vargas and Compagnucci (1983), Richman and Lamb (1985), and Richman (1986), among many others.

Another question can then be raised: What is the rationale behind the rotations that make RPCA a useful tool for identifying the basic structures embedded within a dataset? To answer this question, it is convenient to introduce the concept of *simple structure*, the objective all the existing rotation methods are designed to achieve. The idea of simple structure was proposed by Thurstone (1947), who argued that, to facilitate physical interpretation of a dataset, the minimal number of factors (components) should be used to explain the data variability whenever possible. Rotations guided by the simple

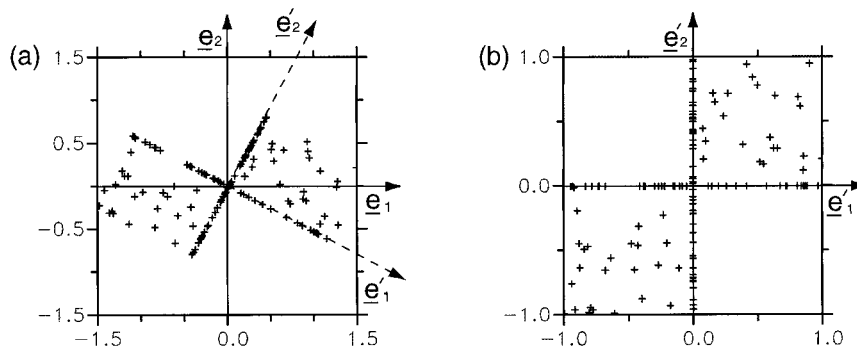


FIG. 3. (a) The scatter diagram of the amplitudes for the two modes extracted by PCA. The abscissa and ordinate represent the amplitudes of the first and second modes, respectively. (b) Same as (a) except for the two modes obtained from simple structure rotation.

structure principle generally help to identify the structures that have the maximal realizability and are thus most likely to be recognized as *basic modes* embedded within a dataset. This argument is illustrated by an example presented in the next section. As for a complete description of the theoretical and technical aspects of the simple structure principle, readers are referred to Thurstone (1947), Mulaik (1972), or Harman (1976).

3. An example of simple structure

In the following, we describe a synthetic dataset that is constructed to show how the simple structure principle helps to identify the basic modes embedded within a dataset. Suppose that there exists a dataset consisting of n observations (in time) of a field, say, Q_1 anomalies, at p vertical levels.³ This dataset can be described by an $n \times p$ array \mathbf{Z} ,

$$\mathbf{Z}^T = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n), \quad (3.1)$$

where the superscript T denotes the transpose operation; and each \mathbf{z}_i , representing the observed vertical profile at the time index i ($1 \leq i \leq n$), is a linear combination of two p -dimensional vectors, \mathbf{a}_1 and \mathbf{a}_2 ,

$$\mathbf{z}_i = \alpha_1^i \mathbf{a}_1 + \alpha_2^i \mathbf{a}_2. \quad (3.2)$$

The k th element of \mathbf{a}_1 denoted by a_1^k , is given by

$$a_1^k = \left[2.5 \times \cos\left(\frac{k-10}{20}\pi\right) \right]^8; \quad (3.3)$$

and the k th element of \mathbf{a}_2 , denoted by a_2^k , is given by

$$a_2^k = \left[2.5 \times \cos\left(\frac{k-20}{20}\pi\right) \right]^8. \quad (3.4)$$

The amplitudes α_j^i ($1 \leq i \leq n$; $j = 1, 2$) in (3.2) are prescribed by

$$\alpha_j^i = \begin{cases} 0 & \text{if } \varepsilon_{i+j} \leq 0 \\ (-1)^{m_i} \varepsilon_{i+j} & \text{if } \varepsilon_{i+j} > 0, \end{cases} \quad (3.5)$$

where ε_{i+j} is a random real number within $(-1, 1)$; and m_i , independent of ε_{i+j} , is an integer randomly assuming the value 0 or 1. With (3.5), α_1^i and α_2^i can be either positive or negative but always have the same sign at each observation i ($1 \leq i \leq n$). With (3.1)–(3.5), the dataset \mathbf{Z} is constructed to have two embedded structures (represented by the p -dimensional vectors \mathbf{a}_1 and \mathbf{a}_2) while, at the same time, its overall pattern is sufficiently simple.

Figure 1 shows a portion of the \mathbf{Z} 's time–height cross section for the case with $n = 250$ and $p = 40$. This dataset is first analyzed by PCA. After deriving p ei-

genectors from the covariance matrix $\mathbf{Z}^T \mathbf{Z}$, \mathbf{Z} can be described by

$$\mathbf{Z} = \mathbf{F} \cdot \mathbf{A}^T, \quad (3.6)$$

where \mathbf{F} is an $n \times p$ matrix with principal components (PCs) as its row vectors, and \mathbf{A} is a $p \times p$ matrix with eigenvectors (EOFs) as its column vectors.⁴ The EOFs are scaled by the square root of their corresponding eigenvalues so that each EOF has a unit length. Note that in this case the variance of the \mathbf{Z} can be completely explained by the first two EOFs since the degree of freedom for the \mathbf{Z} is only 2. These two EOFs, whose associated structures are shown in Fig. 2a, represent two modes of anomalous Q_1 -profiles determined by PCA. The corresponding PCs, a portion of which are shown in Fig. 2b, represent the amplitudes for each modes. The scatter diagram of the amplitudes for each mode is shown in Fig. 3a, in which the abscissa and ordinate indicate the amplitudes of the first and second modes, respectively.

It is noted that, although the dataset \mathbf{Z} can be described by (3.6) through linear combinations of two EOFs as determined through PCA, these two EOFs do not correspond to the modes given by (3.3) and (3.4). We then proceed to apply Thurstone's (1947) simple structure criterion to rotate the derived PCs/EOFs. (By rotation we mean a nonsingular linear transformation on the PCs/EOFs.) We first rewrite (3.6) as

$$\mathbf{Z} = \mathbf{F}(\mathbf{T}\mathbf{T}^{-1})\mathbf{A}^T = (\mathbf{F}\mathbf{T})(\mathbf{T}^{-1}\mathbf{A}^T) \equiv \mathbf{F}^* \mathbf{A}^{*T}, \quad (3.7)$$

where \mathbf{T} is a nonsingular $p \times p$ matrix, and \mathbf{T}^{-1} is the inverse matrix of \mathbf{T} . According to the principle of simple structure, a rotation is considered the best when each \mathbf{z}_i in (3.1), the observed vertical profile at the time index i , can be described by minimum number of the derived EOFs (basic modes). In other words, the rotation seeks to decompose \mathbf{Z} in such a way that the extracted structures have maximum localization *in time*. Algebraically, it is equivalent to maximizing the number of the zero (or near zero) entries of the matrix \mathbf{F}^* ; geometrically, with each \mathbf{z}_i represented by a data point in the space whose coordinate axes are composed of the EOFs, it is equivalent to rotating the coordinate axes in such a way that maximum number of data points fall onto (or close to) the new axes. For the case as represented by Fig. 3a, rotation of the original coordinate axes \mathbf{e}_1 and \mathbf{e}_2 to the new directions along the \mathbf{e}'_1 and \mathbf{e}'_2 seems to yield the most satisfactory simple structure. With such a new set of coordinate axes, it appears that the maximum number of \mathbf{z}_i 's, which are denoted by crosses in Fig. 3a, can be described with only one mode (coordinate). The linear transformation \mathbf{T} that rotates the axes \mathbf{e}_1 and \mathbf{e}_2 to the \mathbf{e}'_1 and \mathbf{e}'_2 is

³ In fact the temporal order of the data will not affect the PCA or RPCA results at all. Here we introduce a dataset consisting of a time series just for the convenience of discussion.

⁴ The notations used in (3.6) follow those used by Richman (1986), who refers to the elements of the matrix \mathbf{A} as principal component loadings and those of the matrix \mathbf{F} as principal component scores.

Total number of data points = 250; power = 4

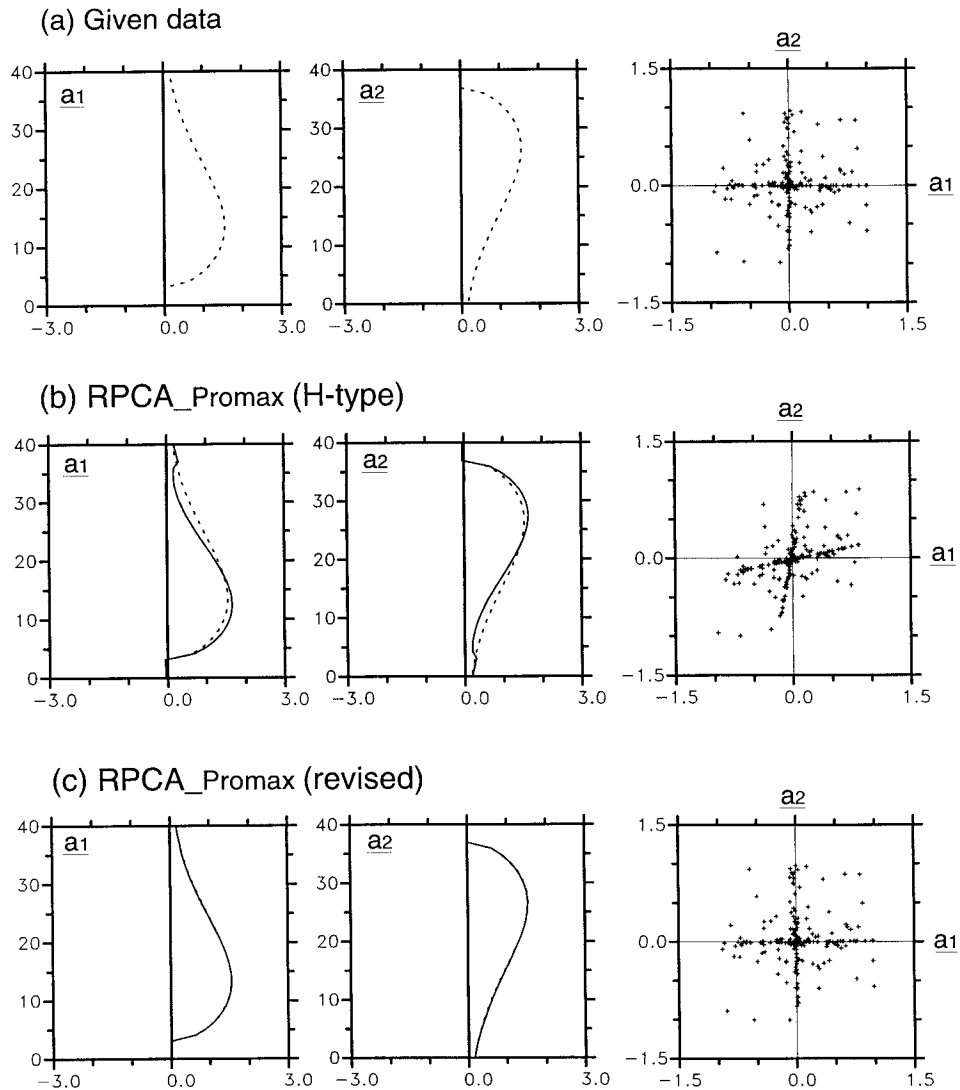


FIG. 5. (a) The prescribed base profiles for data_A (left two panels) and the scatter diagram of the prescribed amplitudes for each base profile (right panel). In the scatter diagram, the abscissa and ordinate are the amplitudes of the first and second base profiles, respectively. (b) The results obtained from performing the *H*-mode $RPCA_{Promax}$ on data_A. Left two panels show the profiles associated with the extracted modes (solid lines) and the prescribed base profiles (dashed lines). The right panel shows the scatter diagram of the data points whose coordinates represent the amplitudes of each modes. (c) Same as (b) except that the results are obtained by performing the revised $RPCA_{Promax}$ on data_A.

$$\frac{1}{\sin(\theta - \Phi)} \begin{pmatrix} \sin\Phi & -\cos\Phi \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad (3.8)$$

where θ is the angle between \mathbf{e}_1 and \mathbf{e}'_1 ($\theta = -29.0^\circ$), and Φ is the angle between \mathbf{e}_1 and \mathbf{e}'_2 ($\Phi = 61.5^\circ$).

