JOURNAL OF CLIMATOLOGY, VOL. 6, 293–335 (1986) 551.501.45:519.23

REVIEW ARTICLE

ROTATION OF PRINCIPAL COMPONENTS

MICHAEL B. RICHMAN

Climate and Meteorology Section, Illinois State Water Survey, Champaign, Illinois 61820-7407, U.S.A.

Received 17 April 1985
Revised 4 September 1985

ABSTRACT

Recent research has pointed to a number of inherent disadvantages of unrotated principal components and empirical orthogonal functions when these techniques are used to depict *individual* modes of variation of data matrices in exploratory analyses. The various pitfalls are outlined and illustrated with an alternative method introduced to minimize these problems via available linear transformations known as simple structure rotations. The rationale and theory behind simple structure rotation and Procrustes target rotation is examined in the context of meteorological/climatological applications. This includes a discussion of the six unique ways to decompose a rotated data set in order to maximize the physical interpretability of the rotated results.

The various analytic simple structure rotations available are compared by a Monte Carlo simulation, which is a modification of a similar technique developed by Tucker (1983), revealing that the DAPPFR and Promax $k = 2$ rotations are the most accurate in recovering the input structure of the modes of variation over a wide range of conditions. Additionally, these results allow the investigator the opportunity to check the accuracy of the unrotated or rotated solution for specific types of data. This is important because, in the past, the decision of whether or not to apply a specific rotation has been a 'blind decision'. In response to this, a methodology is presented herein by which the researcher can assess the degree of simple structure embedded within any meteorological data set and then apply known information about the data to the Monte Carlo results to optimize the likelihood of achieving physically meaningful results from a principal component analysis.

KEY WORDS Orthogonal rotation Oblique rotation Simple structure Procrustes target transformation Principal components Eigenvectors Exploratory analyses Confirmatory analyses Monte Carlo techniques

1. INTRODUCTION

The eigentechniques have been used with increasing frequency since 1980; no fewer than eighty applications of empirical orthogonal functions (EOFs), principal components (PCs) and common factor analysis (CFA) were published in the meteorological literature during this period. A characteristic of this proliferation has been the varied types of ways in which analyses are specified. For instance, there are different dispersion matrices used to relate data in these analyses (e.g. correlation, covariance, cross-products), three basic eigenmodels (EOFs, PCs, CFA), numerous procedures to determine the optimum number of significant eigenmodes, a number of alternative solutions for interpreting the output (e.g. orthogonal rotations, oblique rotations, Procrustes target rotations), and six different ways to specify a data set prior to the analysis, within these alternative solutions, each of which brings out different relationships among the data. Consequently, an analysis of a single data set can yield a number of physically plausible interpretations based on the options outlined, and all of the options need to be carefully considered in order for the analysis to yield physically meaningful results. Furthermore, the eigentechniques have been used to address such diverse objectives as pure data reduction (Weickmann, 1983), a tool to group synoptic (Christenson and Bryson, 1966) and chemical variables (Gatz, 1978), a technique to aid in forecasting atmospheric parameters (Gilman, 1957), a

0196–1748/86/030293–43$21.50

method to aid in determining the variability of atmospheric fields (Hastenrath and Wendland, 1979; Horel, 1981) and a procedure to identify coherent modes of various parameters such as geopotential height (Cohen, 1983) and rainfall (Dyer, 1975).

There has been some lengthy discussion of the effects of choosing various dispersion matrices in the early meteorological literature. Craddock (1965), Kutzbach (1967, 1969) and Craddock and Flood (1969) discuss the relative merits of a covariance matrix which includes the analysis of actual magnitudes of the variances (as opposed to relative variations in a correlation matrix) whereas Gilman (1957), Sellers (1957) and Glahn (1965) discuss the advantages of a correlation input which equally weights all variables (e.g. stations) in an analysis. The cross-products matrix incorporates information on the mean of each variable as well as its variation into the analysis and has been discussed by Resio and Hayden (1975) and Molteni *et al.* (1983). The three eigentechniques have all been used in meteorological research. The most common application has employed EOFs or unit length eigenvectors; this was pioneered in meteorological work by Lorenz (1956) and has remained essentially unchanged ever since. In contrast is the mathematically more elegant principal component model (Hotelling, 1933) which weights the eigenvectors by the square root of the corresponding eigenvalue, so that the weights (known as loadings) represent the correlations (covariances) between each variable and each principal component. The first application of this technique to meteorological data was by Fukuoka (1951). Although both EOFs and PCs have been widely used in the past few years, the common factor model (CFA) has rarely been employed in meteorological applications. This model excludes measurement error (uncorrelated white noise) and variance unique to each individual variable from the analysis, dealing only with that portion of the total variance which can be predicted from the other variables (hence the term common). Walsh *et al.* (1982) and Richman (1983a) apply the common factor model to precipitation data and geopotential height and discuss its relative advantages.

Methods for determining where to stop extracting eigenmodes have been proposed in the meteorological literature by Craddock and Flood (1969) and Preisendorfer and Barnett (1977). Craddock and Flood's LEV test graphs the natural logarithm of each eigenvalue against the mode number. The point where the tail of the plot becomes linear is where no further EOFs are retained, since the linear tail with nearly equal eigenvalues theoretically represents random noise. This test is a variation of the 'scree' test which was proposed by Cattell (1966a), which plots each raw eigenvalue against root number. Cohen (1983) illustrates a variation of the scree test on meteorological data where he rotates a data set, prior to graphing up the eigenvalues, claiming a more distinct break in the curve. Preisendorfer and Barnett (1977) offer a test based on a Monte Carlo criterion where a matrix of random normal deviates, of the same direction as the data matrix, is formed into a dispersion matrix and then decomposed into a series of eigenvalues. A number of trials are completed, confidence intervals constructed and then compared to the data eigenvalues. The point where the data eigenvalues become smaller than the corresponding Monte Carlo confidence interval is where no further PCs/EOFs are retained, as shown by Overland and Preisendorfer (1982). It is a variation of the test that Horn (1965) proposed without confidence intervals, whereas Humphreys and Montanelli (1975) devised another modification of this test for the common factor model. Other tests have appeared in the literature, many of which are reviewed in Rummel's (1970) text, applied to a chemical data set in Thurston and Spengler (1985), to various physical systems in Preisendorfer *et al.*'s (1981) report and recently compared for accuracy by Hakstian *et al.* (1982) and by Zwick and Velicer (1982).

The question concerning the applicability of an alternative solution or linear transformation (known as a 'rotation') has not been addressed fully in the meteorological literature. A few early applications of rotation were employed by Gregory (1975), Dyer (1975), Rogers (1976), Willmott (1977), Gatz (1978), Morin *et al.* (1979) and Salinger (1980) in which a particular criterion (Varimax) was applied to identify homogeneous groups of variables. However, the majority of earlier and recent studies apply the traditional approach of unrotated EOFs to the analysis to attempt to characterize the modes of variation of a field under investigation. All three eigenmodels (EOF, PCA, CFA) allow for an *infinite* number of alternative final solutions which satisfy the models' equations, a fact which will serve as motivation for this paper, which will present the theory behind simple structure rotation and outline the

meteorological benefits of rotating various data sets. In order to facilitate this, section 2 will provide empirical evidence of problems associated with unrotated solutions which are reduced, and in some cases eliminated, by the use of rotation. One eigenmodel (PCA) is then developed and shown to be valid for any non-singular linear transformation (section 3), a characteristic which is capitalized on in the detailed discussion of simple structure and its graphical presentation (sections 4 and 5, respectively). Section 6 provides a discussion of how best to analyse a data matrix to bring out the most physically meaningful meteorological relationships in a rotated analysis and section 7 develops an alternative to simple structure rotations, the Procrustes transformation, and provides some ideas for its applications. The analytic rotations are individually explored and then compared in a Monte Carlo analysis in sections 8 and 9. The resulting findings are presented in section 10, which is useful to meteorologists in that it provides a more finely calibrated tool for the optimal choice of a solution (e.g. unrotated, orthogonally rotated, obliquely rotated) in order to most accurately determine the modes of variation in an analysis.

2. WHY BOTHER WITH ROTATION?

A question which immediately arises is: why not simply interpret the unrotated EOFs/PCs? Unrotated solutions offer a number of potential advantages, such as their economy, ability to extract the maximal variance from a data set, their spatial and temporal orthogonality and their pattern insensitivity to the number of PCs retained, whereas the rotated solutions have a set of jargon associated with them (e.g. hyperplanes, intercomponent correlations, oblique angles, etc.) which can be discouragingly foreign to uninitiated meteorologists. Additionally, these characteristics make unrotated solutions good candidates for situations where pure data reduction is sought, in cases where the PCs will be used for regression without individually interpreting each mode and even in certain exploratory applications. The reason why meteorologists should at least consider the possibility of applying a rotation to a data set is that *unrotated solutions exhibit four characteristics which hamper their utility to isolate individual modes of variation.* These four characteristics are domain shape dependence, subdomain instability, sampling problems and inaccurate portrayal of the physical relationships embedded within the input matrix.

Domain shape dependence

The problem of domain shape dependence has been outlined by Buell (1975, 1979) who illustrates that the topographies of unrotated EOFs are primarily determined by the shape of the domain and *not* by the covariation among the data. Specifically, Buell has shown that *different* correlation functions on a geometrically shaped domain have similar EOF patterns in a predictable sequence which do not reflect the underlying covariation. Numerous published meteorological studies (inadvertently) support Buell's findings as their *unrotated* EOFs/PCs of various unrelated parameters for different geographical areas all show the characteristic patterns outlined by Buell; a limited inventory of which was compiled by Richman (1983a). The results of that inventory for studies with a square/rectangular domain (Table I) indicate that the first unrotated EOF/PC is positive (+) in all of the studies, regardless of the meteorological parameter or geographical area treated, whereas unrotated EOF/PC2 is always a flip-flop (+ −). The third to fifth unrotated EOF/PCs have mixed patterns of either (±), (− + −) or (± ∓) which is exactly what Buell (1975, p. 191) found for another correlation function where the third EOF was (±), the fourth (± ∓) and the fifth (− + −). Overall, Table I supports Buell's (1979, p. 117) contention that the unrotated EOF/PC approach produces results that are domain shape dependent, might not yield realistic modes of variation and should be looked upon with suspicion. The predictable progression of the unrotated EOF/PC patterns may be at least partly due to the relationship between EOFs and harmonics. Cahalan (1983) mentions that they are closely related to spherical harmonics (when represented on a sphere) and on a limited domain (e.g. rectangle) they are related to 'patch harmonics'. North (1985) also discusses this phenomenon for a limited number of physical systems. In order for the *rotated solution* to be potentially more useful, it should not suffer from domain shape

Table I. The morphology of various unrotated EOF/PC studies in comparison to Buell's (1975, p. 190, his Figure 5) rectangular domain. The topography of the EOF/PCs are denoted by (+) for one positive anomaly over the entire domain, (+ −) for a positive anomaly in the west and a negative anomaly in the east, (− + −) for a negative anomaly in the west, positive anomaly in the centre and negative anomaly in the east, (±) for a positive anomaly in the north and negative anomaly in the south and, finally (± ∓) for positive anomalies in the north-west and south-east and negative anomalies in the south-west and north-east. The patterns shown are also valid for their inverses (e.g. (+ −) and (− +) are the same pattern) due to the reflective property of EOF/PCs. The domain can also be rotated 90° to be aligned with the longest axis of the rectangle (e.g. (− + −) and ($\frac{\pm}{\mp}$) are the same pattern)

Author, Year	Parameter	EOF/PC1	EOF/PC2	EOF/PC3	EOF/PC4	EOF/PC5
Buell (1975)	Correlated noise	(+)	(+ −)	(− + −)	(±)	(± ∓)
Klaus (1978)	Precipitation	(+)	(∓)	(+ −)	($\frac{\mp}{\pm}$)	
Willmott (1978)	Precipitation	(+)	(∓)	(− +)	($\frac{\pm}{\mp}$)	
Dyer (1979)	Precipitation	(+)	(∓)	($\frac{\pm}{\mp}$)		
Bedi and Bindra (1980)	Precipitation	(+)	(∓)	(+ − +)		
Walsh and Mostek (1980)	Precipitation	(+)	(+ −)	(− + −)		
Diaz (1981)	Precipitation	(+)	(+ −)	(∓)	($\frac{\pm}{\mp}$)	
Diaz and Fulbright (1981)	Precipitation	(+)	(− +)	(∓)	(+ − +)	
Tabony (1981)	Precipitation	(+)	(∓)	($\frac{\pm}{\mp}$)		
Karl and Koscielny (1982)	Precipitation	(+)	(+ −)	(±)	(+ − +)	(∓ ±)
McGuirk (1982)	Precipitation	(+)	(+ −)	(− + −)		
Lamb and Richman (1982)	Precipitation	(+)	(±)	(− +)	($\frac{\pm}{\mp}$)	(∓ ±)
Lorenz (1956)	Sea-level pressure	(+)	(+ −)	(±)	(+ − +)	
Veitch (1965)	Sea-level pressure	(+)	(+ −)	(∓)	(+ − +)	(± ∓)
Davis (1976)	Sea-level pressure	(+)	(∓)	(− +)	(− + −)	(± ∓)
Brier and Meltesen (1976)	Sea-level pressure	(+)	(∓)	($\frac{\mp}{\pm}$)		
Walsh and Mostek (1980)	Sea-level pressure	(+)	(− +)	(− + −)		
Brinkmann (1981)	Sea-level pressure	(+)	(±)	(+ −)	(∓ ±)	
Walsh and Mostek (1980)	Temperature	(+)	(− +)	(±)		
Walsh and Richman (1981)	Temperature	(+)	(− +)	(±)	(− + −)	
Gray (1981)	Temperature	(+)	(∓)	(+ −)	(± ∓)	($\frac{\pm}{\mp}$)
Karl *et al.* (1982)	Temperature	(+)	(+ −)	(±)		

dependence. Kaiser (1958) and Horel (1981, 1984) both state that, theoretically, rotated solutions are less affected by domain dependence. Walsh and Richman (1981, pp. 780–781) and Richman (1981) tested the domain dependence of both an unrotated solution and a rotated solution for a *known* data set of two sea-level pressure patterns which had a scale of over half the domain size, finding that the unrotated PCs yielded the predictable Buell patterns, whereas the rotated PCs recovered the two underlying pressure patterns. Vargas and Compagnucci (1983) apply this approach of known input modes to a meteorological concept, as they began with an example of three *large scale* flow patterns (zonal, meridional and cyclonic) and their inverses on a square domain (Figures 1(a), (b), (c)). A covariance matrix was formed to relate the 36 grid points for the 3 flow types and their inverses which was then decomposed into three unrotated PCs and the PC loadings plotted at the 36 grid points. The first map of unrotated PC loadings (Figure 1(d)) indicates that all coefficients are of the same sign; this has no resemblance to any of the input patterns but, rather, is a combination, or merging, of the three, which is a direct consequence of the maximum variance restriction placed on the first unrotated PC (Richman, 1981; Karl and Koscielny, 1982). The second map of unrotated PC loadings (Figure 1(e) also does not represent any of the input patterns but has a north-west/south-east (+ −) pattern which agrees with Buell's pattern for a square domain. The third map of unrotated PC loadings (Figure 1(f)) has some similarities to the third input pattern but differs in that there is a flip-flop of loadings with the