After the rotation, the scatter diagram for the \mathbf{z}_i 's is shown in Fig. 3b. The transformed anomalous \mathbf{Q}_1 profiles (modes) are shown in Fig. 4a, and portions of the transformed principal components are shown in Fig. 4b.

The most important consequence of the rotation is probably that both of the extracted modes are similar to the most perceivable patterns, which stand out through their repeated occurrences in Fig. 1. The first mode (left panel of Fig. 4a) has the bimodal structure as observed at the time units 5, 20, 26, 27, 28, and 39. The second mode (right panel of Fig. 4a) has the unimodal structure as observed at the time units 2, 3, 15, 25, 31, and 32. Rotations guided by the simple structure principle thus

Total number of data points = 1000; power = 4

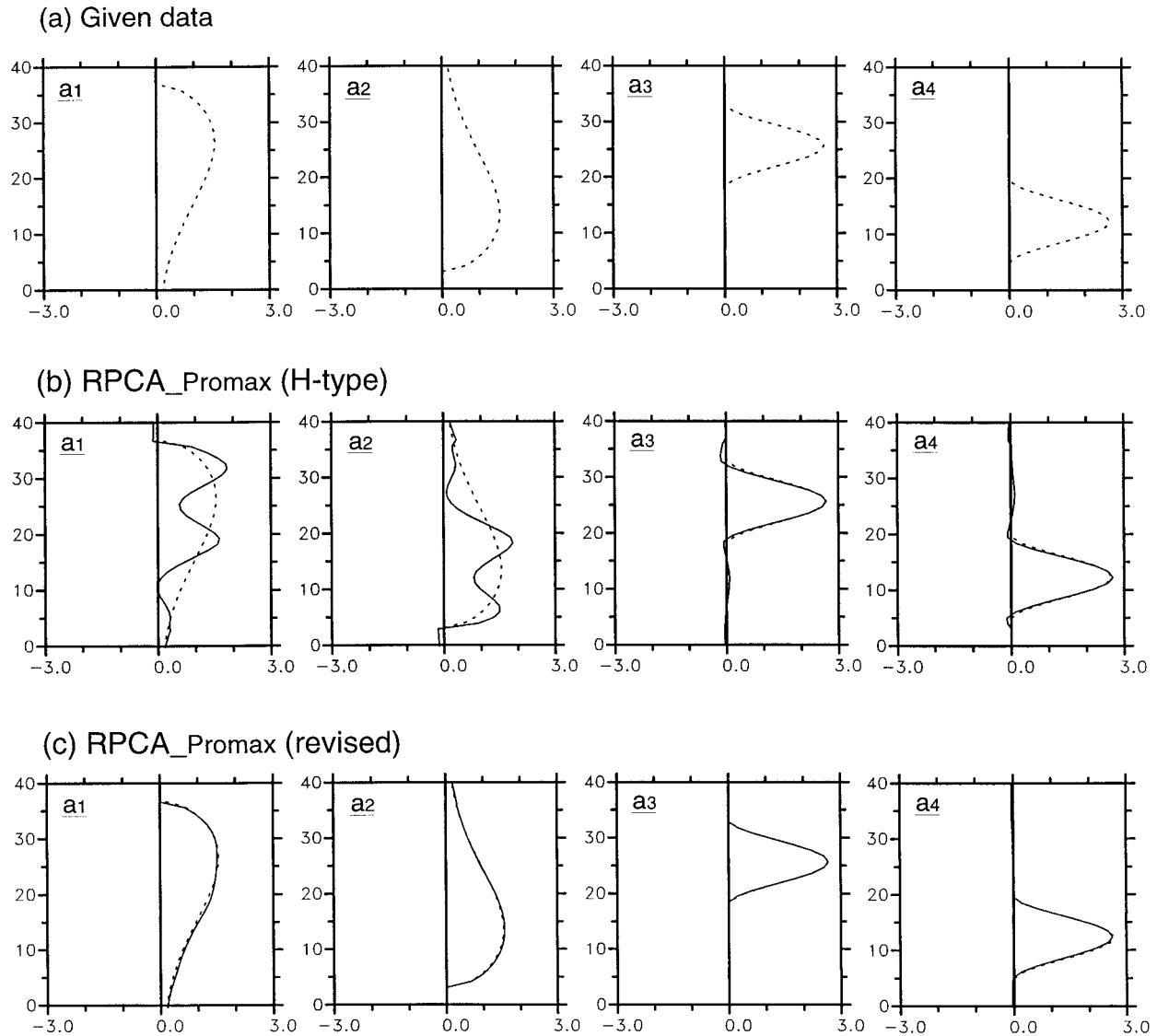


FIG. 6. (a) The prescribed base profiles for data_B. (b) The prescribed base profiles (dashed lines) and the profiles associated with the modes obtained by performing the H -mode $\text{RPCA}_{\text{Promax}}$ on data_B (solid lines). (c) Same as (b) except that the results are obtained by performing the revised $\text{RPCA}_{\text{Promax}}$ on data_B.

maximize the probability that an observed profile is described by only one mode. For this reason, such modes are considered as basic modes embedded within the dataset.

4. The analysis procedure used in this study

For the application of RPCA, many different rotation procedures have been developed to approximate at least a portion of the simple structure criterion using algebraic expressions. The performance of various rotation meth-

ods in recovering the input structures has been assessed by Richman (1986). He concluded that, while no one specific rotation method would always yield the most satisfactory results, Promax rotation (Hendrickson and White 1964) was the most accurate (over a wide range of conditions) among the rotation methods readily available from the major statistical packages. [Refer to Richman (1986) for a list of the available rotation procedures and the major statistical packages.]

However, the usefulness of RPCA based on the Promax rotation ($\text{RPCA}_{\text{Promax}}$) might be limited by the fact

that any linear model inevitably eliminate the ensemble mean of the data. Take the \mathbf{Q}_1 and \mathbf{Q}_2 derived from highly convective events for example. These \mathbf{Q}_1 and \mathbf{Q}_2 are typically positive in the troposphere; thus the magnitude of their ensemble means could be comparable to that of their anomalies. When a linear model, say, EOF/PCA, is used to analyze such a dataset, the derived solutions actually explain the variations of \mathbf{Q}_1 and \mathbf{Q}_2 from their means rather than the \mathbf{Q}_1 and \mathbf{Q}_2 themselves. Thus the modes extracted by a linear model might fail to reflect the embedded basic structures within the data even though the observed variance can be explained by the extracted modes.

Recognizing the potential deficiency of using a linear model, a revised procedure for $\text{RPCA}_{\text{Promax}}$ is developed for use in our analysis. In this procedure, the input data is preprocessed by reducing the magnitudes of the data's ensemble mean. This can be done as follows: Suppose that the input dataset consists of n observations of a field at p vertical levels and can be described by an $n \times p$ array \mathbf{Z} as (3.1), where the column vector \mathbf{z}_i ($1 \leq i \leq n$) represents an observed field. Prior to the analysis, each \mathbf{z}_i is multiplied by a sign function, $\text{Isign}(i)$, which takes a value of either 1 or -1 . The values of $\text{Isign}(i)$ are prescribed so as to minimize the magnitude of \mathbf{z}_i 's ensemble mean, which is measured by L_i ,

$$L_i \equiv \left\| \sum_{l=1}^i \text{Isign}(l)\mathbf{z}_l \right\| \quad 1 < i \leq n. \quad (4.1)$$

(Here the notation $\|\cdot\|$ designates the norm of a vector.) It should be noted that the embedded structures of the data are not affected in the aforementioned process, which does nothing more than redistributing the signs of the input data.

With such a manipulation, we can avoid the major drawback of a linear model and improve the capability of the RPCA in recovering the embedded structures of the data. Readers are referred to appendix A for more details of the revised $\text{RPCA}_{\text{Promax}}$.

a. Evaluation of the revised $\text{RPCA}_{\text{Promax}}$

The performance of the revised $\text{RPCA}_{\text{Promax}}$ is evaluated using a number of synthetic datasets and compared to that of the original $\text{RPCA}_{\text{Promax}}$. The synthetic datasets consist of data points falling strongly along coordinate axes so that the structures represented by the coordinate axes are ideal basic modes as discussed in section 3.

Before presenting the results, we would like to remind readers that there could be two different ways of applying the $\text{RPCA}_{\text{Promax}}$ to a given dataset. Except the method described in section 3, which extracts structures that have maximum localization in *time*, one can also decompose the data to extract structures that have maximum localization in *height*. The former method is hereafter referred to as the *T-mode* $\text{RPCA}_{\text{Promax}}$ and the latter as *H-mode* $\text{RPCA}_{\text{Promax}}$ in this paper. The *H-mode*

$\text{RPCA}_{\text{Promax}}$, which is described in more detail in appendix B, was adopted by Alexander et al. (1993), Liu (1995), and Misra (1997) to analyze the observed \mathbf{Q}_1 and \mathbf{Q}_2 profiles.

The results of the evaluation are summarized using four representative examples.

1) EXAMPLE ONE

The synthetic dataset for the first example (data_A) is generated through a procedure similar to that described by (3.1)–(3.5) with $n = 250$ and $p = 40$ except that the base profiles and their amplitudes are prescribed differently. The base profiles, which are shown in the left two panels of Fig. 5a, are idealized \mathbf{Q}_1 profiles for the large-scale environments with shallow and deep convection, respectively. The amplitudes for each base profile, that is, the α_j^i in (3.2) ($1 \leq i \leq 250$; $1 \leq j \leq 2$), are determined by

$$\alpha_j^i = \varepsilon_{i+j}^p, \quad (4.2)$$

where ε_{i+j} is a random real number within $(-1, 1)$, and p is set to 4. [The amplitudes given by (4.2) are more complex than those given by (3.5). The former allows the two base profiles assuming different signs at each observation i , while the latter does not.] The scatter diagram of the prescribed amplitudes for each base profile is shown in the right panel of Fig. 5a, in which the abscissa and ordinate are the amplitudes of the first and second base profiles, respectively. It is clear from the scatter diagram that this synthetic dataset has the characteristic of simple structure inasmuch as there are many data points that fall onto or close to the coordinate axes a_1 and a_2 . We then apply the revised $\text{RPCA}_{\text{Promax}}$, the *H-mode* $\text{RPCA}_{\text{Promax}}$, and *T-mode* $\text{RPCA}_{\text{Promax}}$ to this synthetic dataset. The results obtained from the *H-mode* $\text{RPCA}_{\text{Promax}}$ are shown in Fig. 5b and the results from the revised $\text{RPCA}_{\text{Promax}}$ are shown in Fig. 5c. The results from the *T-mode* $\text{RPCA}_{\text{Promax}}$ are not shown because they are nearly the same as those obtained from the revised $\text{RPCA}_{\text{Promax}}$. The extracted modes (solid lines) and the prescribed base profiles (dashed lines) are shown in the left two panels. The right panels show the scatter diagram of the data points whose coordinates represent the amplitudes of each mode.

We see that the profiles corresponding to the two modes extracted from the *H-mode* $\text{RPCA}_{\text{Promax}}$ are similar to the embedded base profiles. The scatter diagram obtained from the *H-mode* $\text{RPCA}_{\text{Promax}}$, however, deviates significantly from the prescribed. The profiles and scatter diagrams obtained from the revised $\text{RPCA}_{\text{Promax}}$ and the *T-mode* $\text{RPCA}_{\text{Promax}}$, on the other hand, are almost identical to the prescribed. It suggests that, while the revised $\text{RPCA}_{\text{Promax}}$ and the *T-mode* $\text{RPCA}_{\text{Promax}}$ nearly perfectly identify the embedded base profiles in data_A, the performance of the *H-mode* $\text{RPCA}_{\text{Promax}}$ is also acceptable in that the extracted profiles are not far from the prescribed. Nevertheless, the performance of

the H -mode $\text{RPCA}_{\text{Promax}}$ becomes worse when applied to datasets with higher complexity, as shown by the following example.