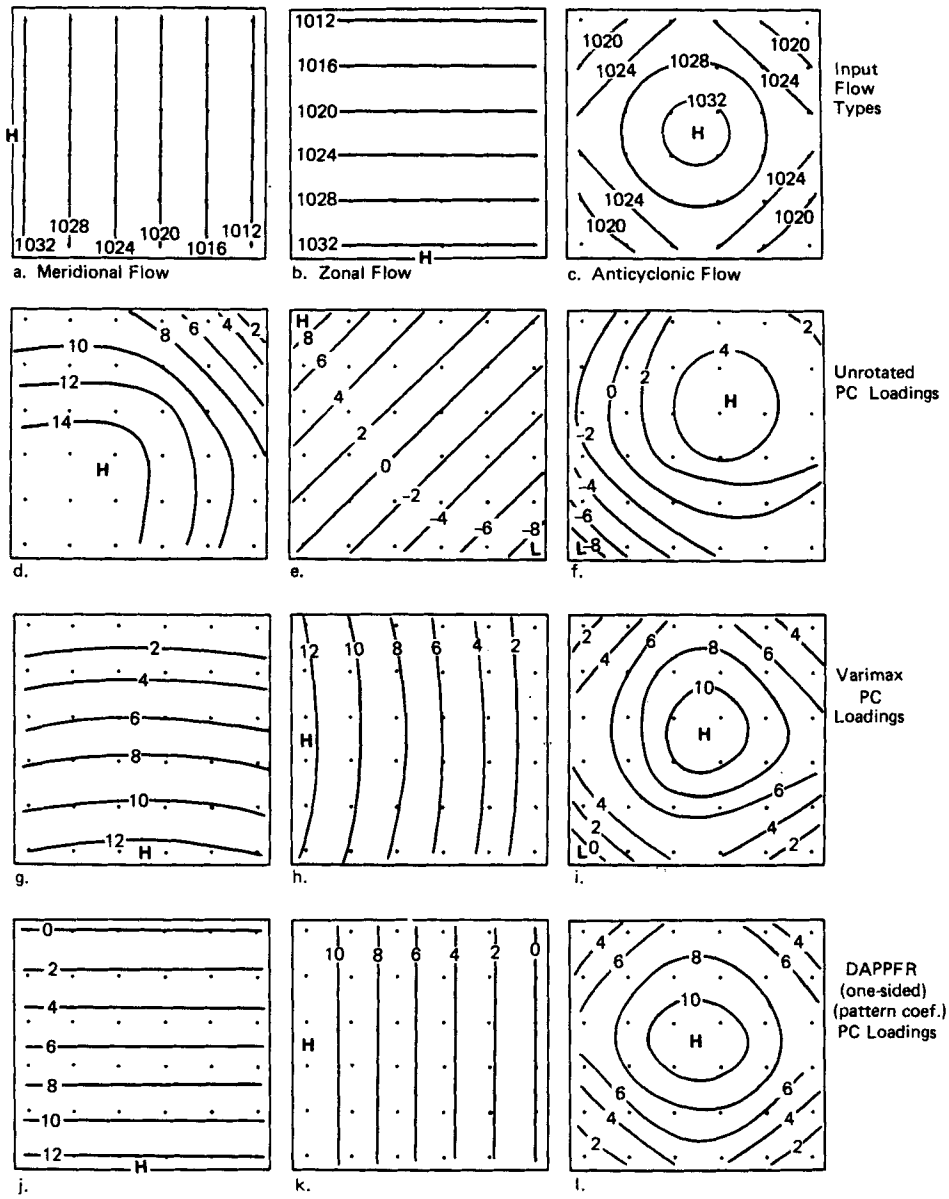

Figure 1. Example of the decomposition of three known input sea-level pressure patterns (in mb) corresponding to meridional, zonal and anticyclonic flow (panels (a), (b), (c)) based on work of Vargas and Compagnucci (1983). The inverses of these three patterns were also included by reflecting the values about a mean of 1012 mb and thus the mean map is constant and can be discarded. Decomposition of the covariance matrix into three unrotated PCs (panels (d), (e), (f)), three orthogonally rotated PCs (panels (g), (h), (i)) and three obliquely rotated PCs (panels (j), (k), (l)) to be compared to the three input flow types

largest magnitude (−8·4) in the south-west portion of the domain, decreasing to lower absolute values in the north-east portion and is similar to what Buell's square domain pattern illustrates. The three unrotated PCs were next rotated orthogonally (Varimax criterion; Kaiser, 1958). The first Varimax map of PC loadings (Figure 1(g)) closely corresponds to input map two, as both represent a zonal flow type, and the second loading map (Figure 1(h)) closely corresponds to input map one with the only difference being the PC loading isolines have a slight curvature. The third Varimax PC loading map (Figure 1(i)) is most similar to the third input map, although there is a small region with one slightly negative PC loading (−0·3) in the south-west corner of the domain. In order to detect if the solution

could be further improved, an oblique rotation (one-sided DAPPFR rotation; Tucker and Finkbeiner, 1982) was applied to the unrotated PCs. The first map of DAPPFR PC loadings (Figure 1(j)) depicts a zonal flow pattern similar to the second input type with a total absence of curvature in the loading isolines and zero PC loadings along the top edge of the domain, corresponding to the zero deviation line in the input mode. The second map of DAPPFR PC loadings (Figure 1(k)) depicts a meridional flow pattern similar to the first input type, with straight isolines throughout the map and the zero isoline on the right edge of the domain, which agrees with the flow type, and the third map of DAPPFR PC loadings (Figure 1(l)) represents anticyclonic flow having all positive loadings and symmetry from the left side of the domain to the right in good correspondence with input type three. This is in better agreement with the input flow types than the unrotated PC loadings (or Varimax) and illustrates how a rotated solution can be used to recover more accurately the input modes with little concern about the shape of the domain.

Subdomain stability

A second disadvantage of unrotated solutions is their subdomain instability, which is a corollary of domain shape dependence and refers to the stability or invariance of the modal patterns as subportions of the domain are examined. If an EOF pattern is to be considered as a mode of variation, it should remain stable if the northern hemisphere is the domain or if only the United States is considered (i.e. as affects the U.S.). Richman and Lamb (1985) present the results of a study of subdomain stability using a United States rainfall data set for 3-day and 7-day totals from May to August. The *unrotated modes* for the full domain (Figure 2(a)) exhibited the Buell-like patterns over the first eight PCs (first four shown). Although they are in a predictable progression, these cannot be simply discounted, as it is possible that the patterns are physically meaningful. The domain was next subdivided into northern and southern portions. The results for each half of the domain (Figure 2(b)) indicate that the PC patterns obtained for the entire domain were simply transferred and compressed into the subareas when the latter were analysed individually. For instance, the first unrotated PC (full domain) had its anomaly centre in Iowa, whereas none of the subdomain modes had a similar pattern. This comparison points to an obvious dilemma created by the unrotated solution—which pattern is the correct mode of decomposition, the full domain or the subdomain? The same experiment was repeated for the *rotated patterns* and is summarized in Figure 3(a) where the 0·4 isoline is plotted. Examination of each half of the domain in Figure 3(a) would lead to the hypothesis that areas 5, 6, 8 and 10 (full domain) should be present in the southern subdomain whereas areas 1, 2, 3, 4, 7 and 9 (full domain) should be present in the northern subdomain. This holds true, as evidenced by Figure 3(b) in which the rotated modes yielded much more encouraging results, as the same basic regionalization emerges in both Figures 3(a) and 3(b). The agreement is nearly identical (maximum difference of $\pm 0{\cdot}02$ at each station loading in each analysis) for the southern half of the domain, whereas for the north only small differences are apparent. The importance of this is that, as a scientific tool, a goal of EOF/PC/CFA is the decomposition of variables into invariant modes of variation, which are considered to be fundamental elements in explaining the relationships among the variables. The extent to which PCs really represent the true modes of variation, however, depends largely on the extent to which the same modes can be found under different conditions.

Sampling errors

Unrotated solutions may also suffer from large sampling errors if neighbouring eigenvalues are close together (Kendall, 1980; North *et al.*, 1982; Storch and Hannoschöck, 1985). In these cases the information from population patterns can become mixed up when samples are drawn with eigenvalues separated by less than a specific value (Hsuing and Newell, 1983). Cahalan (1983) lists some well-referenced papers whose eigenvalue series' lie in the white noise spectrum and, consequently, the EOF patterns *may* be mixed up. To further study the effects of eigenvalue separation, examination of how unrotated and rotated PCs responded to nearly equal eigenvalues (equal out to 15 decimal places)

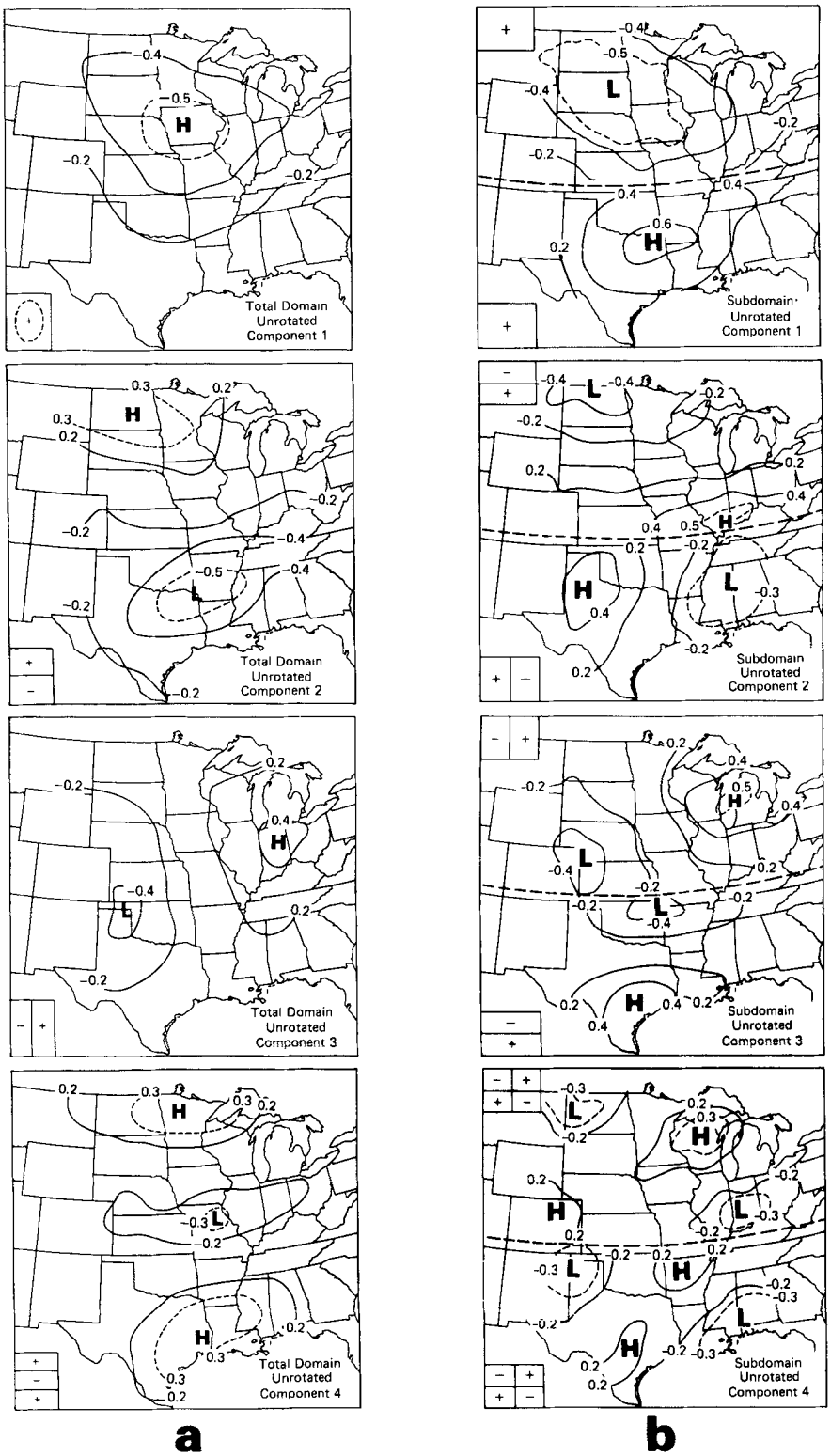

Figure 2. The first four unrotated modes of variation of 3-day totals for May–August precipitation over the central United States with corresponding Buell (1975) patterns shown in left inset. The results for the full domain (a) are to the left and those computed separately for the northern and southern halves of the domain (b) are to the right. The zonal division in (b) is given by the thick broken line. (Taken from Richman and Lamb, 1985)

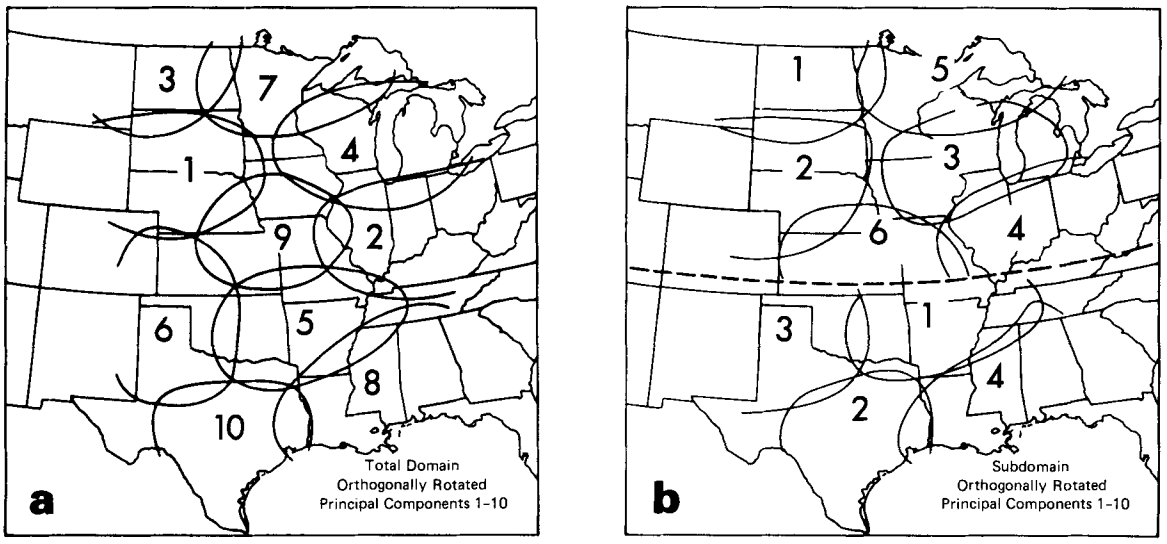

Figure 3. Regionalization of the central United States for 3-day summer rainfall totals for orthogonally rotated (Varimax criterion) principal components. The regional boundaries are the 0·4 loading isopleths for individual PCs. The results for the full domain appear in (a) and those computed separately for the northern and southern halves of the domain are shown in (b). The zonal division is given by the thick broken line. (Taken from Richman and Lamb, 1985)

was performed as a means to investigate their sensitivity to insufficient eigenvalue spacing. In this experiment the sample sizes ranged from 10 to 10,000. The samples for 100 replications of the unrotated and rotated solutions were compared to the appropriate unrotated and rotated population values by using a congruence coefficient (Richman and Lamb, 1985) to match them. The congruence coefficient varies from -1 to $+1$, identical to a correlation coefficient, but does not remove the mean of a variable. Values $\geq |0{\cdot}98|$ are interpretable as an excellent match (see section 9.4 for more information). The results of the study are shown in the upper portion of Table II, indicating that unrotated sampling errors are very severe and relatively insensitive to sample size. Note that increasing the sample size to be as large as 10,000 makes little improvement in the matches. Conversely, when the samples were rotated (DAPPFR criterion) the correct population structure could be recovered for any sample size greater than 25. The *population* loadings of both the unrotated and obliquely rotated solutions are shown at the bottom of Table II (left) and the first replication (of 100) of the *sample* loadings for both solutions are shown for the case of a moderate sample size of 500 (right). Comparisons of the unrotated population loadings with those drawn from the sample indicate little correspondence, which is reflected by the congruence coefficient which had an average root mean square of 0·39847 for the 100 replications at a sample size of 500. Inspection of the obliquely rotated pattern coefficients ('loadings') for the same replication from the population to the sample indicates an excellent match, reflected by the average root mean square congruence coefficient of 0·99902 for the 100 replications. This agrees well with Horel's (1984) preliminary findings of lower sampling errors for rotated patterns compared to unrotated PC loadings and the results in section 10 for data containing strong to moderate 'simple structure'. Cliff and Hamburger (1967, pp. 443–444) in their study of sampling errors in factor analysis note that 'under certain conditions [e.g. closely spaced eigenvalues], sampling errors of *unrotated* loadings are likely to be so large as to render them useless as estimates of anything'. They further conclude that an analyst 'should rotate his factors [PCs] either to a hypothesis or by an analytic procedure if he wants statistical stability'.