2) EXAMPLE TWO

A synthetic dataset (data_B) is created following the previous procedure (for generating data_A) but with four prescribed base profiles. The \mathbf{z}_i in (3.2) is now a linear combination of four p -dimensional vectors instead of two. The first two base profiles (\mathbf{a}_1 and \mathbf{a}_2) are the same as those prescribed in data_A, and the last two (\mathbf{a}_3 and \mathbf{a}_4) are idealized \mathbf{Q}_1 profiles for the large-scale environments with stratiform clouds at higher and lower levels, respectively (Fig. 6a). The number of data points is increased to 1000 in this dataset.⁵ The results obtained from the H -mode $\text{RPCA}_{\text{Promax}}$ and those from the revised $\text{RPCA}_{\text{Promax}}$ are shown in Figs. 6b and 6c, respectively. The results from the T -mode $\text{RPCA}_{\text{Promax}}$ are, again, not shown because they are nearly identical to those obtained from the revised $\text{RPCA}_{\text{Promax}}$.

We see that the H -mode $\text{RPCA}_{\text{Promax}}$ succeeds in extracting the prescribed base profiles \mathbf{a}_3 and \mathbf{a}_4 but fail in extracting the profiles \mathbf{a}_1 and \mathbf{a}_2 . The two extracted profiles that apparently correspond to \mathbf{a}_1 and \mathbf{a}_2 seem to have excessive zero or near zero loadings. This is expected because the H -mode $\text{RPCA}_{\text{Promax}}$ attempts to extract structures that have maximum localization in height. The profiles extracted by the revised $\text{RPCA}_{\text{Promax}}$ and the T -mode $\text{RPCA}_{\text{Promax}}$, on the other hand, are nearly the same as the prescribed. It appears that the revised $\text{RPCA}_{\text{Promax}}$ and the T -mode $\text{RPCA}_{\text{Promax}}$ are reliable in identifying the embedded base profiles even when applied to a dataset as complex as data_B, while the H -mode $\text{RPCA}_{\text{Promax}}$ is not.

3) EXAMPLE THREE

The third example is used to demonstrate our earlier statement that the solutions obtained from any RPCA might not be able to reflect the basic structures embedded within the data if the magnitude of the data's ensemble mean is comparable to that of the anomalies themselves. The synthetic data for this example (data_C) is generated in the same way as data_A except that the amplitudes for each base profile are determined by (4.2) with ε_{i+j} randomly distributed within (0, 1). The scatter diagram for data_C is shown in the right panel of Fig. 7a. As in the case for data_A, many data points fall onto or close to the coordinate axes in the scatter diagram.

⁵ Since the data now span a four-dimensional rather than two-dimensional space, we need 250^2 data points to fill the space with similar density as data_A. For the reason of computational economy, however, we use only 1000 data points. As will be shown presently, the embedded structures can be extracted with such a data density if the method is powerful enough.

But now the coordinates of all the data points are non-negative so that the ensemble means of the coordinates are not trivial. This is different from data_A, in which the distribution of the data points is nearly symmetrical to the origin so that the ensemble means of the coordinates are relatively small.

Figures 7b and 7c show the results obtained by applying the T -mode $\text{RPCA}_{\text{Promax}}$ and the revised $\text{RPCA}_{\text{Promax}}$ to data_C, respectively. We see that the extracted profiles by the T -mode $\text{RPCA}_{\text{Promax}}$ do not resemble the prescribed any more. In fact, there are no visible differences between these profiles and those obtained from PCA (not shown). The scatter diagram also deviates from the prescribed, with the axis \mathbf{a}_1 apparently explaining the largest variance of the dataset. It seems that the T -mode $\text{RPCA}_{\text{Promax}}$ does nothing more than PCA for data_C. On the other hand, the profiles and the scatter diagram obtained from the revised $\text{RPCA}_{\text{Promax}}$ are very close to the prescribed.

4) EXAMPLE FOUR

The performance of the T -mode $\text{RPCA}_{\text{Promax}}$ and the revised $\text{RPCA}_{\text{Promax}}$ have also been evaluated using a synthetic dataset (data_D) that is created following the previous procedure (for generating data_C) but with four base profiles (see Fig. 8a). The results obtained from the T -mode $\text{RPCA}_{\text{Promax}}$ are shown in Fig. 8b and the results from the revised $\text{RPCA}_{\text{Promax}}$ in Fig. 8c. We see that the revised $\text{RPCA}_{\text{Promax}}$ still can recover the embedded base profiles reasonably well although the discrepancies between the identified profiles and the prescribed are more pronounced than the case for data_C. As for the T -mode $\text{RPCA}_{\text{Promax}}$, the extracted profiles are drastically different from the prescribed.

In summary, all the methods examined in this section give satisfactory results when the ensemble means of the data are zeros. However, when applying to a dataset whose ensemble mean's magnitude is comparable to that of the anomalies themselves, their performance becomes worse except for the revised $\text{RPCA}_{\text{Promax}}$. Encouraged by these results, we further test the revised $\text{RPCA}_{\text{Promax}}$ using synthetic datasets with increasing complexity. It is found that the revised $\text{RPCA}_{\text{Promax}}$ performs very well even for datasets with moderate or weak simple structure⁶ (not shown).

b. Theoretical considerations on the application

The evaluations in the previous section reveal that the performance of the revised $\text{RPCA}_{\text{Promax}}$ is not sig-

⁶ The complexity of the synthetic datasets can be adjusted by changing the value of p in (4.2). (The smaller p is, the more scattered the data points are.) The classification of complexity referred to here is based on Richman's (1986) schematic examples (see his Fig. 7), which illustrate various degrees of simple structure (strong, moderate, weak, and random).

Total number of data points = 250; power = 4

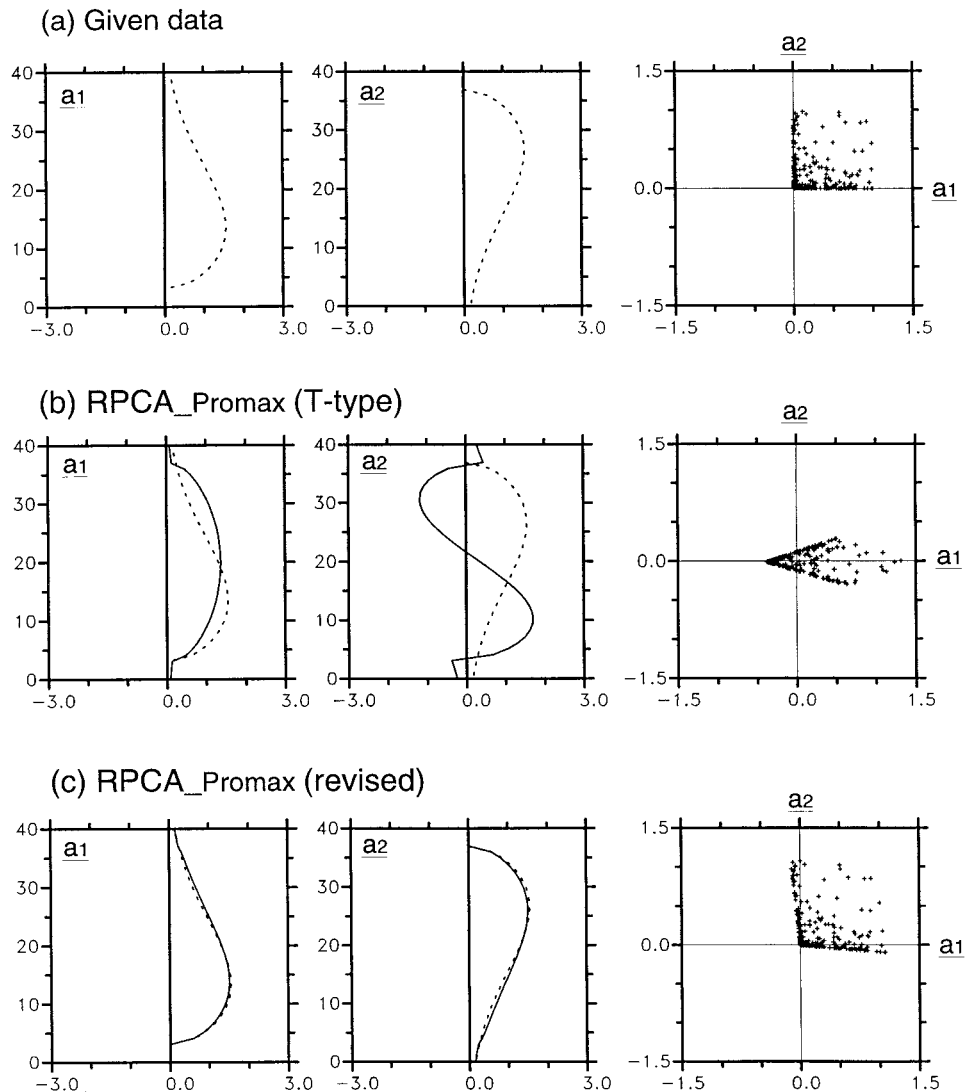


FIG. 7. (a) Same as Fig. 5a except for data_C. (b) Same as Fig. 5b except that the results are obtained by performing the T -mode $\text{RPCA}_{\text{Promax}}$ on data_C. (c) Same as (b) except that the results are obtained by performing the revised $\text{RPCA}_{\text{Promax}}$ on data_C.

nificantly affected, as other RPCA methods, when applying to a dataset whose ensemble mean is comparable to the anomalies themselves. Therefore, in the next section, we will use the revised $\text{RPCA}_{\text{Promax}}$ to extract basic modes of \mathbf{Q}_1 and \mathbf{Q}_2 that are derived from strongly convective regions because, for those regions, the means and standard deviations of \mathbf{Q}_1 and \mathbf{Q}_2 typically have similar magnitudes.

With the revised $\text{RPCA}_{\text{Promax}}$ selected as the tool for our analysis, it is necessary to decide how many PCs/EOFs should be retained in PCA and subsequently rotated according to the Promax method. Numerous ob-

jective rules for determining an optimum number of retained PCs/EOFs have been developed in light of separating the signal from noise in a dataset (e.g., Cattell 1966; Preisendorfer and Barnett 1977). However, it is not our primary concern to retain all the signals for each individual dataset. Rather, we are more interested in whether there exist coupled $\mathbf{Q}_1 - \mathbf{Q}_R$ and \mathbf{Q}_2 profiles that can explain most of the observed variance and at the same time are similar in different regions in the world. Therefore, while retaining more PCs/EOFs is favorable to explain more variance, the similarity among the extracted modes from different regions is considered