Faithfulness to relationships embedded in the dispersion matrix

A final reason why an alternative solution may be desirable is that the rotated solutions sometimes yield results which are more intuitively meaningful in a meteorological sense. The example in Figure 1

Table II. Coefficients of congruence for 100 replications of population versus sample component loadings for cases with nearly equal eigenvalues. An example is shown (bottom) for the correspondence between population values (left) and sample values based on a sample size of 500 (right)

Unrotated solution		Obliquely rotated solution	
Sample size	Coefficient of congruence between sample and population	Sample size	Coefficient of congruence between sample and population
10	0·37138	10	0·92140
25	0·38775	25	0·97974
50	0·38823	50	0·99008
100	0·39521	100	0·99508
250	0·39760	250	0·99804
500	0·39847	500	0·99902
1000	0·39897	1000	0·99951
2500	0·39935	2500	0·99980
5000	0·39951	5000	0·99990
10,000	0·39951	10,000	0·99995

Unrotated and obliquely rotated population values					Unrotated and obliquely rotated sample values				
	Unrotated solution		Obliquely rotated solution			Unrotated solution		Obliquely rotated solution	
Variable	PC 1	PC 2	PC 1	PC 2	Variable	PC 1	PC 2	PC 1	PC 2
1	0·7517	0·6596	0·9999	0·0003	1	0·5929	0·8010	0·9692	0·0003
2	0·7517	0·6596	0·9999	0·0003	2	0·5929	0·8010	0·9692	0·0003
3	0·7517	0·6596	0·9999	0·0003	3	0·5929	0·8010	0·9692	0·0003
4	0·7517	0·6596	0·9999	0·0003	4	0·5929	0·8010	0·9692	0·0003
5	−0·7517	−0·6596	−0·9999	−0·0003	5	−0·5929	−0·8010	−0·9692	−0·0003
6	−0·7517	−0·6596	−0·9999	−0·0003	6	−0·5929	−0·8010	−0·9692	−0·0003
7	0·7517	−0·6596	0·0003	1·0000	7	0·8940	−0·4673	0·0003	0·9521
8	0·7517	−0·6596	0·0003	1·0000	8	0·8940	−0·4673	0·0003	0·9521
9	0·7517	−0·6596	0·0003	1·0000	9	0·8940	−0·4673	0·0003	0·9521
10	−0·7517	0·6596	−0·0003	−1·0000	10	−0·8904	0·4673	−0·0003	−0·9521
11	−0·7517	0·6596	−0·0003	−1·0000	11	−0·8904	0·4673	−0·0003	−0·9521
12	−0·7517	0·6596	−0·0003	−1·0000	12	−0·8904	0·4673	−0·0003	−0·9521
13	0·6000	−0·8000	−0·2071	1·0056	13	0·7761	−0·6360	−0·2007	0·9574
14	0·6000	0·8000	1·0054	−0·2071	14	0·4108	0·9024	0·9746	−0·1972
15	0·5000	−0·8660	−0·3236	0·9891	15	0·6923	−0·7217	−0·3137	0·9417
16	0·5000	−0·8660	0·9889	−0·3237	16	0·2968	0·9437	0·9586	−0·3182

indirectly addresses the idea; however, it is more an example of domain shape dependence and not directly transferable to most analyses since it was based on *known* patterns. A more exact example of this concept with a methodology useful for any exploratory work is now presented using the unrotated and rotated PC patterns of 3-day summer rainfall totals given by Richman and Lamb (1985). This may be the most important advantage of a rotated solution, yet it is often a difficult one to substantiate owing to the exploratory nature of most EOF/PC studies. The first ten *unrotated* PCs are shown in Figure 4(a) and their interpretation is possible, yet difficult, because the patterns are not what one would expect for convective rainfall. For example, PC1 would correspond to either an excess or deficit of 3-day rainfall over the *entire* domain simultaneously. Inspection of the actual rainfall data did not support this pattern, as there were *no* observations where it rained over every station simultaneously or where there was no rain at every station. In order to further verify or reject this observation, a quantitative inspection of the correlation matrix was performed, similar to a procedure used by Wigley *et al.* (1984). The spatial correlations of the station with the highest PC loading on each PC to each of

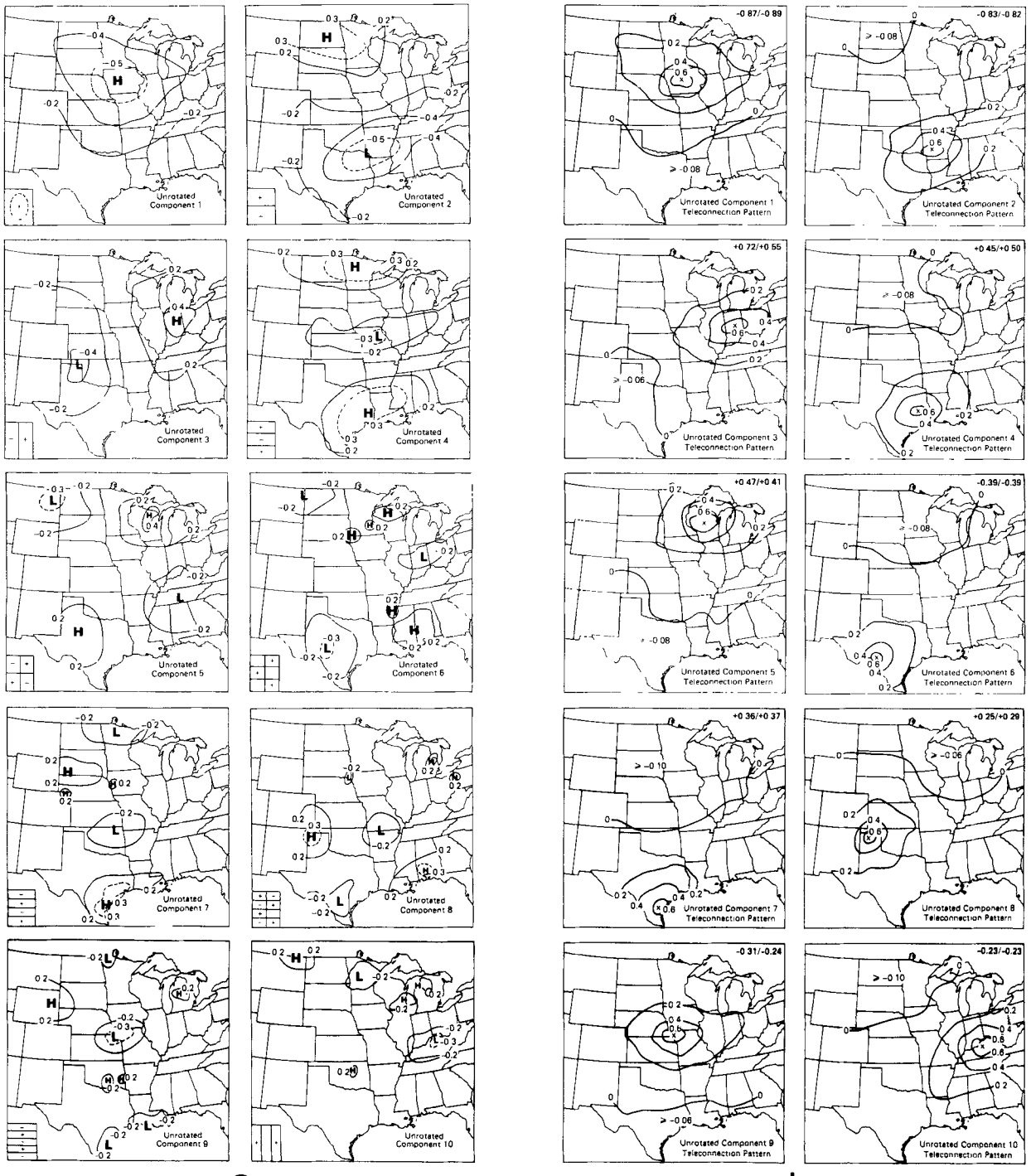

Figure 4. Unrotated principal components of 3-day summer precipitation totals over the central United States. The first 10 components appear in (a). The point correlation teleconnection patterns corresponding to the unrotated PCs appear in (b). The cross in (b) locates station with the highest PC loading in (a). Correlation coefficients between the crossed station in (b) and the remaining 401 stations are isoplethed; those with magnitudes less than approximately 0·08 are not significant at the 99 per cent level according to the Z-test. Numbers in the top right-hand corner of (b) give the pattern correlation/congruence coefficient (see section 9.4) between the correlation field shown and the corresponding PC loading pattern (Taken from Richman and Lamb, 1985)

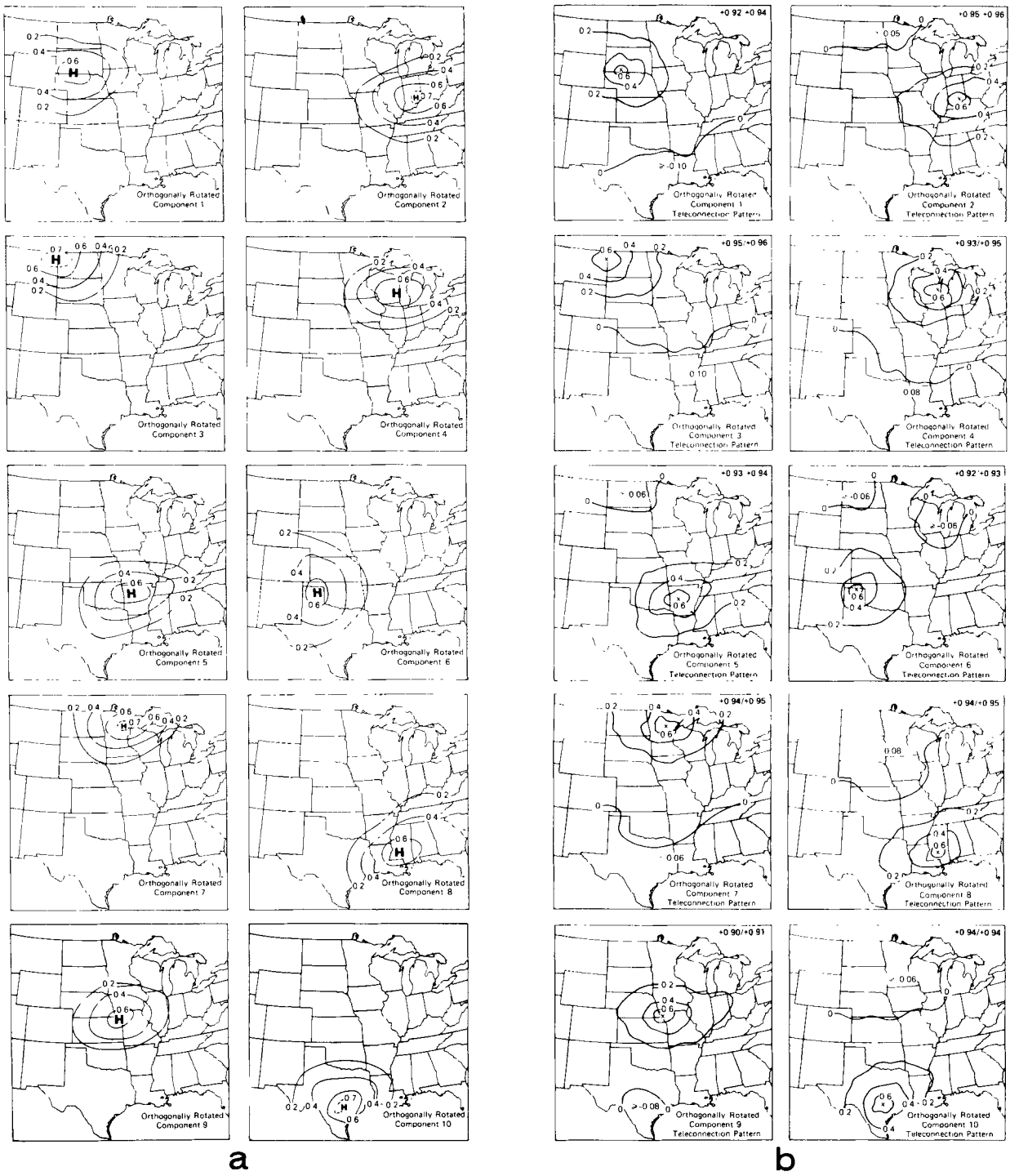

Figure 5. Orthogonally rotated (Varimax criterion) PCs (a) and corresponding point correlation teleconnection patterns (b). Explanation is the same as Figure 4. (From Richman and Lamb, 1985)

the other 401 stations were calculated to provide basic information on the scale and orientation of the anomaly features and compared to the unrotated PCs to assess how well they recovered the underlying input correlation fields. This is shown directly opposite the PC in Figure 4(b), indicating some interesting features, as the first PC has a larger scale (recalling that the PC loadings are correlation coefficients) than the corresponding point-to-point station correlation teleconnection field. Proceeding to PC2, there is a flip-flop evident between the south and north. The physical explanation for this map would be an excess of rainfall in the northern half of the domain and a deficit in the south (or vice versa) yet the point correlations exhibited a surprising pattern in which there was only *one* statistically significant anomaly with no flip-flop. This is in direct disagreement with unrotated PC2 and points to the lack of a climatologically meaningful pattern in the unrotated PC. PC3–PC10 and their corresponding correlation patterns also point to this same defect in the unrotated PCs and result in poor matches. The first PC has a scale larger than the corresponding correlation field supports, whereas unrotated PCs5–10 have scales *smaller* than the corresponding correlation fields indicate. The mean correlation match between the teleconnection maps and the PC loadings (over all 10 maps) was only 0·47, with a standard deviation of 0·22 indicating a wide range of correspondence. The same process was repeated for the ten *rotated* PCs (Figure 5(a)). In rotated PC1, the point correlation field is shown directly opposite the PC in Figure 5(b) and indicates an excellent match between the PC and the correlation pattern. It is noteworthy that the shapes and orientations of the two fields, as well as the *scale* of the region, are very similar. The matches between PCs2–10 and their correlation fields are also excellent, as the sizes of the corresponding regions are almost identical, leading to correlation matches >0·90 in *every* case with a mean match of 0·93 (over all 10 map sets) and a standard deviation of only 0·01 indicating uniformly excellent correspondence of location, size and morphology. The importance of this is that the underlying correlation fields have been shown to be properly depicted by the rotated PCs. The example presented in Figures 4 and 5 is only one application of this procedure, as the methodology can be applied to other spatial parameters (e.g. 500 mb height by Richman and Walsh, 1985) or any type of analysis (e.g. grouping synoptic variables or chemical ions as in Thurston and Spengler (1985)) by identifying the variable with the largest loading on each PC.

As a result of these various potential advantages which may exist for rotated EOF/PCs when individual modes are examined, compared to the unrotated solutions, this work will concentrate on presenting the theory behind simple structure rotation, meteorological applications of these rotations and, finally, a comparison of the various solutions available.

3. PRINCIPAL COMPONENT MODEL

Basic equations

The principal component model will be developed for application to the rotated solutions. Most of the results are directly applicable to EOF and FA also. The PC model can be defined from an $N \times n$ data matrix $\mathbf{Z} = \{z_{ij}: i = 1, \ldots, N; j = 1, \ldots, n\}$ where i indexes individuals or cases and j indexes variables. For every, i, j, z_{ij} denotes the variable value of some physical field or some derived field. The PC model representation of z_{ij} is

$$z_{ij} = \sum_{m=1}^{r} f_{im} a_{mj}^{\mathrm{T}} \qquad i = 1, \ldots, N; j = 1, \ldots, n; m = 1, \ldots, r; r \leqslant n \tag{1}$$

where f_{im} is the rth principal component score for the ith individual and a_{jm} is the rth principal component loading on the jth variable.