Total number of data points = 1000; power = 4

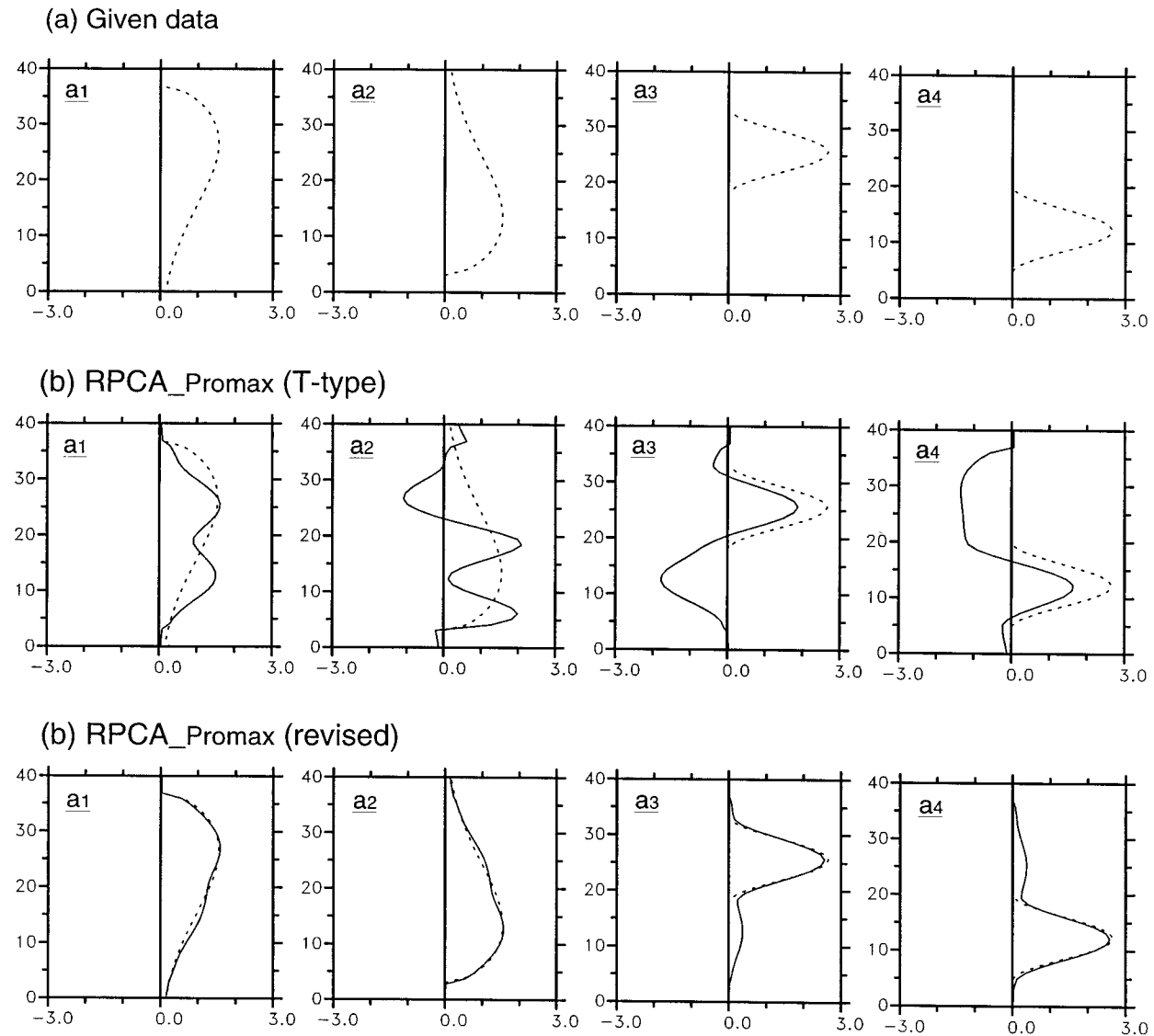


FIG. 8. (a) The prescribed base profiles for data_D. (b) The prescribed base profiles (dashed lines) and the profiles associated with the modes obtained by performing the T -mode $\text{RPCA}_{\text{Promax}}$ on data_D (solid lines). (c) Same as (b) except that the results are obtained by performing the revised $\text{RPCA}_{\text{Promax}}$ on data_D.

most important for determining the number of retained PCs/EOFs.

5. Analysis of the observational data

a. The data

Three datasets (GATE Phase III, TOGA COARE IOP, and ECMWF Re-Analysis) are analyzed using the revised $\text{RPCA}_{\text{Promax}}$ in this study.

1) GATE PHASE III

The wind, temperature, and relative humidity data are analyzed by Ooyama and Esbensen using a statistical interpolation scheme (Ooyama 1987; Esbensen et al. 1982). The grid system has a resolution of $0.5^\circ \times 0.5^\circ$ within the domain $4^\circ\text{--}13^\circ\text{N}$, $19^\circ\text{--}28^\circ\text{W}$. Three-hourly data over a 20-day period (30 August–18 September 1974) are originally given at 41 pressure levels from 1012 to 70 mb (Sui and Yanai 1986). The radiative heating rates, Q_R , were calculated by Cox and Griffith

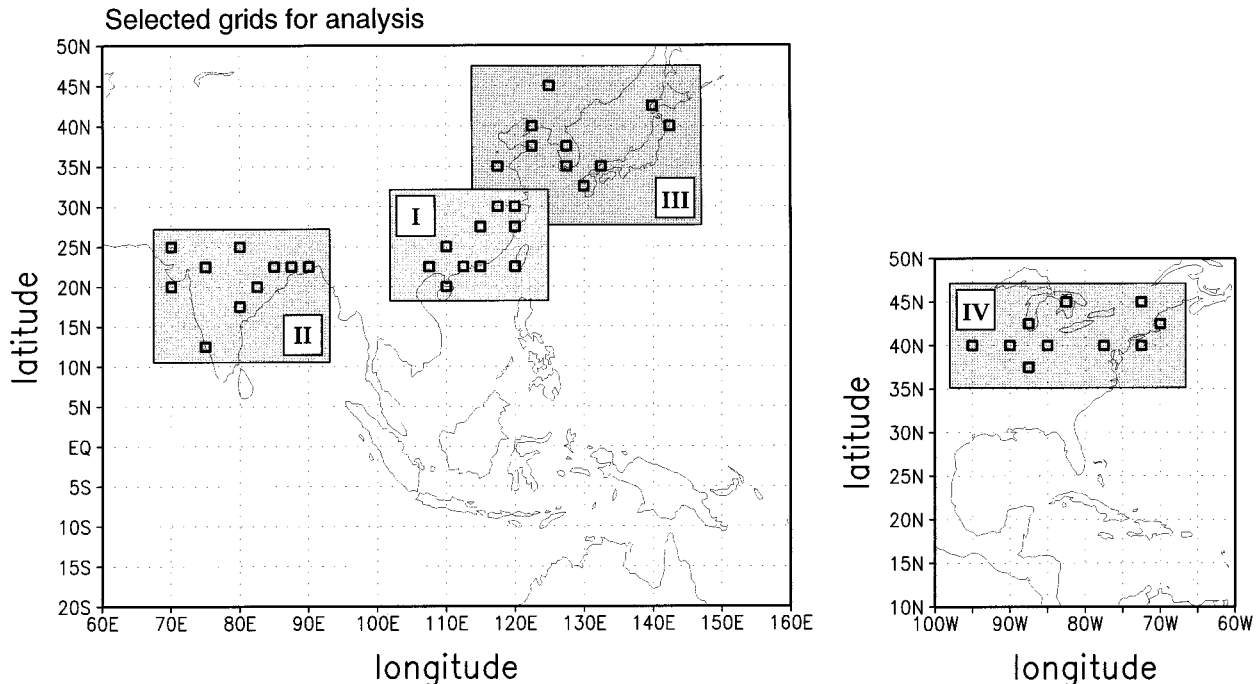


FIG. 9. The grids used in our analysis. They are chosen from the following regions (indicated by rectangular with shading): (I) southeastern China; (II) India; (III) northeastern China; Korea; Japan; and (IV) northeastern United States.

(1979) and $Q_1 - Q_R$ and Q_2 were calculated by Chen (1989). All these data are averaged over the center $3^\circ \times 3^\circ$ box before being used for the analysis.

2) TOGA COARE IOP

The wind, temperature and relative humidity data are analyzed by Tung et al. (1999). A brief summary of this dataset is given as follows. The ECMWF Re-Analysis was used as the first guess and refined with data from 43 sounding stations by successive correction method in the domain of 30°S – 30°N , 90°E – 170°W . Twelve-hourly data over a 4-month period (November 1992–February 1993) were on a $2.5^\circ \times 2.5^\circ$ grid and interpolated into 20 levels with a 50-mb interval from 1000 to 50 mb. The Q_1 and Q_2 averaged over 5°S – 5°N , 150° – 160°E , which covers the intensive flux array (IFA) region, are used in this study.⁷

3) ECMWF RE-ANALYSIS

This dataset covers a 15-yr period from 1979 to 1993, including global 4 times daily zonal wind, meridional wind, temperature, specific humidity, and geopotential

height on a $2.5^\circ \times 2.5^\circ$ grid at 17 levels from 1000 to 10 mb. The details of this data have been described by Gibson et al. (1997). In our analysis, Q_1 and Q_2 were calculated for selected convectively active regions during their summertime (July) using the data from the entire 15 yr. These regions are 1) southeastern China; 2) India; 3) northeastern China; Korea; Japan; and 4) northeastern United States. For each region, 10 grid points are selected for analysis (Fig. 9). These grid points are deliberately chosen from areas with relatively high density of neighboring stations to ensure that much weight has been given to the observations rather than the parameterization schemes in the reanalysis.

The aforementioned three datasets, originally on different pressure levels, are interpolated into a common height coordinate with 500-m interval prior to analysis. The $Q_1 - Q_R$ and Q_2 for the GATE Phase III data are on 25 levels from 1.0 to 13.0 km. Chen (1989) did not provide the data above the 13.0-km level because the calculation of Q_1 is sensitive to the vertical p velocity (ω) near the tropopause. (The static stability is very large near the tropopause so that small errors in ω would lead to erroneous values of Q_1 .) As for the TOGA COARE IOP and ECMWF Re-Analysis data, Q_1 and Q_2 are on 29 levels from 1.0 to 15.0 km. The ω at the tropopause was calculated with the assumption of adiabatic condition ($Q_1 \approx Q_R$) (Nitta 1977), where $Q_R \approx 0$ is further made based on Liou (1992, chapter 6).

While the data for the entire periods of GATE Phase III and TOGA COARE IOP are used for our analysis,

⁷ Here Q_1 , instead of $Q_1 - Q_R$, is used because the radiation data are not available. The best we can hope is that Q_R is much smaller than Q_1 so that the results with RPCA applied to Q_1 might not be significantly different from those with RPCA applied to $Q_1 - Q_R$. Such an argument also applies to the ECMWF Re-Analysis data.

TABLE 1. A list of the data analyzed by the revised RPCA_{Promax} in this study.

Dataset	Selected region	Sample size	Cumulative variance (%) explained with 1/2/3 modes
GATE Phase III	Tropical eastern Atlantic	159	82.9/88.5/93.1
TOGA COARE IOP	IFA region	238	87.4/93.6/95.8
ECMWF Re-Analysis	Southeastern China	630	83.8/90.5/93.9
Same as above	India	1055	77.6/87.1/91.0
Same as above	Northeastern China; Korea; Japan	762	77.6/86.8/91.1
Same as above	Northeastern United States	352	75.9/86.0/91.3

the ECMWF Re-Analysis data are screened to select the portion that represent strong convective events and suffer fewer problems in data inconsistency.⁸ We reject the cases whose vertical mean of Q_1/C_p and Q_2/C_p are less than 2.5 K day^{-1} . Additional rejections of the data were made in view of the correction made in calculating ω . In constructing this dataset, the divergence at each grid point and at each observation time was corrected uniformly in the vertical to obtain a reasonable value of ω at the tropopause following Tung et al.'s (1999) approach. We rejected the data for which the correction of the vertical mean divergence exceeded 2.5% of the maximum divergence of the layers. The size of the screened data for each region is listed in Table 1.

b. Analysis of the GATE Phase III data

Prior to the analysis, each of the $Q_1 - Q_R$ and Q_2 profiles were first normalized by their ensemble means of vertically integrated values and combined into a vector. The matrix \mathbf{Z} , which is formally described by (3.1) with the \mathbf{z}_i substituted by the combined vectors,⁹ was analyzed by the revised RPCA_{Promax}, and the derived EOFs/PCs are then denormalized for $Q_1 - Q_R$ and Q_2 , respectively.¹⁰ As a result, \mathbf{Z} is expressed as a product of EOFs (profiles normalized by their variances) and PCs (time series of amplitudes for each normalized profiles). The unit for the entries of \mathbf{Z} and EOFs is K day^{-1} .

We first retain two EOFs, which together explain 88.5% of the total variance. The vertical profiles of $Q_1 - Q_R$ and Q_2 corresponding to these two rotated EOFs are shown in Fig. 10a. The maxima of $Q_1 - Q_R$ (solid line) and Q_2 (dashed line) of the first mode are both around the 5-km level, with the overall $Q_1 - Q_R$ profile higher than the Q_2 profile. The second mode, with a

greater difference between the $Q_1 - Q_R$ and Q_2 profiles, has a more convective structure than the first mode. The maxima of Q_2 is centered near the 2-km level, and the $Q_1 - Q_R$ is almost evenly distributed between the 2-km and the 8-km levels. Similar results were obtained earlier by Liu (1995) with an *H*-mode RPCA applied to the same dataset.