In matrix form (1) becomes

$$\mathbf{Z} = \mathbf{F}\mathbf{A}^{\mathrm{T}} \tag{2}$$

The principal component scores (sometimes called 'amplitudes' in the meteorological literature) and loadings can be developed with the following definitions (assumed to be in standardized form and

divided by the number of cases although not a necessary condition):

$$\mathbf{R} \equiv \mathbf{Z}^{\mathrm{T}}\mathbf{Z}\ (\mathbf{n} \times \mathbf{n}) \text{ (data correlation matrix)} \tag{3}$$

$$\mathbf{\Phi} \equiv \mathbf{F}^{\mathrm{T}}\mathbf{F}\ (r \times r) \text{ (principal component score correlation matrix)} \tag{4}$$

$$\mathbf{S} \equiv \mathbf{Z}^{\mathrm{T}}\mathbf{F}\ (n \times r) \text{ (principal component primary structure matrix)} \tag{5}$$

Since $\mathbf{Z}^{\mathrm{T}} = \mathbf{A}\mathbf{F}^{\mathrm{T}}$ can be substituted in (5),

$$\mathbf{Z}^{\mathrm{T}}\mathbf{F} = \mathbf{A}\mathbf{F}^{\mathrm{T}}\mathbf{F} \tag{6}$$

$$\mathbf{S} = \mathbf{A}\mathbf{\Phi} \tag{7}$$

which is an alternative solution of the PC primary structure matrix (5) which represents the correlations between the PCs and the variables. The equations (3), (4) and (6) are used to define the PC model in terms of the correlation matrix ($\mathbf{R}$) representation: $\mathbf{R} = \mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{A}\mathbf{F}^{\mathrm{T}}\mathbf{F}\mathbf{A}^{\mathrm{T}}$

$$\mathbf{R} = \mathbf{A}\mathbf{\Phi}\mathbf{A}^{\mathrm{T}} \tag{8}$$

This relationship (8) is the fundamental PC equation in terms of the correlation matrix and is a combination of $\mathbf{A}$ (the *PC primary pattern matrix*) and $\mathbf{A\Phi}$ (the *PC primary structure matrix*). In the initial solution (and under orthogonal rotation) $\mathbf{\Phi} = \mathbf{I}_r$, therefore $\mathbf{A\Phi} = \mathbf{A}$ and both are referred to as *PC loadings*.

Transformational indeterminacy

In the introduction it was stated that an infinite number of alternative solutions satisfied the PC model. Applying the relationships which define the PC model in (2) and (8), it will be shown that an *inherent indeterminacy* exists, since a unique solution is not possible without additional constraints. This arises because PCs may be linearly transformed to an infinite number of other sets of PCs having the same properties by using *arbitrary non-singular linear transformations* ($\mathbf{T}$) such as

$$\mathbf{F}^* = \mathbf{F}\mathbf{T}^{\mathrm{T}} \tag{9}$$

$$\mathbf{A}^* = \mathbf{A}\mathbf{T}^{-1} \tag{10}$$

The data matrix representation in (2) then becomes

$$\mathbf{Z} = \mathbf{F}\mathbf{A}^{\mathrm{T}} = \mathbf{F}\mathbf{T}^{\mathrm{T}}(\mathbf{T}^{\mathrm{T}})^{-1}\mathbf{A}^{\mathrm{T}} = \mathbf{F}^*\mathbf{A}^{*\mathrm{T}} \tag{11}$$

For the correlation matrix representation in (8), $\mathbf{\Phi}^*$ can be defined as $\mathbf{\Phi}^* = \mathbf{T}\mathbf{\Phi}\mathbf{T}^{\mathrm{T}}$ and application of (10) yields

$$\mathbf{R} = \mathbf{A}\mathbf{\Phi}\mathbf{A}^{\mathrm{T}} = \mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{\Phi}\mathbf{T}^{\mathrm{T}}(\mathbf{T}^{\mathrm{T}})^{-1}\mathbf{A}^{\mathrm{T}} = \mathbf{A}^*\mathbf{\Phi}^*\mathbf{A}^{*\mathrm{T}} \tag{12}$$

From (11) it can be seen that $\mathbf{F}$ and $\mathbf{F}^*$ have the same properties and are indistinguishable, accounting for the same amount of total variance. From (12) it can be seen that $\mathbf{A}$ can be linearly transformed to $\mathbf{A}^*$ to yield a new set of PC loadings. To determine a unique set of PCs, a set of r basis vectors must be defined so that the PCs can be referred to these vectors. The basis vectors can be chosen to be orthogonal to each other, so that the PCs are uncorrelated in an *orthogonal rotation,* or can be chosen to be obliquely related, so that the PCs are correlated in an *oblique rotation.* In both orthogonal and oblique solutions, the basis vectors are typically chosen such that the PC loadings ($\mathbf{A}^*$) [orthogonal case] or the PC primary pattern matrix ($\mathbf{A}^*$) [oblique case] have specific variance properties. The matrix $\mathbf{A}^*$ is typically transformed with the constraint that it contains the properties of a concept known as 'simple structure'.

4. SIMPLE STRUCTURE

4.1. *General concept*

Thurstone (1947) developed a set of five requirements or criteria for a transformed solution, which he called simple structure, to help ensure that the matrix $\mathbf{A}^*$ is easily interpretable. The definition of these five requirements will be presented in Section 4.3. Some background information will now be offered to make the exact definition more meaningful and logical. The goal of simple structure is to aid in 'the scientific problem of discovering an underlying order in the domain'. This implies that the concept is useful in exploratory work. *Thurstone reasoned that the variable correlations* (**R**) *defined a geometrical configuration of variable vectors without a reference frame.* Since the PCs could be represented by the axes of the frame, it was necessary to locate a reference frame in the variable configuration where it could be used for meaningful scientific interpretation of the configuration.

Thurstone noted that, in the majority of analyses, variable vectors would cluster or group together. Furthermore, the angular separations between the variable vectors within a group were relatively small, whereas the separations between groups of clustered vectors were markedly larger. *When these subgroups or clusters were examined, the specific variable vectors within any one were frequently so highly related that one causal or summary influence could be assigned to each of the variables within a cluster.* Rotating the PCs into such a position where they explained these clusters (were co-linear with them) would facilitate easier interpretation since each PC would represent one summary influence (or a subset of influences) and would not be related to other summary influences in other data clusters. Thurstone's concept of simple structure was specifically tailored to common factor analysis (as opposed to PC), which only analyses the common or shared variance between variables. Thurstone states that all of the common variance of a variable should be accounted for by fewer than r common factors. Consequently, the definitions in this paper have been modified to agree with PC terminology. The important point is that the advantages gained by a simple structure rotation of common factors are completely transferable to a simple structure rotation of PCs.

4.2. *Technical aspects of simple structure*

The fundamental rationale behind simple structure is the compact scientific explanation of a variable with the *smallest* number of PCs. For example, if there were n variables and r PCs retained in an analysis, the solution would fulfil the principles of simple structures when each variable required *fewer* than r PCs to account for most of the variable's variance. Moreover, the individual PCs would have a compact interpretation if each PC was associated with only a subgroup of the original variables. Since a simple structure solution requires that much of the variables' variance must be accounted for by a linear combination of fewer than r PCs, this directly implies that to achieve simple structure, the variables' variance must exist in a subspace having no more than $r-1$ dimensions in r-dimensional PC space. *By definition, a subspace of* $r-1$ *dimensions is known as a hyperplane or co-ordinate hyperplane.* For example, a three-dimensional region is a hyperplane lying in four-space, analogous to a plane lying in three-space.

Under simple structure rotation, the problem is to find those subspaces which completely contain the significant variance of various subsets of the original variables. The device which is used to locate these subspaces is known as a *reference axis.* A reference axis (as opposed to a primary axis) is a hypothetical random variable of unit length in PC space which is orthogonal to a hyperplane in the r-dimensional PC space. Since a reference axis is orthogonal to every vector contained in the hyperplane, a vector in PC space which is orthogonal to an observed variable is potentially a reference axis and identifies a corresponding co-ordinate hyperplane to which it is orthogonal (Mulaik, 1972). Given variables for which simple structure applies, then, since there are n observed variables and, at most, r linearly independent reference axes in r-dimensional PC space, some of the observed variables must be orthogonal to the same reference axes, and the significant part of these variables' variances must share the same subspaces of the PC space. Mulaik (1972) suggests a way to find a simple structure solution

using the $n \times r$ *reference structure matrix* (**V**)

$$\mathbf{V} = \mathbf{R}_{ZY} = \mathbf{Z}^{T}\mathbf{Y} \tag{13}$$

where **Z** is a matrix of n row vectors $\mathbf{z}_1^T, \ldots, \mathbf{z}_n^T$ which represent the normalized variables $Z_1, \ldots, Z_n$ and **Y** is a matrix consisting of r row vectors $\mathbf{y}_1^T, \ldots, \mathbf{y}_r^T$ representing the r reference axes. The entries in **V** are correlation coefficients between the vectors $\mathbf{z}_1^T, \ldots, \mathbf{z}_n^T$ and the vectors $\mathbf{y}_1^T, \ldots, \mathbf{y}_r^T$.

4.3. *Explicit definition of simple structure*

From the definition of **V** in (13), a simple structure can be determined which satisfies Thurstone's (1947, p. 335) five criteria. Thurstone notes that variables of complexity r should be eliminated prior to a final rotation and interpretation since these variables do *not* contribute to the location of the reference frame. This point is frequently overlooked. Additionally, the five criteria specifically refer to an oblique solution, since it yields the cleanest interpretation as no artificial orthogonality constraint is applied, although the criteria are also valid for orthogonal solutions. All that these criteria require of the r reference axes is that they be a set of linearly independent vectors in PC space.

1. Each row of the [oblique] reference structure matrix **V** should have at least one zero. This criterion satisfies the basic concept (explained in section 4.1).
2. For each column m of the matrix **V** there should be a distinct set of at least r linearly independent observed variables whose correlations (as found in the mth column of **V**) with the mth reference axis variable are zero. This criterion is needed to overdetermine the reference axis, since each column m of **V** is determined by a hyperplane $\mathbf{H}_k$.
3. For every pair of columns in **V** there should be several variables, whose entries are zero in one column but not in the other. This criterion ensures both the overdetermination and the distinctness of the reference axes and their corresponding subspaces of $r - 1$ dimensions in PC space. [If two columns of **V** contained the same zero entries and if these exceeded $r - 2$ in number, then the two corresponding hyperplanes would be identical.]
4. With four or more PCs, each pair of columns of **V** should have a large proportion of corresponding zero entries. This criterion ensures that the reference axis is pertinent to only a few of the observed variables, thereby guaranteeing a separation of the variables into distinct clusters. [This criterion is unimportant when there are only two or three PCs retained as the complexity would be less than or equal to 1 since, at most, one PC would be measured.]
5. For every pair of columns in **V** there should preferably be only a small number of entries in both columns which do not vanish. This criterion further ensures the simplicity of the variables.

Thurstone notes that when these conditions are met a *graphical plot* of each pair of columns of **V** shows (i) a large concentration of points in two radial streaks, (ii) a large number of points at or near the origin, and (iii) only a small number of points off the two radial streaks. For a PC analysis of r dimensions, there would be $\frac{1}{2}r\,(r-1)$ graphical plot diagrams. Thurstone (1947, p. 335) goes so far as to note that 'In the last analysis it is the appearance of the diagrams that determines, more than any other criterion, which of the hyperplanes of the simple structure are convincing and whether the whole configuration is to be accepted as stable and ready for interpretation.'

A common point of confusion is that the reference vectors (**Y**) mentioned above are *not* the PCs (**F**). The reference axes are treated as if they were PCs in a simple structure solution since they form a basis in PC space. The PCs are determined after the hyperplanes have been determined by the reference axes. *The reference axes form a biorthogonal system of vectors with the PC vectors.* Mulaik (1972) presents a detailed mathematical comparison of the interrelationships of these two vectors. The two can be related by the following equation:

$$\mathbf{F} = \mathbf{YT} \tag{14}$$

where **T** is an $r \times r$ non-singular linear transformation matrix. Finding **T** such that **F** and **Y** have unit length involves letting the covariances of the reference axes (**Y**), $\mathbf{C}_{YY} = \mathbf{Y}^{T}\mathbf{Y}$ and $\mathbf{D}_Y^2 = [\text{diag}\,\mathbf{C}_{YY}^{-1}]$.

Then **T** transforms **Y** into **F** if

$$\mathbf{T}^{\mathrm{T}} = \mathbf{D}_{\mathrm{Y}}^{-1}\mathbf{C}_{\mathrm{YY}}^{-1} \tag{15}$$

The PCs have unit variances because their covariance matrix yields a correlation matrix. If $\mathbf{C}_{\mathrm{YY}}$ is not a diagonal matrix (that is, the reference axes are correlated), the PCs are also correlated, so $\mathbf{C}_{\mathrm{FF}}$ or $\mathbf{\Phi}$ is not diagonal.

Typically, simple structure is assessed in the PC loading matrix (for orthogonal solutions) or in the PC primary pattern matrix (oblique solutions). The relationships can be seen by substituting **Y** as defined in (14), **V** as defined in (13) and $\mathbf{T}^{\mathrm{T}}$ as defined in (15):

$$\begin{aligned} \mathbf{Z} &= \mathbf{F}\mathbf{A}^{*\mathrm{T}} = \mathbf{Y}\mathbf{T}\mathbf{A}^{*\mathrm{T}} \\ (\mathbf{Z}^{\mathrm{T}}\mathbf{Y}) &= (\mathbf{A}^{*}\mathbf{T}^{\mathrm{T}}\mathbf{Y}^{\mathrm{T}})\mathbf{Y} = \mathbf{A}^{*}\mathbf{T}^{\mathrm{T}}\mathbf{Y}^{\mathrm{T}}\mathbf{Y} \\ \mathbf{V} &= \mathbf{A}^{*}\mathbf{T}^{\mathrm{T}}\mathbf{C}_{\mathrm{YY}} = \mathbf{A}^{*}\mathbf{D}_{\mathrm{Y}}^{-1}\mathbf{C}_{\mathrm{YY}}^{-1}\mathbf{C}_{\mathrm{YY}} = \mathbf{A}^{*}\mathbf{D}_{\mathrm{Y}}^{-1} \\ \mathbf{A}^{*} &= \mathbf{V}\mathbf{D}_{\mathrm{Y}} \end{aligned} \tag{16}$$

As a result of (16), the zeros in $\mathbf{A}^{*}$ correspond to the zeros in **V** and the PC primary pattern matrix ($\mathbf{A}^{*}$) has coefficients which are proportional by columns to the matrix of reference structure coefficients. Therefore, in a definition of simple structure configurations, either the reference structure matrix or the primary pattern matrix can be used. This will allow identification of simple structure in the pairwise graphical plots directly from the PCs.