The time series of the amplitudes for the two modes are shown in Fig. 10b. It appears that the relative importance of each mode, in terms of its contribution to the variation of the $Q_1 - Q_R$ and Q_2 profiles,¹¹ is different from time to time. Compared with the documented major convective events observed during GATE Phase III [see, e.g., Table 1 in Sui and Yanai (1986)], the first mode tends to be associated with nonsquall clusters, while the second mode tends to be associated with squall clusters. This impression is partially supported by the comparison between Fig. 10b and the time-height cross section of horizontal wind fields for the same time period (Fig. 10c). It is observed that the first mode is more dominant in the periods with weak low-level wind shear, while the second mode is more dominant during or immediately after the periods with strong low-level wind shear. The differences between the $Q_1 - Q_R$ and Q_2 profiles associated with squall cluster and those associated with nonsquall clusters were pointed out earlier by Cheng and Yanai (1989).

Rotation is further made on the first three EOFs, which explain 93.1% of the total variance. The vertical profiles of $Q_1 - Q_R$ and Q_2 corresponding to these three rotated EOFs are shown in Fig. 11a, and their amplitudes are shown in Fig. 11b. It appears that the first mode and its amplitudes are nearly identical to those obtained earlier with only two modes retained. The second mode changes slightly, however, with larger $Q_1 - Q_R$ in upper levels and smaller Q_2 in the middle levels. The third mode is new. With its maxima of $Q_1 - Q_R$ and Q_2 both around the 3-km level, the third mode represents a weaker convective structure than those associated with the first two modes. It is noted that the first mode still tends to be associated with nonsquall clusters and the second

⁸ The ECMWF Re-Analysis dataset is treated differently because it is derived using less information from direct observations. However, its huge dataset size allows us to retain, for our analysis, only the portion that we believe is most reliable.

⁹ This approach is different from that adopted by Frank et al. (1996) and Misra (1997), who analyzed the observed Q_1 and Q_2 profiles separately.

¹⁰ Normalization of the input data is a standard practice in the eigenanalysis. In fact we later found that, without normalization, our analysis results are almost indistinguishable from those following the standard practice. This is understandable because the means of vertically integrated $Q_1 - Q_R$ and Q_2 are almost the same.

¹¹ Here the contribution of each mode to the variation of the $Q_1 - Q_R$ and Q_2 is evaluated by its sum of squares of their amplitudes. The relative percentage of contribution from each mode is indicated in the upper-right corner of each panel in Fig. 10a.

The GATE data (two modes)

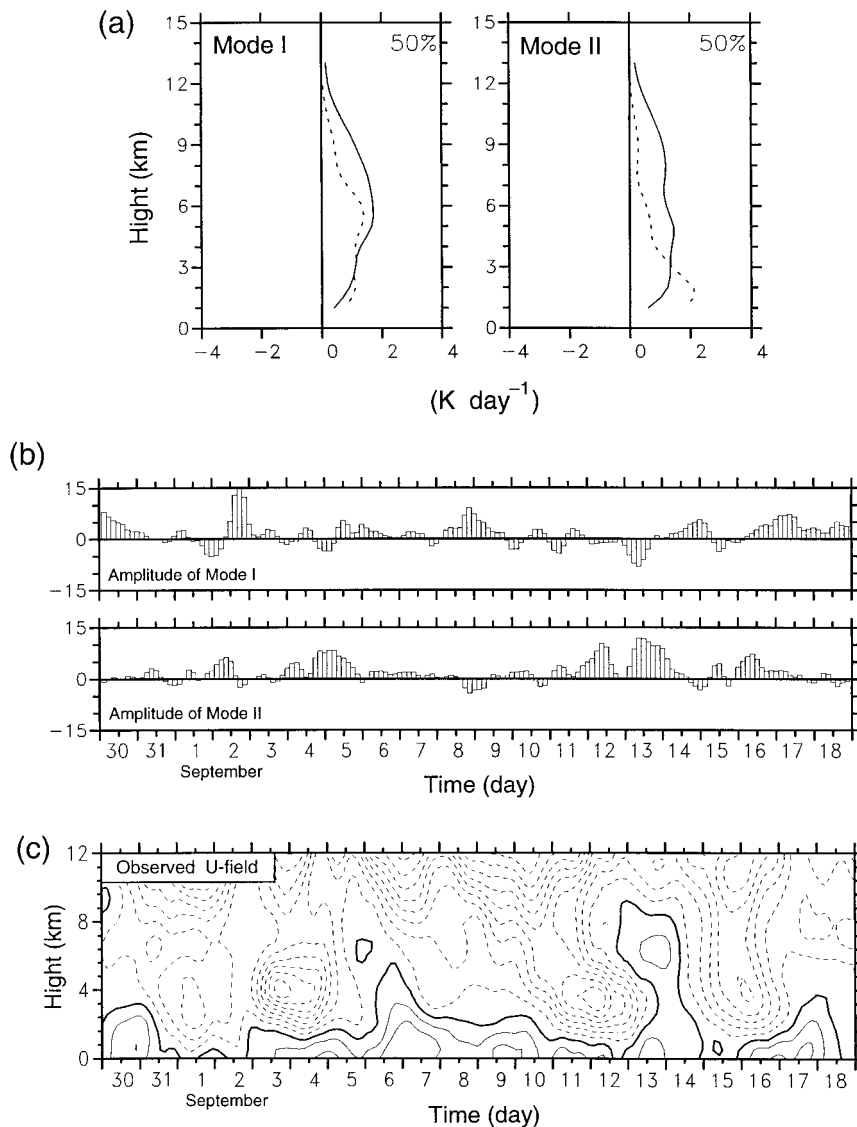


FIG. 10. (a) Vertical profiles of the combined vector $(\mathbf{Q}_1 - \mathbf{Q}_R, \mathbf{Q}_2)$ associated with the two modes extracted by performing the revised $\text{RPCA}_{\text{PROMAX}}$ on the GATE Phase III data. Solid lines are for $\mathbf{Q}_1 - \mathbf{Q}_R$ and dashed lines are for \mathbf{Q}_2 . The relative percentage of contribution from each mode is indicated in the upper-right corner of each panel. Refer to footnote 11 for more explanation. (b) The time series of the amplitudes for each of the two modes. (c) The time-height cross section of the observed zonal wind averaged over the center $3^\circ \times 3^\circ$ box during the GATE Phase III. Solid and dashed contours indicate westerly and easterly, respectively, with interval of 2 m s^{-1} . [Adapted from Sui and Yanai (1986).]

mode tends to be associated with squall clusters, as indicated by the comparison between Fig. 11b and Fig. 10c. However, the relationship between the third mode and the documented convective events is not clear.

It is noted that, from time to time, the basic modes have negative amplitudes that are relatively small compared to positive amplitudes in general (Figs. 10b and

11b). Moreover, the periods with negative amplitudes of one mode are very often accompanied by positive amplitudes of other modes. These periods thus mostly correspond to weak signals of $\mathbf{Q}_1 - \mathbf{Q}_R$ and \mathbf{Q}_2 , whose accuracy is very likely obscured by observational errors or errors in estimating \mathbf{Q}_R . The amplitudes within such periods, we believe, may not have any physical meaning

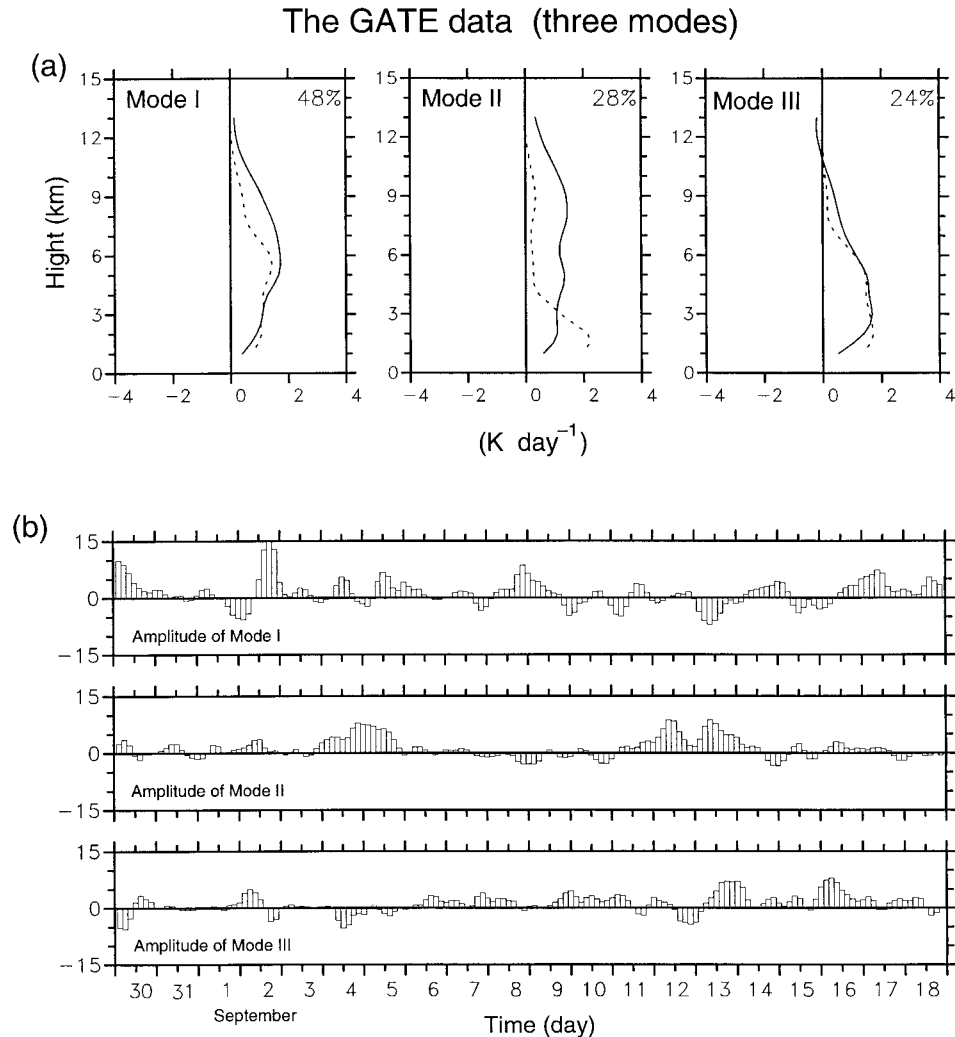


FIG. 11. (a) Same as Fig. 10a except that the results are obtained with three modes retained in the analysis. (b) The time series of the amplitudes for each of the three modes.

but simply result from fitting the residual of $Q_1 - Q_R$ and Q_2 into the retained basic modes.

c. Analysis of the TOGA COARE IOP data

The same analysis procedure described in the last section was applied to the TOGA COARE data over the IFA region except that the combined vectors (Q_1, Q_2), instead of ($Q_1 - Q_R, Q_2$), were used as the input. We first retain two EOFs, which together explain 93.6% of the observed Q_1 and Q_2 variance (Table 1). The Q_1 and Q_2 profiles corresponding to the two rotated EOFs are shown in Fig. 12. While these two modes are very similar to the two modes extracted from GATE (Fig. 10a), it is noticed that the second mode has a significantly smaller contribution to the explained Q_1 and Q_2 variance (27% vs 50% in GATE). This is consistent with our prior impression that the second mode tends to be associated with squall line clusters, as less-organized con-

vection is more common in COARE, and when convective systems are organized into squall lines, they are typically shorter lived than those in GATE (LeMone et al. 1998).

Rotation is also made on the first three EOFs, which explain 95.8% of the total variance (Table 1). Figure 13 shows the Q_1 and Q_2 profiles corresponding to these three EOFs. The first two modes are similar to those in Fig. 12. As for the newly identified third modes, the Q_1 profile has nearly the same shape as the Q_2 profile, with its maximum centered around the 4-km level. We suspect that, with Q_R typically negative throughout the troposphere, the $Q_1 - Q_R$ and Q_2 profiles for the third mode might be slightly separated, as the case for the third mode in GATE, and represents a weak convective structure. However, we are not certain about the significance of this mode because it contributes only 6% of the Q_1 and Q_2 variation.