4.4. *Objective analytic rotation*

The simple structure criteria listed in section 4.3 have been used as a springboard to create analytic computer programs to *approximate* at least a portion of the concept. There is a difficulty in defining the five criteria objectively since they are topological concepts originally designed for hand-rotating the axes. The advantages gained with analytic rotation programs arise from the objectivity gained by automating an algebraic formula which replaces human judgement in placing the reference axes (hence is more easily reproduced independently) and the time saved by computers (as what may now take 30 seconds of computer time would take days or weeks to accomplish by hand). This last feature has been such a strong motivating force that no fewer than *nineteen* orthogonal and oblique rotations exist today; they are listed in Table III. Most of the analytic rotations are simple algebraic expressions which attempt to approximate simple structure through the application of specially designed mathematical algorithms which distribute the PC loadings such that the dispersion of the loadings is maximized by maximizing the number of large and small coefficients (e.g. Varimax). This, in turn, has led to the common misconception that simple structure involves maximizing the number of near-zero and large loadings (for example, Johnston (1981, p. 217) states '. . . theory of simple structure, which says that components/factors should be clearly identified by their loadings. In an orthogonal solution, all loadings should be either ±1·0 or 0·0; in an oblique solution, all variables should have one loading of ±1·0'). This is *not* what simple structure attempts to do since simple structure makes no assumptions about how large the salient coefficients are; maximization of the near-zero coefficients is the most important aspect. Nevertheless, the advantages gained by such analytic programs has resulted in their widespread use. The important point to remember is that all of the analytic rotation algorithms listed in Table III were created as attempts to fulfil Thurstone's simple structure principles, as simple structure itself was considered to be a highly desirable concept by the various authors. There is a potential danger in blindly accepting the output of such analytic programs as, occasionally, their accuracy in recovering known patterns is no better than randomly spinning the PCs (see section 10) or they fail to converge in exploratory analyses (Jennrich and Sampson, 1966). One method to detect most of the problems is to plot up the rotated results as outlined in section 5 below. Cattell (1966b, p. 188) sums up the danger in blindly accepting the analytic results without inspecting the pairwise plots of rotated loadings when he states 'The goal of a single "push button" rotation to final simple structure is naturally so seductive that many proceed as if it existed, but the resources at present claiming to do this are theoretically and practically misleading'.

Table III. Analytic simple structure rotations

Rotation name	Type	Reference(s)
Quartimax	Orthogonal	Carroll (1953); Neuhaus and Wrigley (1954)
Varimax	Orthogonal	Kaiser (1958)
Transvarimax	Orthogonal	Saunders (1962)
Parsimax	Orthogonal	Crawford (1967)
Equamax	Orthogonal	Kaiser (1974)
Quartimin	Oblique	Carroll (1953)
Biquartimin	Oblique	Carroll (1957)
Covarimin	Oblique	Kaiser (1958)
Binormamin	Oblique	Kaiser and Dickman (1959, 1977)
Maxplane	Oblique	Cattell and Muerle (1960)
Oblimax	Oblique	Saunders (1961)
Harris–Kaiser Class II	Oblique	Harris and Kaiser (1964)
Harris–Kaiser Class III	Oblique	Harris and Kaiser (1964)
Promax	Oblique	Hendrickson and White (1964)
Direct oblimin	Oblique	Jennrich and Sampson (1966)
Functionplane	Oblique	Katz and Rohlf (1974)
H–A (Types A and B)	Oblique	Hakstian and Abell (1974)
Casey's method	Oblique	Kaiser and Cerny (1978)
DAPPFR	Oblique	Tucker and Finkbeiner (1982)

5. GRAPHICAL PRESENTATION OF SIMPLE STRUCTURE

5.1. *Construction of pairwise plots*

The importance of the above statement naturally leads to the question: how can it be determined whether an analytic rotation is yielding results close to an ideal simple structure solution? Thurstone provides a clue when he states that the inspection of graphical pairwise plots of PCs is of utmost importance in the interpretation of an analysis (see Section 4.3). Such a plot involves setting two PCs as the x- and y-axes. The rotated PC loadings (or pattern coefficients) are plotted up with respect to the two PCs as a reference frame. An example is provided in Figure 6 which plots up the PC primary pattern coefficients in Table IV for the three sets of pairwise plots. In addition to this, Figure 6(a) has been constructed as a guideline to illustrate the hyperplanes which are orthogonal to the PCs. Variables which fall close to a hyperplane are termed 'in the hyperplane of that PC' (see below for definition of hyperplane width); variables located near the origin are termed 'in the hyperplanes of both PCs', and variables which have large coefficients on both PCs are termed 'complex variables' (not to be confused with $\sqrt{(-1)}$). Examining Figure 6(b), which plots PC1 vs. PC2, indicates that variables 2, 3, 6, 9, 12, 14, 15, 19 and 20 are nearly zero with respect to PC1. Strictly speaking, the hyperplane is infinitely narrow and can only be formed by an exact zero loading. The question arises of how the problem of sampling errors and other noise affects the magnitude of the loadings as it is unlikely that an *exact* zero loading will occur in any analysis. Since a subset of variables are 'close' to zero, Tucker (1955) suggested that variables considered to be in the hyperplane should have coefficients falling within a certain narrow

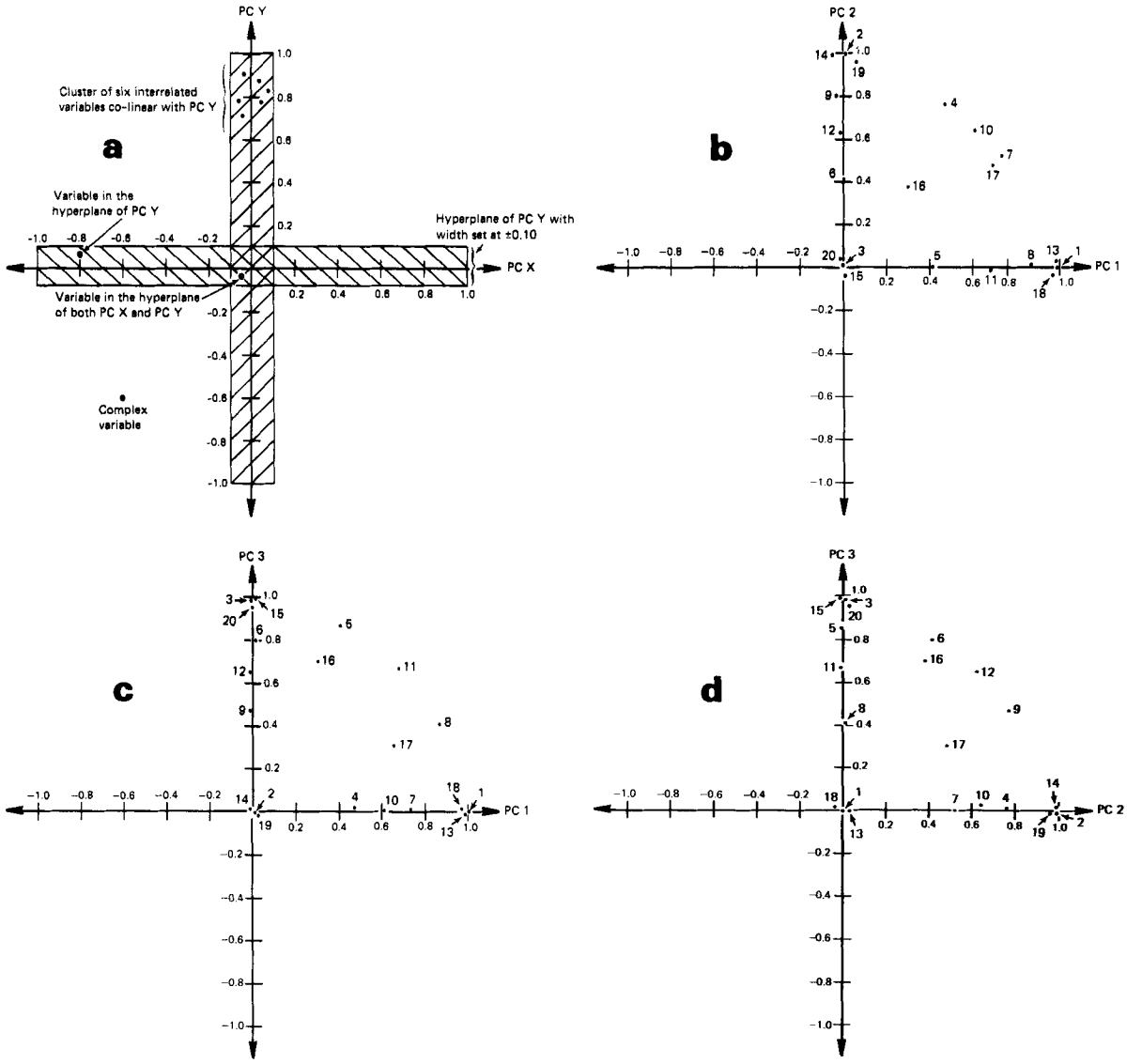

Figure 6. Graphical pairwise plots of the obliquely rotated primary pattern coefficients for the three PCs listed in Table IV and an illustrative guideline for interpretation (panel (a)). The plot of PC1 vs. PC2 is panel (b), PC1 vs. PC3 is panel (c) and PC2 vs. PC3 is panel (d). Note the clustering of subsets of variables along the PC axes. The numbering of the points in panels (b), (c), (d) corresponds to variable numbers in Table IV. See Table V for results of this analysis.

band on either side of zero. Typically, values between ±0·05 and ±0·20 are used with the *hyperplane width* being directly proportional to the number of input observations and the amount of noise in the analysis. For this example a value of ±0·10 will be used, as shown in Figure 6(a). Returning to Figure 6(b), variables 2, 3, 6, 9, 12, 14, 15, 19 and 20 all define the hyperplane of PC1. Similarly, variables 1, 3, 5, 8, 11, 13, 15, 18 and 20 define the hyperplane of PC2, whereas variables 1, 2, 4, 7, 10, 13, 14, 18, and 19 define the hyperplane of PC3 (Figure 6(c)). PC1 clusters variables 1, 5, 8, 11, 13 and 18 whereas PC2 clusters variables 2, 6, 9, 12, 14 and 19 (Figure 6(b)). Similar interpretations can be made by examining Figures 6(c) and (d).

This solution was obtained with an analytic oblique rotation. Does it fulfil the principles of simple structure? Examining the plots in Figure 6 indicates that there are near-zero loadings around the origin with two radial streaks and not too many complex variables as Thurstone mentions. If the five criteria

Table IV. Hypothetical PC loadings for graphical plot example shown in Figure 6 (Harman, 1976, p. 359)

Variable	PC1	PC2	PC3
1	0·997	0·000	−0·002
2	0·008	0·994	−0·006
3	−0·001	0·013	0·985
4	0·471	0·759	0·007
5	0·406	−0·008	0·863
6	0·008	0·411	0·800
7	0·728	0·522	0·001
8	0·863	0·014	0·411
9	−0·013	0·773	0·474
10	0·615	0·643	0·009
11	0·678	−0·001	0·672
12	−0·007	0·624	0·645
13	0·990	0·030	−0·011
14	−0·047	0·993	0·050
15	0·019	−0·046	0·996
16	0·293	0·378	0·697
17	0·645	0·472	0·301
18	0·970	−0·036	0·009
19	0·055	illegible	−0·055
20	−0·009	0·038	0·950

are examined individually, the simple structure can also be assessed and this rotation can be satisfactorily interpreted. This example is an actual analysis of physical data in which 20 non-linear measurements (variables) were taken of a series of boxes, as shown in Table V, based on Thurstone's (1947, pp. 140–144) experiment. These measurements were formed into a correlation matrix and analysed, and the final rotated solution identifies the three PCs as length, width and height of the boxes (X, Y, Z), which indicates that the analysis can recover the underlying order in a set of variables.

Table V. Various measurements taken for Thurstone's (1947) 20 variable box problem

Variable	Measurement	Variable	Measurement
1	X^2	11	$2X + 2Z$
2	Y^2	12	$2Y + 2Z$
3	Z^2	13	$\log X$
4	XY	14	$\log Y$
5	XZ	15	$\log Z$
6	YZ	16	XYZ
7	$\sqrt{(X^2 + Y^2)}$	17	$\sqrt{(X^2 + Y^2 + Z^2)}$
8	$\sqrt{(X^2 + Z^2)}$	18	e^X
9	$\sqrt{(Y^2 + Z^2)}$	19	e^Y
10	$2X + 2Y$	20	e^Z

5.2. *Amount of simple structure present in an analysis*

Using the graphical plots presented in Figure 6 as a guide, the amount of simple structure present in an analysis will now be schematically illustrated in Figure 7. This is vital to the development of the Monte Carlo technique in sections 9 and 10. *Strong simple structure* is present in Figure 7(a), as there are many variables in the hyperplane with distinct clusters of variables radiating out along the axes and few complex variables. Figure 7(b) illustrates *moderate simple structure* in which there are more complex variables than the previous case, and slightly fewer variables in the hyperplane. Figure 7(c) depicts *weak simple structure* in which there are a good number of complex variables, poorly defined

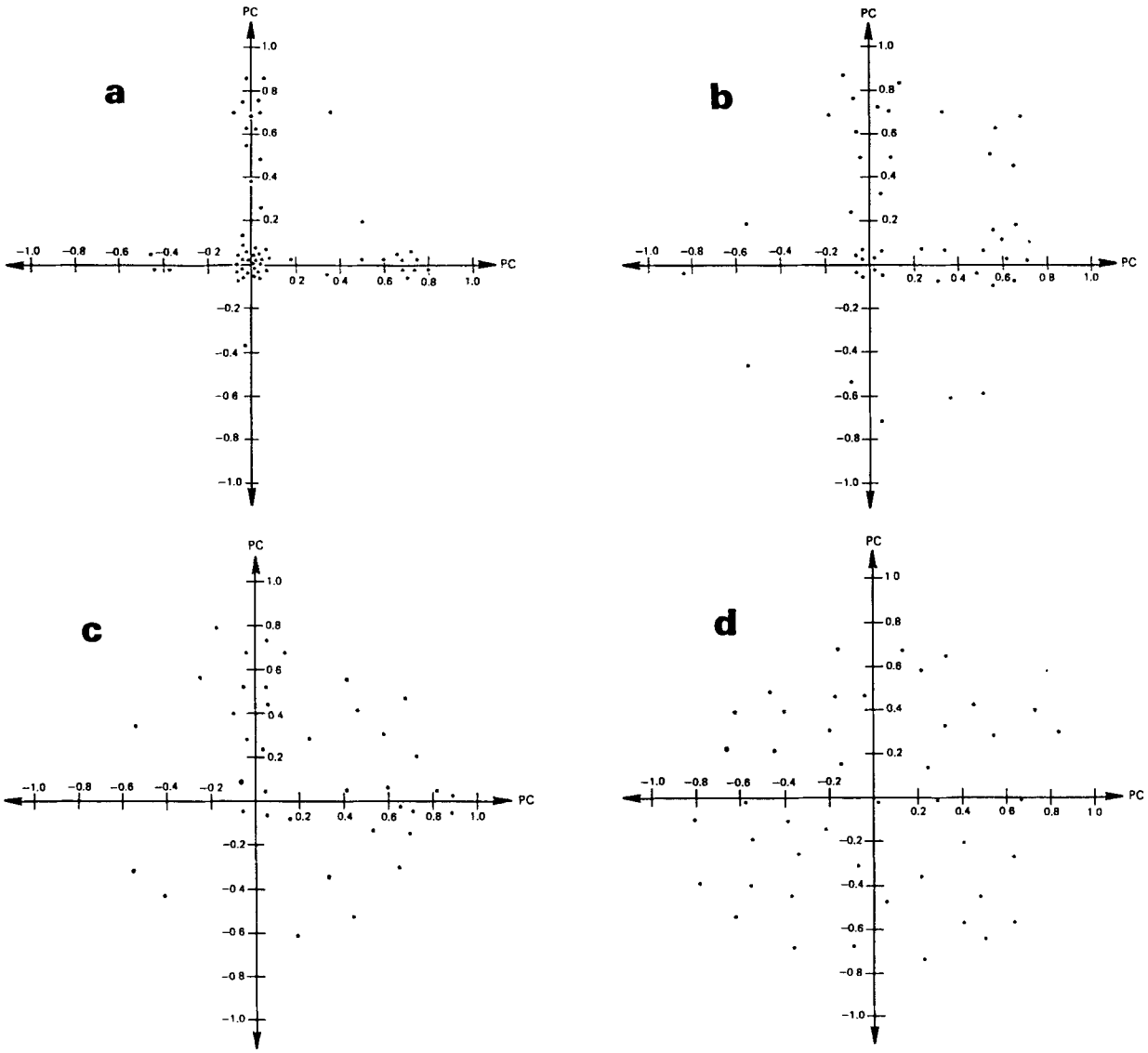

Figure 7. Hypothetical schematic examples of various degrees of simple structure using the graphical pairwise plots of rotated PC loadings as a guide. Panel (a) corresponds to a strong simple structure. Panel (b) corresponds to a moderate simple structure. Panel (c) corresponds to a weak simple structure. Panel (d) corresponds to a random configuration of input correlations. These plots can be compared to actual pairwise graphical plots of meteorological data to assess the degree of simple structure in any rotated data set

hyperplanes and indistinct clusters. Finally, Figure 7(d) depicts a *random configuration* with many *complex* variables and no clusters, in which case any PC position would be equally valid to describe the variables. Fortunately, meteorological data are rarely in a random configuration; those that are should not be subject to an eigenanalysis (particularly EOF/PC), as it is possible for the coefficients to appear as signal when they are artifacts of the techniques (see Richman, 1985, for illustrations and screening devices).