More detailed analysis of the COARE data, in which

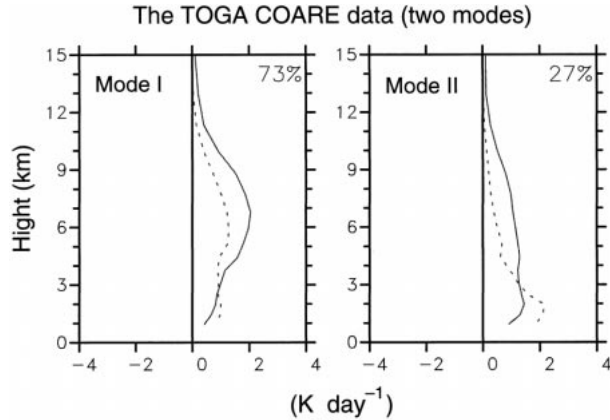


FIG. 12. Vertical profiles of the combined vector (Q_1 , Q_2) associated with the two modes extracted by performing the revised $RPCA_{Promax}$ on the TOGA COARE data over the IFA region. Solid lines are for Q_1 and dashed lines are for Q_2 . The relative percentage of contribution from each mode is indicated in the upper-right corner of each panel.

the time series of the extracted modes are related to other observational studies (e.g., DeMott and Rutledge 1998; Johnson et al. 1999) has been presented by Tung et al. (1999). Of the three basic modes identified by them, one mode is very similar to the first mode in Figs. 12 and 13 and is found to be the dominant mode during the convective phase of Madden-Julian oscillation.

d. Analysis of the ECMWF Re-Analysis data

The revised $RPCA_{Promax}$ was further applied to the combined vectors (Q_1 , Q_2) derived from four subsets of the ECMWF Re-Analysis data. About 86.0%–90.5% of the observed Q_1 and Q_2 variance can be explained with the first two EOFs retained (Table 1). The Q_1 and Q_2 profiles corresponding to the two rotated EOFs are shown in Fig. 14. We see that similarity seems to exist among the modes identified from different regions. For

the first mode, the maxima of Q_1 appear approximately at the 6-km level while the overall Q_2 profiles are slightly lower than the Q_1 profiles except for the northeastern United states, for which the separation between the maximum of Q_1 and that of Q_2 is nearly 3 km wide. As for the second modes, their maxima of Q_2 are generally centered near the 2-km level and the vertical distributions of Q_1 are relatively uniform compared to the Q_2 .

When the first three EOFs are retained, they together explain 91.0%–93.9% of the total variance (Table 1). Figure 15 shows the Q_1 and Q_2 profiles corresponding to these three EOFs. The first two modes are similar to those in Fig. 14 except for the northeastern United states, for which the Q_2 of the first mode is larger than Q_1 at all levels and the Q_2 of the second mode is nearly zero above the 5-km level. As for the newly identified third modes, the Q_1 profiles all have similar shapes with maxima centered in the middle levels but the Q_2 profiles are different from one region to another.

Rotations on larger numbers of the EOFs have also been performed for GATE, TOGA COARE, and ECMWF Re-Analysis data. The results, however, are not presented here for the reasons that will be explained in the next section.

6. Discussion and conclusions

In this study we attempt to empirically identify the basic modes of cumulus heating and drying profiles, which can be used as a partial closure for cumulus parameterization. With such an objective in mind, we analyzed the Q_1 and Q_2 derived from observations using the rotated principal component analysis based on the Promax rotation ($RPCA_{Promax}$). We revise the $RPCA_{Promax}$ in such a way that the distortion of identified structures due to the use of a linear model is minimized. The revised $RPCA_{Promax}$, together with some selected statistical tools, are evaluated using synthetic datasets before it is applied to the observations. The results indicate

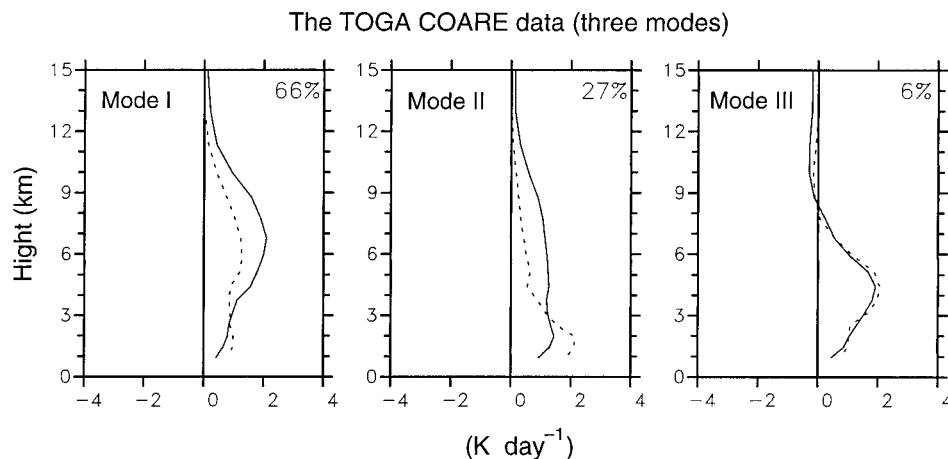


FIG. 13. Same as Fig. 12 except that the results are obtained with three modes retained in the analysis.

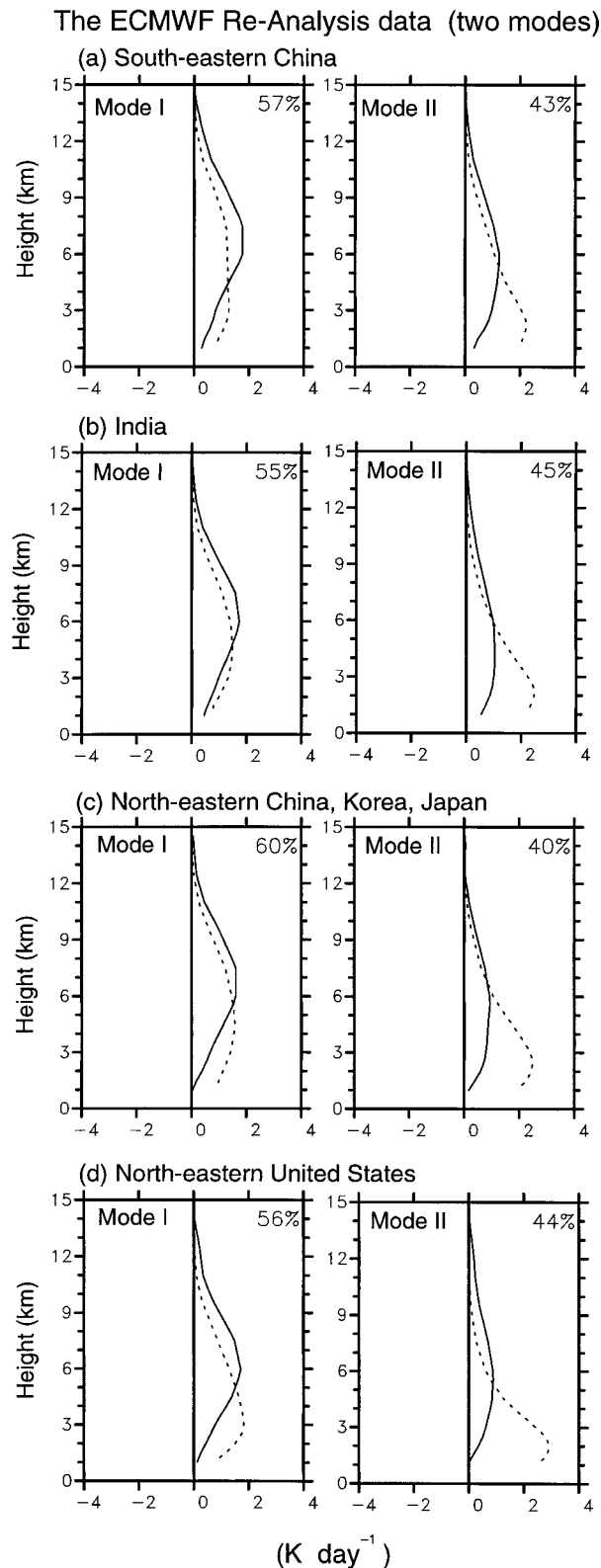


FIG. 14. Vertical profiles of the combined vector (Q_1 , Q_2) associated with the two modes extracted by performing the revised $RPCA_{Promax}$ on the four subsets of the ECMWF Re-Analysis data: (a) southeastern China; (b) India; (c) northeastern China; Korea; Japan;

that, in general, the revised $RPCA_{Promax}$ is a useful tool for identifying the basic modes embedded within a given data.

Ideally, the identified basic modes should be universal to a certain degree so that they can be used as a partial closure for cumulus parameterization. With this consideration in mind, the revised $RPCA_{Promax}$ is performed on the data from different convectively active regions, including the GATE Phase III data, TOGA COARE IOP data over the IFA region, and four subsets of the ECMWF Re-Analysis data that cover areas ranging from tropical to midlatitude continents. The results obtained from retaining two and three modes have been presented for these datasets separately in section 5.

When two modes are retained, they can explain 88.5% of the variance for the GATE Phase III data, 93.6% for the TOGA COARE over the IFA region, and 86.0%–90.5% for the ECMWF Re-Analysis data (Table 1). A comparison between the Q_1 and Q_2 profiles obtained from these six sets of data indicate similarity among the first modes as well as the second modes (Figs. 10a, 12, and 14). Minor differences, however, do exist between the first mode derived from the northeastern United States and those derived from the other three subsets of the ECMWF Re-Analysis data (as already discussed in section 5). Discrepancies, although secondary, are also visible between the second mode derived from the GATE Phase III and those from the four subsets of the ECMWF Re-Analysis data: Q_2 of the former is generally smaller than Q_1 while it is the opposite for the latter.

When three modes are retained, they can explain 93.1% of the variance for the GATE Phase III data, 95.8% for TOGA COARE over the IFA region, and 91.0%–93.9% for the ECMWF Re-Analysis data (Table 1). While additional 2%–5% of the variances are explained by adding the third modes, the similarity among the Q_1 and Q_2 profiles obtained from these six sets of data becomes questionable (Figs. 11a, 13, and 15). Except for the differences already discussed in section 5c, the second and third modes derived from the GATE Phase III are found to deviate significantly from any of the modes derived from the four subsets of the ECMWF Re-Analysis data. The lack of similarity among the derived Q_1 and Q_2 profiles is simply too high a price to pay for explaining additional 2%–5% of the variances, particularly from the viewpoint of providing a closure for cumulus parameterization. We thus discontinue further discussion of the results obtained from retaining larger number of modes, whose corresponding Q_1 and Q_2 profiles exhibit even greater discrepancies among each other (not shown).

The two modes derived from the GATE Phase III

←

and (d) northeastern United States. Solid lines are for Q_1 and dashed lines are for Q_2 . The relative percentage of contribution from each mode is indicated in the upper-right corner of each panel.