5.3. *Relating real data to the amount of simple structure present*

Using real examples of meteorological data, the amount of simple structure present in pairwise plots of PCs can be determined. The first example, Figure 8(a), is taken from the Lamb and Richman

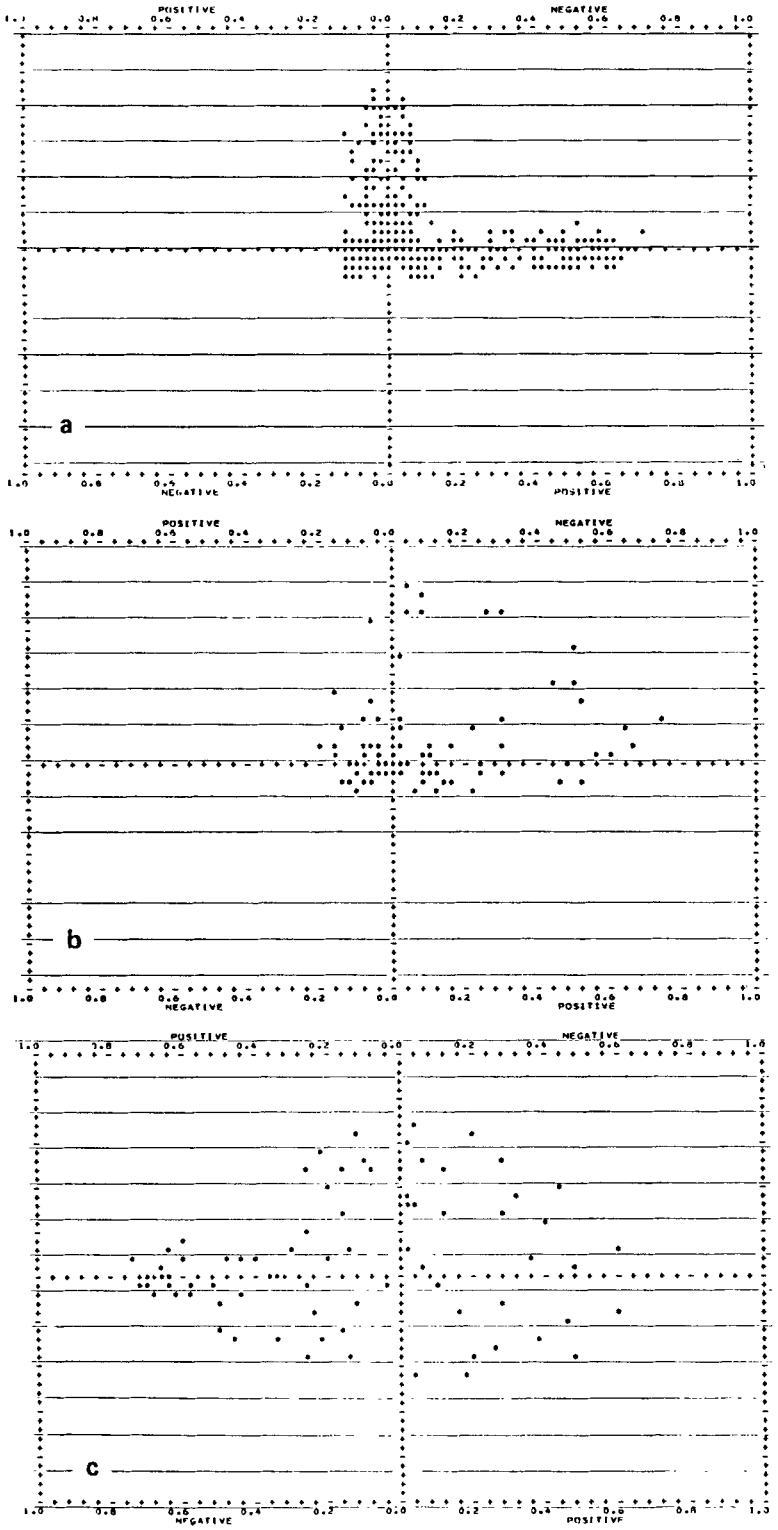

Figure 8. Computerized output of three representative graphical plots of meteorological/climatological data sets. Panel (a) is taken from Lamb and Richman's (1983a, b) 3-day summer rainfall study and illustrates a strong degree of simple structure with a Varimax rotation using the information in Figure 7 as a guideline. Panel (b) is taken from Walsh and Richman's (1981) sea-surface temperature data set and illustrates moderate simple structure with a Varimax rotation. Panel (c) is from Richman's (1981) geopotential height data set and illustrates a very weak simple structure with a Varimax rotation

(1983a) precipitation analysis, where PC loadings refer to climatological stations, exhibiting a large number of near-zero loadings in the hyperplane and clusters of salient loadings aligned with the PCs indicating a prime example of strong simple structure. The second example, Figure 8(b), is taken from the Walsh and Richman (1981) sea-surface temperature data set which exhibits moderate simple structure, as there are a number of complex variables between the PCs and clusters of variables which are not as well defined as in Figure 8(a). The third example, Figure 8(c), is taken from a study by Richman (1981) of 500 mb geopotential data (not shown in that article) in which there are more ambiguous clusters of variables with fewer variables in the hyperplanes corresponding to a weak simple structure. One point should be made clearly: the results in Figure 8 do *not* imply that all precipitation data exhibit strong simple structure, all sea-surface temperature data exhibit moderate simple structure or all geopotential data exhibit weak simple structure. In any given analysis some pairwise plots may exhibit strong simple structure, whereas others may exhibit moderate simple structure; therefore, *all* possible $\frac{1}{2}r\ (r-1)$ pairs of plots should be examined in order to assess the overall degree of simple structure in the data (as illustrated by Richman and Lamb, 1985).

Certain types of meteorological data are more amenable to simple structure rotation than others. Recalling that simple structure identifies *subsets* of variables is one important factor in deciding if a rotation will aid in the specific interpretation of an analysis. *This requires some meteorological insight prior to the analysis.* For example, the Richman and Lamb (1985) precipitation study was expected to yield more intuitively meaningful results when rotated, since warm season rainfall anomaly patterns tend to occur on a subset (or subregions) of the United States at any given 3-day time interval [(Figure 8(a)) attests to the strong simple structure and distinct clustering]. Conversely, the geopotential height data indicated little simple structure when rotated (Figure 8(c)) as many of the pairwise plots had almost randomly located variable points. This one example does *not* mean that geopotential data are not suitable for a simple structure rotation, as Cohen (1983), Barnston and Livezey (1985) and Horel (1981) present geopotential analyses where rotation was useful, only that the orthogonal rotation used (Varimax) failed to capture a simple structure for this one particular data set. At this point an oblique rotation could be applied to discern if the artificial orthogonality constraint was causing the problem. If the plots were similar to Figure 8(c), this one particular data set might, *in its present form,* have to be considered unsuitable for a simple structure interpretation. This leaves two alternatives (besides analysing the unrotated results): (i) attempt to analyse the transposed data matrix with a simple structure solution, or (ii) use a non-simple-structure rotation known as Procrustes target rotation. Each of these possibilities will be briefly examined.

6. THE SIX MODES OF DECOMPOSITION

A PC analysis can be specified in at least six basic operational modes, depending on which parameters are chosen as variables, individuals and fixed entities. These six modes have been defined as O, P, Q,

Table VI. The six modes of decomposition and how these relate to the PC model as shown in equation (1). Matrix configurations shown in Figure 9

PC mode	Variable index j denotes	Individual index i denotes	Fixed entity
O	time	field	station
P	field	time	station
Q	station	field	time
R	field	station	time
S	station	time	field
T	time	station	field
Matrix display	$\{a_{jm}, j=1,\ldots,n;\ m=1,\ldots,r\}$ PC loadings	$\{f_{im}, i=1,\ldots,N;\ m=1,\ldots,r\}$ PC scores	

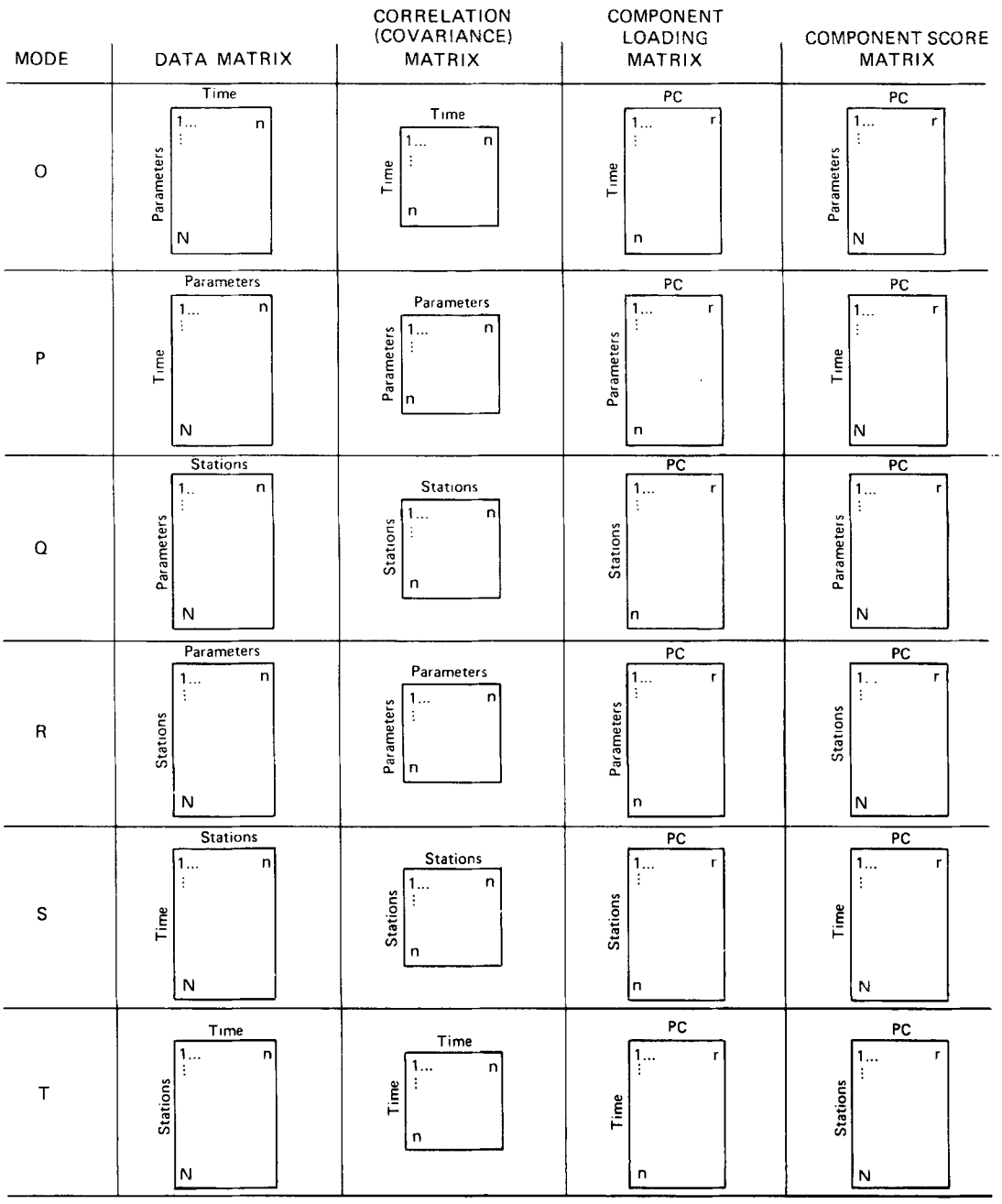

Figure 9. Data matrix (with columns treated as variables), dispersion matrix, PC loading matrix and PC score matrix under each of the six modes of variation. The application of a simple structure rotation to each of the six modes will uniquely cluster and simplify the variables (rows) in the PC loading matrix, corresponding to the information in Table VI

R, S and T by Cattell (1952) and will result in a *unique* clustering of variables when simple structure rotation is applied. In studies of meteorological fields in space and time, there are three entities: a meteorological field (or parameter), time and a location (or station). A PC analysis of a meteorological field can therefore be made by varying any two of these three entities and holding the third fixed. Examination of (1) reveals two basic matrices which are examined in PC. These are suggested by a_{jm} (the PC loadings) and f_{im} (the PC scores). These are applied to the six modes in Table VI. For example,

suppose **Z** has been rendered into an *S-mode* representation. From Table VI note that j indexes stations and i indexes times, of some fixed field. Then a geographical plot of a_{jm}, for fixed m, as a function of j would be a map of the PC loadings as distributed over the spatial extent of the data of **Z**. The time plot of f_{im}, for fixed m, as a function of i, would be a time series of the rth PC score. Alternatively, if *T-mode* had been used to analyse **Z**, the PC loadings, a_{jm}, would be a time series for each m whereas f_{im} would plot up as a geographical map of PC scores for each m. The matrix configurations of the data (columns are treated as variables), dispersion matrices, PC loading matrices (rows are treated as variables) and PC score matrices corresponding to all six modes in Table VI are shown in Figure 9.

Applying this information to meteorological data provides an important *insight* into the type of information gained via each mode of decomposition. In the typical S-mode decomposition, the simple structure rotation attempts to isolate subgroups of stations which covary similarly. Horel (1981) and Richman and Lamb (1985) provide this type of analysis, which is useful for regionalization. By *transposing* the data matrix, **Z**, so that the variables refer to the individual time observations, a T-mode analysis can be performed, and when rotated it would isolate subgroups of observations with similar spatial patterns and, thereby, simplify the time series. It is important that the mode of decomposition be given serious thought prior to the analysis in order to optimize the potential for rotation to uncover physically meaningful relationships. For example, in Richman (1983a) the data were 700 mb heights and the aim of the study was to develop a catalogue of flow types which resembled weekly 700 mb anomaly averages over the western half of the Northern Hemisphere. *Examination of the input data prior to the analysis was vital in this case as it gave insight into the nature of the height anomalies over the domain.* Inspection of the weekly 700 mb height anomalies indicated that 3–6 discrete anomaly centres were present on 98 per cent of the weekly mean maps examined. If an S-mode analysis with rotation was applied, the rotation would attempt to isolate as few anomalies as possible per component; Cohen (1983, p. 1981, his Figure 2) illustrates an example of this effect with the Varimax rotation depicting only one anomaly per PC map. This is not consistent with what was observed in the 700 mb maps, hence S-mode was rejected, *a priori*, and a T-mode analysis with rotation was considered. In T-mode the time series would be simplified into discrete groups of clustered weeks, which better fits the idea of discrete map types, and the spatial series (component scores) are left in a complex form which can then represent the complex waves evident at 700 mb. This is exactly what occurred when the T-mode rotated analysis was applied to the data set and yielded height anomaly patterns with typically 3–6 anomaly centres, in agreement with the input data. The pairwise graphical plots were examined and indicated a strong to moderate degree of simple structure, since the time series clustered into groups of weekly observations with similar spatial patterns. This example combined with Table VI and Figure 9 underlines the need to carefully consider which of the six modes of decomposition should yield the best theoretical representation of the data for the objective at hand *prior* to running an analysis.

7. PROCRUSTES TARGET ROTATION

Suppose that simple structure in any of the six modes is not theoretically suitable for a field under investigation. An alternative technique, Procrustes target rotation, may be useful (Hurley and Cattell, 1962). The various rotations listed in Table III are all simple structure transformations, which are most useful when trying to identify relationships, yet have limited use in trying to confirm existing theory. Procrustes target rotation is useful in confirmatory applications where a hypothesized component structure is of interest and involves designating the hypothesis as a target and rotating the unrotated PCs to an optimized fit with these loadings. The transformation involves a least-squares solution between transformed PC weights or loadings and the hypothesized PC weights. The goodness-of-fit between these two sets of weights is related to the validity of the hypothesis. Mathematically, the Procrustes equation is

$$\mathbf{B} = \mathbf{AT} + \mathbf{E} \tag{17}$$

where **B** is the $n \times r$ target matrix, **A** is the $n \times r$ PC unrotated loading matrix, **T** is the $r \times r$ transformation matrix and **E** in the $n \times r$ matrix of discrepancies or residuals. We wish to find a transformation matrix (**T**) such that

$$\text{trace } (\mathbf{E}^{\mathrm{T}}\mathbf{E}) = \text{trace } (\mathbf{B} - \mathbf{AT})^{\mathrm{T}}(\mathbf{B} - \mathbf{AT}) \text{ is a minimum} \tag{18}$$

Mulaik (1972) provides such a solution by taking the partial derivative with respect to **T** and setting (18) to zero as $\partial/\partial\mathbf{T}\ [\text{trace } (\mathbf{E}^{\mathrm{T}}\mathbf{E})] = -2\mathbf{A}^{\mathrm{T}}\mathbf{B} + 2\mathbf{A}^{\mathrm{T}}\mathbf{AT} = \mathbf{0}$. This can be rewritten as $\mathbf{A}^{\mathrm{T}}\mathbf{AT} = \mathbf{A}^{\mathrm{T}}\mathbf{B}$ and solving for **T** yields

$$\mathbf{T} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{B} \tag{19}$$

T is typically normalized such that diag $(\mathbf{T}^{\mathrm{T}}\mathbf{T}) = \mathbf{I}_r$. The matrix (**AT**) is analysed and comparison to the residuals (**E**) is of utmost importance when using Procrustes to gauge the validity of the hypothesis.

Procrustes offers numerous applications, such as in weather modification research, where the expected modification region could be specified as the 'target area' to note how well stations within this region clustered. Another application would be to verify if some hypothesized synoptic pattern from one data set existed in a second database, or at a different time interval (Richman, 1983a). In T-mode, Procrustes can be employed to isolate specific time periods in the target and used to study the effects of quasi-periodic events (such as El Niño) on meteorological parameters (such as surface temperature). This approach has advantages over harmonics since *aperiodic intervals* can be specified as the target.

8. ANALYTIC SIMPLE STRUCTURE ROTATION

8.1 Availability

In section 4, the concept of simple structure was introduced along with a method of assessing the amount of simple structure present in a database (section 5). The focus of this portion of this work involves comparing the analytic rotation techniques available on the statistical packages. Thurstone's (1947) five simple structure criteria have been partially incorporated into algebraic expressions which attempt to approximate a simple structure solution of which 19 were identified in Table III. A survey of the major mathematical/statistical packages (BMDP, 1981; IMSL, 1979; SAS, 1982; SPSS, 1979) revealed that, of the 19, only a subset of eight orthogonal and oblique rotations are readily obtainable (Table VII) because some are outdated whereas others are so new that they are not yet available. Of these eight, all five oblique rotations and the most accurate orthogonal rotation will be examined and compared for accuracy over a variety of different data sets. Algebraically, the orthogonal rotations

Table VII. Analytical simple structure rotations available on the four major mathematical/statistical packages. Asterisk denotes Procrustes which is generally used as a confirmatory transformation

Packages			
BMDP (1981)	IMSL (1979)	SAS (1982)	SPSS (1979)
Orthogonal rotations			
Equimax	Equimax	Equimax	Equimax
Quartimax	Quartimax	Quartimax	Quartimax
Varimax	Varimax	Varimax	Varimax
Oblique rotations			
Direct oblimin	Harris–Kaiser II IC	Harris–Kaiser II IC	Direct oblimin
Harris–Kaiser II IC	Harris–Kaiser II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$	Harris–Kaiser II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$	
Harris–Kaiser II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$	Procrustes*	Procrustes*	
	Promax	Promax	

must satisfy equation (10), $\mathbf{B} = \mathbf{A}^* = \mathbf{A}\mathbf{T}^{-1}$, where $\mathbf{B}$ is an $n \times r$ rotated PC loading matrix corresponding to $\mathbf{A}^*$ in (10) and (16), $\mathbf{A}$ is an $n \times r$ initial unrotated loading matrix and $\mathbf{T}$ is an $r \times r$ orthonormal transformation matrix which attempts to impart the properties of simple structure on $\mathbf{B}$. (Some of the packages define $\mathbf{T}^{-1}$ as $\mathbf{T}$; therefore, in these situations, $\mathbf{B} = \mathbf{AT}$). Since $\mathbf{T}$ is an orthonormal matrix, $\mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{T}\mathbf{T}^{\mathrm{T}} = \mathbf{I}_r$. The oblique rotations also use (10) but relax the condition that $\mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{T}\mathbf{T}^{\mathrm{T}} = \mathbf{I}_r$ so that $\mathbf{\Phi}$ may have off-diagonal values. One other characteristic of rotated solutions is that they typically resolve nearly the same total variance as unrotated solutions (see Richman and Lamb (1985) for empirical evidence, where variance retained for the Varimax solution was 99·9995 per cent of the unrotated solution). Harman (1976, p. 268) presents a means to calculate variances of oblique solutions.

8.2. *Varimax orthogonal rotation (Kaiser, 1958, 1959)*

Varimax will be the only orthogonal package rotation examined since it is generally accepted as the most accurate analytic algrbraic orthogonal rotation when applied to 'known' data sets (Rummel, 1970; Mulaik, 1972; Harman, 1976) and is the only orthogonal rotation which has been widely used. Varimax attempts to simplify the columns or PCs of the matrix $\mathbf{B}$ as a route to achieving simple structure. The simplicity (V^*) of a PC r is defined in terms of the variance of the squared loadings by

$$V^* = \left[n \sum_{i=1}^{n} (b_{ij}^2)^2 - \left(\sum_{i=1}^{n} b_{ij}^2 \right)^2 \right] \Big/ n^2 \qquad j = 1, \ldots, r \tag{20}$$

where n is the number of variables, r is the number of PCs, and the bs are the PC loadings. Kaiser reasoned that when the variance was at a maximum, the PC had its greatest simplicity (i.e. interpretability) as the b's tended to either 0 or 1. This is in some contrast to what simple structure states, as only near-zero loadings are important to a simple structure, yet is frequently a reasonable approximation. From (20), the normalized Varimax criterion (version available on packages) can be developed as

$$V = \sum_{j=1}^{r} \frac{n \sum_{i=1}^{n} (b_{ij}^2/h_i^2)^2 - [\sum_{i=1}^{n} (b_{ij}^2/h_i^2)]^2}{n^2} \tag{21}$$

where h_i^2 is the communality of the ith variable (the amount of the ith variables variance accounted for by the r PCs retained). The PCs are rotated *in pairs* until the value V is maximized to a given convergence criterion. The V criterion in (21) is used to determine the matrix $\mathbf{B}^* = \mathbf{H}^{-1}\mathbf{A}\mathbf{T}_{\mathrm{V}}$, where $\mathbf{H}^2$ is an $n \times n$ diagonal matrix of communalities, $\mathbf{A}$ is a matrix of the $n \times r$ unrotated loadings and $\mathbf{T}_{\mathrm{V}}$ is an $r \times r$ orthonormal transformation matrix determined when V is maximized. This is then renormalized to the original metric of the variables by a premultiplication by a diagonal matrix of square roots of the communalities as $\mathbf{B} = \mathbf{H}\mathbf{B}^*$.

8.3. *Direct oblimin (Jennrich and Sampson, 1966)*

The Direct oblimin method of oblique PC rotation operates directly on the primary pattern matrix $\mathbf{B}$ rather than the reference structure matrix $\mathbf{V}$ as shown in (16). Simple structure is approximated by minimizing a function of the primary pattern matrix $\mathbf{B}$, $F(\mathbf{B})$ of the form

$$F(\mathbf{B}) = \sum_{j<q=1}^{r} \left(\sum_{i=1}^{n} b_{ij}^2 b_{iq}^2 - \frac{\delta}{n} \sum_{i=1}^{n} b_{ij}^2 \sum_{i=1}^{n} b_{iq}^2 \right) \tag{22}$$

where δ is a specification parameter which is typically set at zero (default value for the various packages). The problem then becomes finding a transformation matrix $\mathbf{T}$ which minimizes $F(\mathbf{A}\mathbf{T}^{-1})$ under the condition diag $(\mathbf{T}^{\mathrm{T}}\mathbf{T}) = \mathbf{I}_r$. Rotations are performed systematically using all possible pairs of PCs until $F(\mathbf{B})$ converges to a set value (typically 0·00001). The value δ can be changed to alter the obliquity of the PCs. The permissible range for δ is $-\infty$ to $+1$, although values less than -10 typically

result in nearly orthogonal solutions whereas values greater than approximately +0·5 can result in a 'collapsed solution' as the PCs become highly correlated when the function $F(\mathbf{B})$ goes highly negative. If the graphical plots of the coefficients of $\mathbf{B}$ are examined, δ can be altered over a more moderate range (e.g. −1 to +0·5 in 0·1 increment steps) to visually assess which solution yields the best structure. Karl and Koscielny (1982) illustrate a meteorological example where δ was set at +0·25 to achieve a solution.

8.4. Harris–Kaiser Case II (Harris and Kaiser, 1964)

This oblique rotational approach finds a $\mathbf{T}$ by using orthonormal and diagonal transformation matrices using the Eckart–Young (1936) theorem with an arbitrary transformation matrix $\mathbf{M}$ which may be expressed as the product

$$\mathbf{M} = \mathbf{T}_2\mathbf{D}_2\mathbf{T}_1\mathbf{D}_1 \tag{23}$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ are orthonormal $r \times r$ matrices and $\mathbf{D}_1$ and $\mathbf{D}_2$ are $r \times r$ diagonal matrices. The Eckart–Young theorem states that, for any real matrix $\mathbf{A}$, two orthogonal matrices $\mathbf{P}$ and $\mathbf{Q}$ can be found for which $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{Q}$ is a real diagonal matrix, $\mathbf{D}$, with no negative elements (Mulaik, 1972). Rewriting (23) as

$$\mathbf{M} = \mathbf{M}^*\mathbf{D}_1 \tag{24}$$

which is in the form $\mathbf{B} = \mathbf{A}\mathbf{T}$, and applying the Eckart–Young theorem, any $\mathbf{M}^*$ can be expressed as $\mathbf{M}^* = \mathbf{T}_2\mathbf{D}_2\mathbf{T}_1$ where the relationships $\mathbf{T}_2 = \mathbf{P}$, $\mathbf{T}_1 = \mathbf{Q}^{\mathrm{T}}$ and $\mathbf{D}_2 = \mathbf{D}$ are defined by the theorem. The matrix $\mathbf{D}_1$ in (23) and (24) only serves to make the rotated PCs unit length. Therefore, $\mathbf{D}_1 = [\mathrm{diag}\,(\mathbf{M}^{*\mathrm{T}}\mathbf{M}^*)^{-1}]^{\frac{1}{2}}$. As a result, any transformation matrix $\mathbf{M}$ can be described in terms of a series of orthonormal and length-adjusting transformations of the PCs.

There are two types of case II procedures recommended by Harris and Kaiser: the independent cluster (IC) solution and the proportional ($\mathbf{B}^{\mathrm{T}}\mathbf{B}$) solution. The IC solution is best applied when the clusters of variables do *not* overlap and each variable loads highly on only one PC. When the data are more complex, the $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ solution is recommended. The graphical pairwise plots can be used to help determine which solution yields a better structure. The case II IC procedure begins with the following equation:

$$\mathbf{M} = (\mathbf{I})\mathbf{D}_2\mathbf{T}_1\mathbf{D}_1 \tag{25}$$

where $\mathbf{T}_2$ is set to $\mathbf{I}$. (The definition of any case II procedure involves setting $\mathbf{T}_2 = \mathbf{I}$). The columns of $\mathbf{A}$ are rescaled, followed by an orthogonal transformation, and finally a rescaling again of the resulting columns. The process involves a principal axes solution $\mathbf{A} = \mathbf{V}\mathbf{D}^{\frac{1}{2}}$ where $\mathbf{V}$ is an $n \times r$ matrix of retained eigenvectors and $\mathbf{D}$ is an $r \times r$ diagonal matrix of retained eigenvalues. $\mathbf{D}_2$ in (23) is set to $\mathbf{D}^{-\frac{1}{2}}$. Then the resulting primary pattern matrix ($\mathbf{AM}$) is represented as

$$(\mathbf{AM}) = \mathbf{V}\mathbf{T}_1\mathbf{D}_1 \tag{26}$$

which can be obtained by letting $\mathbf{T}_1$ be any one of the orthogonal simple structure solutions of the eigenvectors $\mathbf{V}$ (e.g. Varimax) [*not* $\mathbf{V}\mathbf{D}^{\frac{1}{2}}$]. The matrix $\mathbf{D}_1$ normalizes the PCs.

The second type of Harris–Kaiser solution, the $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ [sometimes denoted $\mathbf{P}^{\mathrm{T}}\mathbf{P}$], is more complex and used for overlapping data clusters. In this case $\mathbf{D}_2$ is chosen to be some power of $\mathbf{D}$ typically, $\mathbf{D}_2 = \mathbf{D}^{-\frac{1}{4}}$. Then, in this case, the primary pattern matrix ($\mathbf{AM}$) is represented as

$$(\mathbf{AM}) = \mathbf{V}\mathbf{D}^{\frac{1}{4}}\mathbf{T}_1\mathbf{D}_1 \tag{27}$$

Again, $\mathbf{T}_1$ can be any orthogonal simple structure rotation of $\mathbf{V}\mathbf{D}^{\frac{1}{4}}$ (as opposed to $\mathbf{V}$ in the IC solution). The name proportional solution arises as $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{D}_1\mathbf{T}_1^{\mathrm{T}}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{T}_1\mathbf{D}_1$ where $\mathbf{\Lambda}$ is the matrix of eigenvalues. This shows that $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ is proportional to the PC intercorrelations ($\mathbf{\Phi}$), within a scaling constant $\mathbf{D}_1$. It is also possible to raise $\mathbf{V}\mathbf{D}$ to other powers (e.g. $\mathbf{V}\mathbf{D}^{\frac{1}{8}}$) and examine the pairwise plots

to assess the degree of simple structure, as Hakstian (1974) illustrates. Richman and Lamb (1985) apply the Harris–Kaiser case II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ solution to a rainfall data set.

8.5 Promax (Hendrickson and White, 1964)

The Promax oblique simple structure rotation uses an oblique Procrustes transformation to achieve a simple structure solution. The Procrustes equation (17): $\mathbf{B} = \mathbf{AT} + \mathbf{E}$ where $\mathbf{B}$ is the $n \times r$ target matrix, $\mathbf{A}$ is the $n \times r$ initial unrotated matrix, $\mathbf{T}$ is the $r \times r$ transformation matrix and $\mathbf{E}$ is the $n \times r$ matrix of discrepancies is developed in equations (17)–(19). Minimizing the discrepancies ($\mathbf{E}^{\mathrm{T}}\mathbf{E}$) by taking the partial derivative with respect to $\mathbf{T}$ and setting the solution to zero yields equation (19): $\mathbf{T} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{B}$ which Promax uses to obliquely rotate to simple structure under the assumption that an orthogonal simple structure solution is usually fairly close to an optimum oblique solution. As a result, functions of the loading elements of an orthogonal solution (e.g. Varimax) are used to construct a hypothetical oblique PC primary pattern matrix to be approximated by an oblique Procrustes transformation of the orthogonal PC primary pattern matrix. Mulaik (1972) offers a solution.