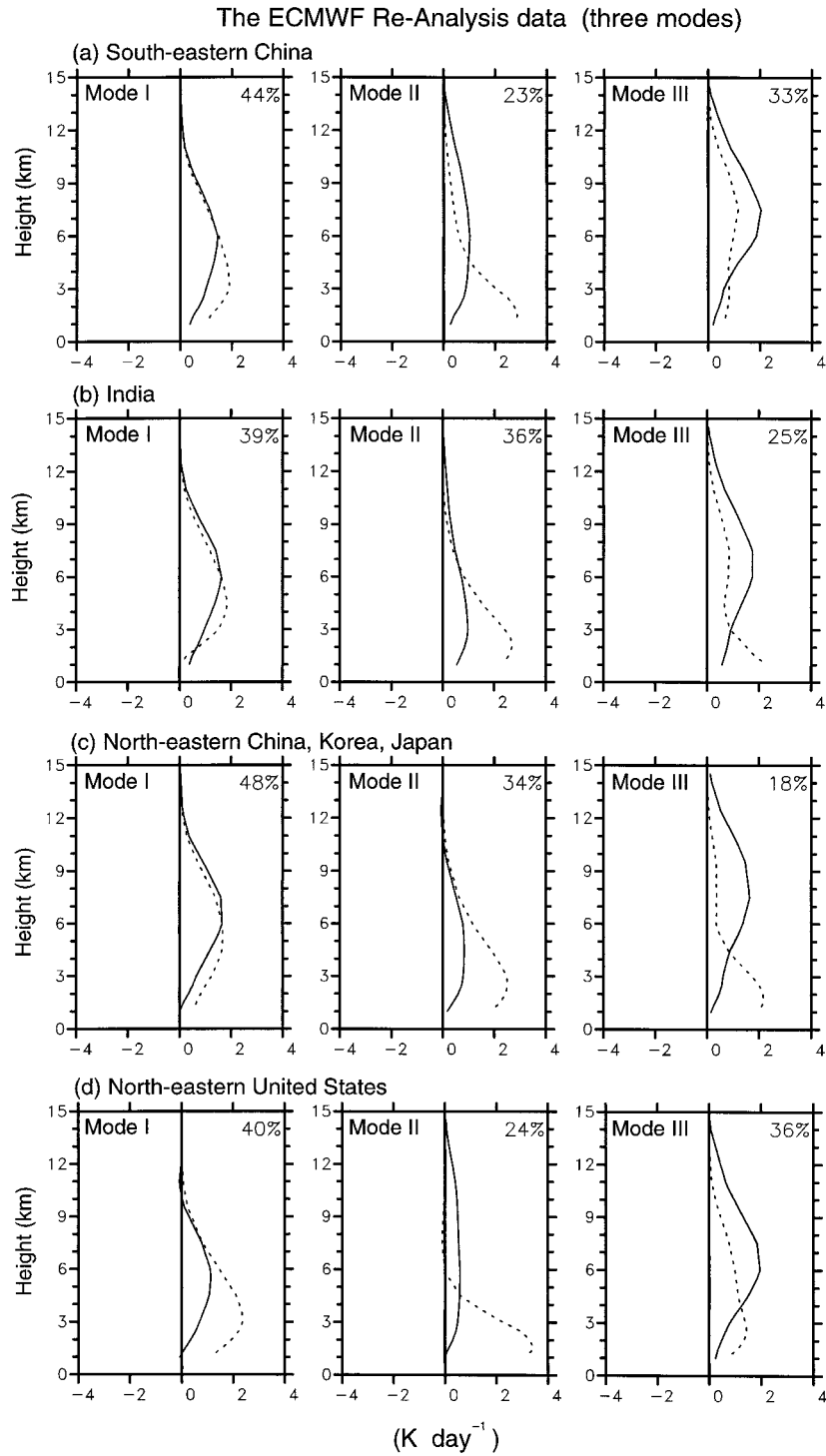


FIG. 15. Same as Fig. 12 except that the results are obtained with three modes retained in the analysis.

data (Fig. 10a), the TOGA COARE data over the IFA region (Fig. 12), or the ECMWF Re-Analysis data (Fig. 14) are probably the best candidates for the basic modes mentioned in the beginning of this paper. These two

modes are favorable not only for their apparently small regional differences but also for their capability of explaining the observed Q_1 and Q_2 variances (86.0%–93.6%). It is further noted that these two modes have

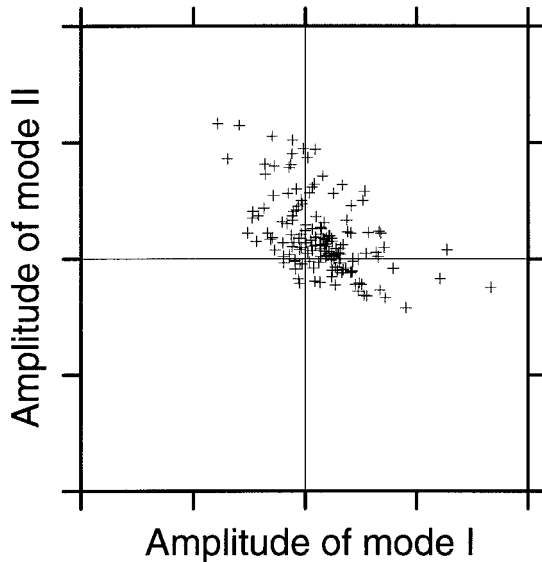


FIG. 16. The scatter diagram of the amplitudes for the two modes extracted from the GATE Phase III data. The abscissa and ordinate represent the amplitudes of the first and second mode. (These amplitudes are readily available from the time series shown in Fig. 10b.)

different amplitudes in time for each region. In other words, their relative contributions to the observed Q_1 and Q_2 variations change in different environments. In this sense, these two modes are similar to the cloud types in the spectral cumulus ensemble model (Arakawa and Schubert 1974), in which the amplitude of mass flux for each cloud type is determined by the concurrent large-scale conditions.

In connection with the spectral cumulus ensemble model, further comment on these two basic modes can be made on the scatter diagram of the basic modes' amplitudes. As an example, Fig. 16 shows the scatter diagram in which the abscissa and ordinate are the amplitudes of the first and second modes from the GATE data. (These amplitudes are readily available from the time series shown in Fig. 10b.) Compared to Richman's (1986) schematic examples of various degrees of simple structure, the distribution in Fig. 16 might be considered moderately simple, with a portion of the points significantly falling off the axes. Since the $RPCA_{\text{Promax}}$, as illustrated in section 4, should be able to give a strong simple structure if such a structure does exist, the moderate complexity of Fig. 16 may indicate the inevitable destiny of adopting a spectral representation of convection even though maximum localization of basic modes in time is preferred.

While the cloud types in the spectral cumulus ensemble model can be interpreted as groups of saturated updrafts with the same maximum heights (Lin and Arakawa 1997; Lin 1999), our understanding of these two modes is still limited. What we have learned from our analysis is that an approximate correspondence between the two modes and squall (nonsquall) clusters seems to

exist. The two modes probably can be interpreted as deep upward motions with moist and dry middle-tropospheric humidity corresponding to the times of nonsquall and squall clusters. Collaborating with our colleagues at the University of California, Los Angeles, we continue to explore the link between the basic modes and their corresponding convective events documented in other observational studies. While having published some preliminary results in this effort (see Tung et al. 1999), we feel that more supporting evidence should be sought to avoid falsely interpreting the results obtained from statistical eigenanalysis as cautioned by Newman and Sardeshmukh (1995).

In conclusion, our analysis results are encouraging in that they suggest a possible avenue for replacing the normalized cumulus heating and drying profiles of each cloud type in a CMCP by those empirically determined. However, readers should be reminded that these results are subject to the problem of data accuracy particularly for ω , which is computed using the divergent component of observed wind profiles. The usefulness of the results should be ultimately judged by the performance of their future applications. Moreover, since our analysis are based on the data from convectively active regions, the extracted basic modes may fail to include some cloud regimes such as trade-wind cumulus, which might not be important in disturbed cases but dominant in other situations. Observations taken from more diverse large-scale environments should be included for the analysis in order to better represent the entire cloud spectrum.

Last but not least, we still need to ask the following question: Are the regional differences of the two basic modes significant despite of our claiming their apparent similarity? From the viewpoint of cumulus parameterization, this question can be answered only after evaluating how sensitive the atmospheric dynamics is to the cumulus heating and drying profiles. If the fields simulated by a large-scale numerical model change significantly when different profiles are used, the differences of basic profiles would be considered significant. In this regard, we have been planning to investigate such sensitivities and hope to report the results in the future.

Acknowledgments. The authors would like to thank Professor Michio Yanai and Ms. Wen-Wen Tung for providing the Q_1 and Q_2 that they derived from the TOGA COARE data. The authors also thank two anonymous reviewers for their helpful comments and suggestions, which have led to significant improvement of the original manuscript. This research is jointly supported by NOAA Grant NA66GP0218, NSF Grant ATM-9613979, and NASA Grant NAG5-4424. The computations were performed at the Department of Atmospheric Sciences at UCLA and the NCAR Scientific Computing Division. NCAR is sponsored by the NSF.

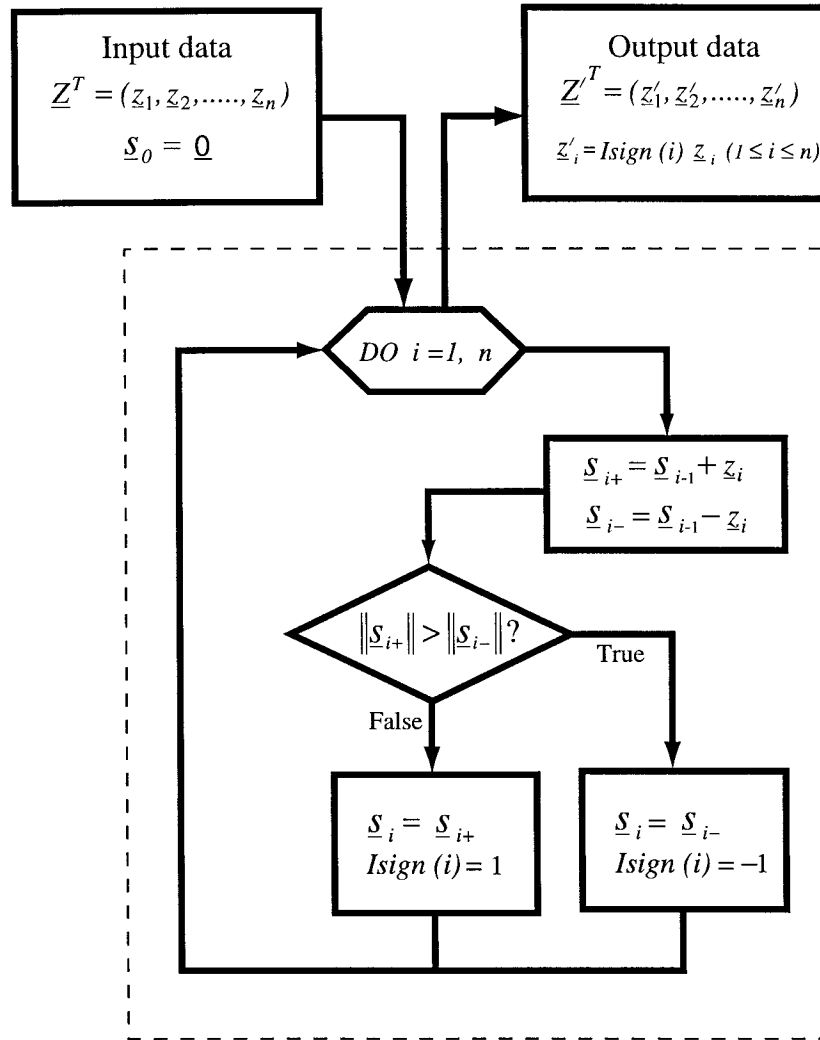


FIG. A1. Flow diagram for the procedure used to reduce the magnitude of the data's ensemble mean prior to analysis.

APPENDIX A

The Analysis Procedure Used in This Study

In this section, the same notations introduced in section 3 will be used when we describe the analysis procedure, which is a revised version of the RPCA based on the Promax rotation (RPCA_{Promax}).

The analysis procedure begins, as a typical approach in other RPCA, with the dataset **Z** first decomposed into a product of **F** and **A**^T [see (3.6)] through PCA. The Promax rotation is then performed on the **F** and **A**^T to obtain the **Z**'s embedded structures, which realize as column vectors in the matrix **A*** [see (3.7)]. The most important assumption made in the Promax rotation is probably that the **F*** obtained by an orthogonal rotation is fairly close to that obtained by an optimum oblique rotation. Therefore the orthogonal solution for **F*** can

be used as a basis for constructing a hypothetical solution, **G***, that satisfies the simple structure criterion even better and is thus treated as the target solution to be approximated.

The orthogonal rotation that is selected to give the preliminary solution is the Varimax rotation (Kaiser 1958). The normalized Varimax criterion, which is expressed by Eq. (21) in Richman (1986), is used in this study following the recommendation by Mulaik (1972). In short, the Varimax rotation (using the normalized Varimax criterion) maximizes the variance of the squared loadings in the **F*** in an attempt to produce a **F*** with as many zero (or near zero) loadings as possible. This method usually gives solutions considered to be very good orthogonal simple structure and is perhaps the only orthogonal rotation that has been widely used (Mulaik 1972; Harman 1976; Richman 1986).

Let the preliminary solution for \mathbf{F}^* be denoted by $\mathbf{F}_{\text{init}}^*$ and their elements denoted by f_{ij} . The target matrix \mathbf{G}^* can be defined with elements g_{ij} such that

$$g_{ij} = \frac{|f_{ij}^{m+1}|}{f_{ij}}, \quad (\text{A.1})$$

where m is a power to which the element f_{ij} is raised. The reason for raising the elements of the preliminary solution to the m th power ($m > 1$) is to make those elements near zero in magnitude approach zero more rapidly than elements farther away from zero. As a result, the \mathbf{G}^* would better satisfy the simple structure criterion since the number of near zero elements in the \mathbf{G}^* is increased through the adjustment described by (A.1). We set m equal to 2 following Richman's (1986) recommendation.

The last step of the Promax rotation is to find a rotation matrix, \mathbf{T}_p , such that the $\mathbf{F}_{\text{init}}^* \mathbf{T}_p$ fits into the target matrix \mathbf{G}^* as closely as possible. This procedure is usually referred to as the Procrustean rotation (Hurley and Cattell 1962), in which the rotation matrix \mathbf{T}_p is determined through a regression technique. [A description of the regression technique can be found in Mulaik (1972), Harman (1976), or Richman (1986).]

As mentioned in section 4, the performance of $\text{RPCA}_{\text{Promax}}$, as a linear model, might be affected by removing the ensemble mean of the data if the magnitude of the ensemble mean is comparable to that of the anomalies themselves. We therefore propose reducing the magnitude of the data's ensemble mean prior to analysis. In practice, this can be done through multiplying each \mathbf{z}_i ($1 \leq i \leq n$) in (3.1) by a sign function $\text{Isign}(i)$, which is determined as follows (Fig. A1): Let $\mathbf{s}_0 = 0$ and \mathbf{s}_i ($1 \leq i \leq n$) be defined by

$$\mathbf{s}_i = \begin{cases} \mathbf{s}_{i+} \equiv \mathbf{s}_{i-1} + \mathbf{z}_i & \text{if } \|\mathbf{s}_{i+}\| \leq \|\mathbf{s}_{i-}\| \\ \mathbf{s}_{i-} \equiv \mathbf{s}_{i-1} - \mathbf{z}_i & \text{if } \|\mathbf{s}_{i+}\| > \|\mathbf{s}_{i-}\|, \end{cases} \quad (\text{A.2})$$

where the notation $\|\cdot\|$ designates the norm of a vector. Starting from $i = 1$, we can obtain \mathbf{s}_i in sequence and determine the value of $\text{Isign}(i)$ through

$$\text{Isign}(i) = \begin{cases} 1 & \text{if } \mathbf{s}_i = \mathbf{s}_{i+} \\ -1 & \text{if } \mathbf{s}_i = \mathbf{s}_{i-}. \end{cases} \quad (\text{A.3})$$

With such a manipulation, the ensemble mean of the newly transformed \mathbf{z}_i (i.e., the \mathbf{z}'_i in Fig. A1) will be reduced and the aforementioned deficiency of a linear model would be mitigated. (In fact the ensemble mean will be nearly zero when a sufficiently large sample of data are used as input.) The embedded structures of the data, on the other hand, are not affected through multiplying the \mathbf{z}_i by $\text{Isign}(i)$, which takes a value of either 1 or -1 . The i th principal component derived by the $\text{RPCA}_{\text{Promax}}$, however, should be multiplied by $\text{Isign}(i)$ again so that (3.7) still holds for the input array \mathbf{Z} .

APPENDIX B

The H -mode $\text{RPCA}_{\text{Promax}}$

With a dataset \mathbf{Z} given as described in section 3, we can derive n eigenvectors (EOFs) from the covariance matrix $\mathbf{Z}\mathbf{Z}^T$, and the dataset \mathbf{Z} can be formally written as (3.6). Here \mathbf{A} is an $n \times p$ matrix with PCs as its row vectors, representing the derived vertical profiles; \mathbf{F} is an $n \times n$ matrix with EOFs as its column vectors, representing time variation of the amplitudes for each vertical profile. To seek structures that have maximum localization in height, we rotate EOFs/PCs in the same manner as described by (3.7) but with \mathbf{A}^{*T} targeted as the object of simple structure. For more details of this approach, readers are referred to Alexander et al. (1993), Liu (1995), or Misra (1997).

REFERENCES

- Alexander, G. D., G. S. Young, and D. V. Ledvina, 1993: Principal component analysis of vertical profiles of Q_1 and Q_2 in the Tropics. *Mon. Wea. Rev.*, **121**, 535–548.
- Arakawa, A., 1993: Closure assumptions in the cumulus parameterization problem. *The Representation of Cumulus Convection in Numerical Models*, Meteor. Monogr., No. 46, Amer. Meteor. Soc., 1–15.
- , and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.*, **31**, 674–701.
- , and J.-M. Chen, 1987: Closure assumptions in the cumulus parameterization problem. *Short- and Medium-Numerical Prediction*, T. Matsuno, Ed., Meteorological Society of Japan, 107–131.
- Cattell, R. B., 1966: The scree test for the number of factors. *Mult. Behav. Res.*, **1**, 245–276.
- Chen, J.-M., 1989: Observational study of the macroscopic behavior of moist-convective processes. Ph.D. dissertation, University of California, Los Angeles, 264 pp. [Available from UCLA, Los Angeles, CA 90095.]
- Cheng, M.-D., and M. Yanai, 1989: Effects of downdrafts and mesoscale convective organization on the heat and moisture budgets of tropical cloud clusters. Part III: Effects of mesoscale convective organization. *J. Atmos. Sci.*, **46**, 1566–1588.
- Cooley, W. W., and P. R. Lohnes, 1971: *Multivariate Data Analysis*. Wiley, 364 pp.
- Cox, S. K., and K. T. Griffith, 1979: Estimates of radiative divergence during Phase III of the GARP Atlantic Tropical Experiment. Part I: Methodology. *J. Atmos. Sci.*, **36**, 576–585.
- DeMott, C. A., and S. A. Rutledge, 1998: The vertical structure of TOGA COARE convection. Part I: Radar echo distributions. *J. Atmos. Sci.*, **55**, 2730–2747.
- Esbensen, S. K., E. I. Tollerud, and J.-H. Chu, 1982: Cloud-cluster-scale circulations and the vorticity budget of synoptic-scale waves over the eastern Atlantic intertropical convergence zone. *Mon. Wea. Rev.*, **110**, 1677–1692.
- Frank, W. M., H. Wang, and J. L. McBride, 1996: Rawinsonde budget analyses during the TOGA COARE IOP. *J. Atmos. Sci.*, **53**, 1761–1780.
- Gibson, J. K., P. Källberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA description. ECMWF Re-Analysis Final Report Series, Vol. 1, 71 pp.
- Harman, H. H., 1976: *Modern Factor Analysis*. The University of Chicago Press, 487 pp.
- Harris, C. W., and H. F. Kaiser, 1964: Oblique factor analytic solutions by orthogonal transformations. *Psychometrika*, **29**, 347.
- He, H., J. W. McGinnis, Z. Song, and M. Yanai, 1987: Onset of the

- Asian summer monsoon in 1979 and the effects of the Tibetan Plateau. *Mon. Wea. Rev.*, **115**, 1966–1995.
- Hendrickson, A. E., and P. O. White, 1964: Promax: A quick method for rotation to oblique simple structure. *Br. J. Stat. Psychol.*, **17**, 65–70.
- Hurley, J. R., and R. B. Cattell, 1962: The Procrustes program: Producing direct rotation to test a hypothesized factor structure. *Behav. Sci.*, **7**, 258.
- Johnson, R. H., T. M. Rickenbach, S. A. Rutledge, P. E. Ciesielski, and W. H. Schubert, 1999: Trimodal characteristics of tropical convection. *J. Climate*, **12**, 2397–2418.
- Jolliffe, I. T., 1986: *Principal Component Analysis*. Springer-Verlag, 271 pp.
- , 1987: Rotation of principal components: Some comments. *J. Climatol.*, **2**, 313–329.
- Kaiser, H. F., 1958: The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187.
- Kuo, H.-L., 1965: On formation and intensification of tropical cyclones through latent heat release by cumulus convection. *J. Atmos. Sci.*, **22**, 40–63.
- , 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.*, **31**, 1232–1240.
- LeMone, M. A., E. J. Zipser, and S. B. Trier, 1998: The role of environmental shear and thermodynamic conditions in determining the structure and evolution of mesoscale convective systems during TOGA COARE. *J. Atmos. Sci.*, **55**, 3493–3518.
- Lin, C., 1999: Some bulk properties of cumulus ensembles simulated by a cloud-resolving model. Part II: Entrainment profiles. *J. Atmos. Sci.*, **56**, 3736–3748.
- , and A. Arakawa, 1997: The macroscopic entrainment processes of simulated cumulus ensemble. Part II: Testing the entraining-plume model. *J. Atmos. Sci.*, **54**, 1044–1053.
- Liou, K.-N., 1992: *Radiation and Cloud Processes in the Atmosphere: Theory, Observation and Modeling*. Oxford Monogr. on Geology and Geophysics, No. 20, Oxford, 487 pp.
- Liu, Y.-Z., 1995: The representation of the macroscopic behavior of observed moist-convective processes. Ph.D. dissertation, University of California, Los Angeles, 227 pp. [Available from UCLA, Los Angeles, CA 90095.]
- Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Sci. Rep. 1, Statistical Forecasting Project, Department of Meteorology, Massachusetts Institute of Technology, 48 pp. [Available from MIT, Cambridge, MA 02139.]
- Manabe, S., J. Smagorinsky, and R. F. Strickler, 1965: Simulated climatology of a general circulation model with a hydrological cycle. *Mon. Wea. Rev.*, **93**, 769–798.
- McBride, J. L., B. W. Gunn, G. J. Holland, T. D. Keenan, N. E. Davidson, and W. M. Frank, 1989: Time series of total heating and moistening over the Gulf of Carpentaria radiosonde array during AMEX. *Mon. Wea. Rev.*, **117**, 2701–2713.
- Misra, V., 1997: A statistically based cumulus parameterization scheme that makes use of heating and moistening rates derived from observations. Ph.D. dissertation, The Florida State University, 184 pp. [Available from The Florida State University, Tallahassee, FL 32306.]
- Mulaik, S. A., 1972: *The Foundations of Factor Analysis*. McGraw-Hill, 453 pp.
- Newman, M., and P. Sardeshmukh, 1995: A caveat concerning singular value decomposition. *J. Climate*, **8**, 352–360.
- Nitta, T., 1977: Response of cumulus updraft and downdraft to GATE A/B-scale motion systems. *J. Atmos. Sci.*, **34**, 1163–1186.
- Ooyama, K., 1987: Scale-controlled objective analysis. *Mon. Wea. Rev.*, **115**, 2476–2506.
- Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, 425 pp.
- , and T. P. Barnett, 1977: Significance tests for empirical orthogonal functions. Preprints, *Fifth Conf. on Probability and Statistics in Atmospheric Sciences*, Las Vegas, NV, Amer. Meteor. Soc., 169–172.
- Richman, M. B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293–335.
- , and P. J. Lamb, 1985: Climatic pattern analysis of three- and seven-day summer rainfall in the central United States: Some methodological considerations and a regionalization. *J. Climate Appl. Meteor.*, **24**, 1325–1343.
- Sui, C.-H., and M. Yanai, 1986: Cumulus ensemble effects on the large-scale vorticity and momentum fields of GATE. Part I: Observational evidence. *J. Atmos. Sci.*, **43**, 1618–1642.
- Thurstone, L. L., 1947: *Multiple Factor Analysis*. The University of Chicago Press, 535 pp.
- Tung, W.-W., C. Lin, B. Chen, M. Yanai, and A. Arakawa, 1999: Basic modes of cumulus heating and drying observed during TOGA-COARE IOP. *Geophys. Res. Lett.*, **26**, 3117–3120.
- Vargas, W. M., and R. H. Compagnucci, 1983: Methodological aspects of principal component analysis in meteorological fields. Preprints, *Second Int. Meeting on Statistical Climatology*, Lisbon, Portugal, National Institute of Meteorology and Geophysics, 5.3.1.
- Yanai, M., S. Esbensen, and J. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611–627.