Let $\mathbf{A}^*$ be defined as an $n \times r$ orthogonal simple structure pattern matrix (Varimax PC loadings are the same as the primary PC pattern matrix under orthogonal rotation). The rows of $\mathbf{A}^*$ are then normalized as follows by h_i to construct an $n \times r$ matrix $\mathbf{G}$: $g_{ij} = a^*_{ij}/h_i$ where

$$h_i^2 = \sum_{j=1}^{r} a_{ij}^{*2}$$

A matrix $\mathbf{A}$ can now be defined which normalizes $\mathbf{G}$ by columns so that in each column of $\mathbf{G}$ the largest absolute value of an element is 1·00. Other elements of $\mathbf{G}$ are rescaled in $\mathbf{A}$ as $a_{ij} = g_{ij}/k_j$ where k_j is the absolute value of the element with the largest absolute value in the jth column. Now a hypothetical simple structure primary pattern matrix $\mathbf{B}$ can be defined with elements b_{ij} such that

$$b_{ij} = \frac{|a_{ij}^{m+1}|}{a_{ij}} \tag{28}$$

where m is a power to which the element a_{ij} in $\mathbf{A}$ is raised. The matrix $\mathbf{B}$ represents $\mathbf{A}$ raised to the mth power, but with the original sign of a_{ij}. The rationale for raising the elements of $\mathbf{A}$ to the mth power is to make those elements near-zero in magnitude approach zero more rapidly than elements further away from zero. Hendrickson and White recommend that initially $m = 4$ in most instances, although Hakstian and Abell (1974) mention that the more complex the initial orthogonal solution is, with less spread between the low and high loadings, the lower the value at which m should be set ($m = 2$ in complex solutions). The power constant m is denoted as k in section 10 to agree with the packages.

In Promax, the matrix $\mathbf{A}^*$ is the matrix to be transformed and the matrix $\mathbf{B}$ is target matrix to be approximated by application of a Procrustes target transformation. The transformation matrix $\mathbf{T}^*$ is then obtained by application of equation (19):

$$\mathbf{T}_1^* = (\mathbf{A}^{*\mathrm{T}}\mathbf{A}^*)^{-1}\mathbf{A}^{*\mathrm{T}}\mathbf{B} \tag{29}$$

Since unit length orthogonal vectors are used in $\mathbf{A}^*$, the matrix $\mathbf{A}^*\mathbf{T}_1^*$ will not be a primary pattern matrix since the associated vectors are not unit length. Therefore a transformation is needed to normalize $\mathbf{T}^*$: $\mathbf{T} = \mathbf{T}_1^*\mathbf{D}_1$ where $\mathbf{D}_1^2 = [\mathrm{diag}\,(\mathbf{T}_1^{\mathrm{T}}\mathbf{T}_1)^{-1}]$. Since $[\mathrm{diag}\,(\mathbf{T}^{\mathrm{T}}\mathbf{T})^{-1}] = [\mathrm{diag}\,\mathbf{D}_1^{-1}(\mathbf{T}_1^{\mathrm{T}}\mathbf{T}_1)^{-1}\mathbf{D}_1^{-1}] = \mathbf{I}$, this is the proper normalization of $\mathbf{T}$. The desired oblique PC primary pattern matrix, $\mathbf{A}_{\mathrm{p}}^*$, is given by

$$\mathbf{A}_{\mathrm{p}}^* = \mathbf{A}^*\mathbf{T} = \mathbf{A}^*\mathbf{T}_1\mathbf{D}_1 \tag{30}$$

Hayden (1983) applies a Promax solution to a cyclone frequency data set, and Lins (1985) applies it to streamflow data.

8.6. *DAPPFR* (*Tucker and Finkbeiner*, 1982)

The Direct Artificial Personal Probability Factor Rotation differs from the other four simple structure rotations examined for two reasons: (i) it is an oblique topological approach to simple structure solution rather than the simple maximization of an algebraic formula and, (ii) it is the only newly developed rotation tested which is not yet available on the packages. This method represents an attempt to quantify the detailed judgements involved in Thurstone's (1947) graphical approach to simple structure by fitting hyperplanes to the variable configurations and is based on earlier work by Tucker (1940, 1944, 1955). As such, it closely parallels graphically hand-rotating the factors (PCs) to achieve simple structure, as subgroups of variables determine hyperplane locations in PC space and these variables will therefore have small projections on the normal to the hyperplane. The determination of exactly which variables constitute these subgroups is sometimes difficult, since this implies that total groups of variables are partitioned into two subgroups on the basis of projections on the normal to the hyperplane. One subgroup is composed of those variables whose projections on the normal lie in the near-zero range, whereas the second subgroup contains the remaining variables. Partitioning the variables as so described amounts to a 'step function' based on the projections of the variables on the normal to the hyperplane.

In Figure 6, for example, the hyperplane width is shown as ±0·10, and this implies that variables with projections of greater than +0·10 or less than −0·10 are considered to lie outside the hyperplane, whereas variables within the ±0·10 range are considered in the hyperplane. This approach can be thought of as a weighted least-squares fit of the hyperplane to the entire set of variables when the weights are either 0 or 1 as determined by the step function, but presents a theoretical problem as the question arises: why should a variable with a projection just within the interval (e.g. 0·09) be given the full weight in the determination of the hyperplane whereas a slightly larger value (e.g. 0·11) is given a zero weight? A continuous function would help eliminate the problem if it could represent the *personal probability* that a given variable should be considered in the subgroup defining the hyperplane, which would be a function of the projection of the variable on the normal to the hyperplane. The name of the rotation arises from these concepts as the primary factor pattern matrix is operated on (hence 'direct' as in Direct oblimin), and artificial personal probability refers to the probabilities of specific variables lying in the hyperplane. In DAPPFR, the hyperplane location is defined by a weighted least-squares principle which minimizes the mean weighted projection of the variables when the weights are the personal probabilities.

The probability (p) of a theoretical common factor loading being within the hyperplane: $p\ (c = 0 \mid b)$ is calculated (after Tucker and Finkbeiner, 1982). Let $f(b \mid \beta)$ be a conditional probability density function of b (observed factor loading) given β (theoretical factor loading). The analyses of real data use observed factor loadings and do not deal directly with the theoretical loadings. Divergence of the observed factor loadings from the theoretical loadings occurs as a result of sampling error and lack of fit of the factor model to the data. The bivariate probability density function of b and β is

$$f(b, \beta) = f(\beta) f(b \mid \beta) \tag{31}$$

where $f(\beta)$ is the probability density function (p.d.f.). For a given value of b, define:

$$g_{0b} = \int_{-\beta_0}^{\beta_0} f(b, \beta)\, d\beta \quad \text{and} \quad g_{1b} = \int_{-\infty}^{-\beta_0} f(b, \beta)\, d\beta + \int_{\beta_0}^{\infty} f(b, \beta)\, d\beta \tag{32}$$

where β_0 is the hyperplane width. Then

$$p(c = 0 \mid b) = g_{0b}/(g_{0b} + g_{1b}) \tag{33}$$

The personal probability of a factor loading being within the hyperplane is given by (33). These are then employed as weights, and a weighted least-squares solution is applied to the variables in a series of trials until a stable position is obtained. A trial matrix of vectors normal to the hyperplanes (**N**) is formed and a shift matrix applied as shown by Tucker and Finkbeiner. The relationship between the

final $\mathbf{N}$ and the transformation matrix used in the determination of the simple structure solution $\mathbf{B} = \mathbf{A}\mathbf{T}^{-1}$ is given by

$$\mathbf{T}^{-1} = \mathbf{N}^{\mathrm{T}}\mathbf{D}^{-1} \tag{34}$$

where diag $(\mathbf{N}\mathbf{N}^{\mathrm{T}}) = \mathbf{I}$ and diag $((\mathbf{N}\mathbf{N}^{\mathrm{T}})^{-1}) = \mathbf{D}^{-2}$. Therefore, $\mathbf{D}$ is a diagonal matrix used to normalize the columns of $\mathbf{T}^{-1}$ such that $\mathbf{B}$ represents a primary pattern matrix.

One other important aspect of the rotation is that there are two separate procedures, one for cases when salient projections might be positive or negative (two-sided personal probability function) and the other for cases when a positive manifold is to be assumed (one-sided personal probability function). The development of both functions was accomplished by a trial-and-error basis for which Tucker and Finkbeiner list the specific constants employed. The physical importance of the option of choosing a one- or a two-sided rotation is that the specific case can be chosen to agree with the data. For example, the Lamb and Richman (1983a) 3-day rainfall total correlation matrix exhibited all significant correlations of the same sign as indicated in Figure 5(b). Consequently, all PC loadings were expected to be of one sign and a one-sided DAPPFR solution could be applied. In other cases, as in the analysis of geopotential height over a large domain where standing waves would be manifested by both large positive and negative correlations, a two-sided DAPPFR solution could be applied. It is up to the analyst to use physical insight in order to figure out which of the two procedures best corresponds to his data. Richman (1983a) applies a two-sided DAPPFR rotation to a 700 mb data set.

8.7. Orthogonal Procrustes (Schönemann, 1966)

Schönemann (1966) develops a family of techniques which are most useful in confirmatory work. However, similar to the way Promax uses oblique Procrustes as a means to achieve an oblique simple structure solution, one facet of this approach can be used as a means to achieve an orthogonal simple structure solution. This is an experimental approach to achieving an orthogonal solution by first using an oblique rotation to arrive at a target matrix and then taking advantage of Procrustes to yield a least-squares fit to the target. Conceptually, this is the reverse of the Promax rotation (which uses an initial orthogonal rotation to form the target). In this transformation, the Procrustes equation (17), is applied in the form $\mathbf{A}\mathbf{T} = \mathbf{B} + \mathbf{E}$ and solves for $\mathbf{T}$ in a least-square sense as trace $(\mathbf{E}^{\mathrm{T}}\mathbf{E})$ is minimized under the restriction that $\mathbf{T}$ be orthonormal: $\mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{T}\mathbf{T}^{\mathrm{T}} = \mathbf{I}$. There has been little or no application of orthogonal Procrustes in an exploratory simple structure mode to date; however, there may be circumstances when the statistical need for an orthogonal solution outweighs the disadvantages incurred (e.g. if the PC scores are to be used in a regression equation on independent data as by Walsh and Richman, 1981). This was the motivation for testing this new type of orthogonal solution: to discern if this approach could retain more of the accuracy advantage of an oblique rotation over a conventional orthogonal rotation such as Varimax. In this specific application the primary pattern matrix from the DAPPFR transformation was directly inserted (without rescaling) as a target matrix, $\mathbf{B}$. The resulting Procrustes output $\mathbf{A}\mathbf{T}$ was used as the orthogonal loading matrix. Since this is a new, experimental and relatively untested procedure, the results will be shown in section 10 in the hope that other investigators will apply it along with Varimax when statistical requirements indicate that an orthogonal solution is necessary.

9. MONTE CARLO SIMULATION

9.1. Rationale

In the past, the choice of which analytic rotation(s) to apply has been a 'blind' decision; that is, there have been few studies comparing how accurately (see Section 9.4 for specifics) the rotations can uncover the underlying structure for different types of data in various analyses. An initial analysis would typically be followed by a readily accessible orthogonal or oblique rotation. This was the motivating force behind this Monte Carlo study: to examine how well various factor solutions (i.e.

unrotated, orthogonally rotated, obliquely rotated and random spin) can uncover input factor structures which were randomly generated. The basic Monte Carlo procedure and format were developed by Tucker (1983) to test a new rotational technique and the theoretical development is given by Tucker *et al.* (1969) and Tucker and Finkbeiner (1982). The technique uses factor analysis as opposed to PC. Although the factor model may have a slight advantage in accurately uncovering the underlying modes of variation, the results found will be valid for either technique. A brief discussion of some of the parameters will ensue in Section 9.3.

9.2. How Thurstone's simple structure concept was incorporated into the simulation

Thurstone's simple structure concept can be seen by examining pairwise graphical PC plots as shown in Section 5.2. The relationships shown in Figure 7 will form the basis for objectively stating the amount of simple structure in the variable configurations. Recalling that the simple structure plots have many near-zero loadings and radial streaks or clusters aids in objectifying the concept. The values for the plots in the Monte Carlo procedure are randomly generated numbers between 0 and ±1. These input random numbers are used to simulate realistic looking simple structure fields by treating them as if they are factor pattern coefficients. In order to create these 'mock' fields, all values which are less than the specific pre-set value shown in Table VIII are changed to zero to ensure well-defined hyperplanes. To ensure that the hyperplane values are separate from the 'clusters', a minimum number of high loadings on one factor which are low loadings on the second must be present, and a specific number of low loadings must be present. All three of these values vary with the degree of simple structure as determined by Tucker (1983) and shown in Table VIII.

Table VIII. Relationships for various degrees of simple structure in first-order factor loadings for 25 and 50 variable simulations (Tucker, 1983)

	Simple Structure		
25 *variables*/5 *factors*	Strong	Moderate	Weak
Critical loading to be set to zero	≤0·70	≤0·55	≤0·40
Minimum number of zero loadings	15	12	10
Minimum number of loadings high on factor X and low on factor Y	5	4	3
50 *Variables*/10 *factors*			
Critical loading to be set to zero	≤0·70	≤0·55	≤0·40
Minimum number of zero loadings	30	25	20
Minimum number of loadings high on factor X and low on factor Y	5	4	3

9.3. Outline of simulation

From the preceding, it can be seen that the basic configuration contains the properties of simple structure as defined by the random numbers and the amount of simple structure required. The purpose of the simulation is to test various procedures (i.e. unrotated, rotated solutions) in their ability to retrieve the mock simple structure fields generated according to the program. To make the results closer to actual data (which contain signal and noise) a random noise component is added to the mock simple structure fields. In order to ensure control over the obliquity of the primary factors, the matrix $\mathbf{\Phi}$ is defined initially. Since $\mathbf{\Phi}$ is a correlation matrix of factor scores, it can be defined as $\mathbf{\Phi} = \mathbf{A}_2^T\mathbf{A}_2$ where $\mathbf{A}_2$ are the randomly generated loadings for the second order (see Richman (1983a) for discussion on higher-order factoring). The program at present allows for $\mathbf{\Phi} = \mathbf{I}$ if orthogonal first-order factors are desired. The procedure sets up second-order factor loadings which cluster similar first-order factors. This controls the obliquity of the first-order factors within limits similar to the simple structure guidelines shown in Table VIII and ensures distinct second-order factors. Once these have been

formed, $\mathbf{\Phi}$ is calculated. The program uses a form of the correlation equation (8) to generate unrotated common factors from the mock simple structure pattern coefficients:

$$\mathbf{B\Phi B}^{\mathrm{T}} = \mathbf{BVV}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}$$
$$\mathbf{\Phi} = \mathbf{VV}^{\mathrm{T}} \tag{35}$$

$$\mathbf{A} = \mathbf{BV} \tag{36}$$

where $\mathbf{B}$ is an $n \times r$ matrix of simple structure factor pattern coefficients, $\mathbf{\Phi}$ is the $r \times r$ factor score intercorrelation matrix, $\mathbf{V}$ is the $r \times r$ transformation matrix and $\mathbf{A}$ is the $n \times r$ unrotated factor loading matrix. At this point the white noise mentioned above is added to the matrix $\mathbf{A}$, as specified by the user. The matrix $\mathbf{A}$ is then transposed and multiplied by itself to yield a product matrix, $\mathbf{C}$.

$$\mathbf{C} = \mathbf{A}^{\mathrm{T}}\mathbf{A} \tag{37}$$

This product matrix is then decomposed into a real eigensolution via a principal axis factor solution and the eigenvalue series is checked to ensure full rank. *The unrotated output and the outputs of the various rotations with noise are all compared to the original mock simple structure input (without noise).* A random rotation is also used, which spins the axes by randomly generated values prior to the comparison, for a baseline. The parameters which can be varied are (i) whether or not the factor loadings will all be positive (positive manifold), (ii) number of variables, (iii) number of first-order factors, (iv) number of second-order factors, (v) the upper limit for the contribution of the variables' variances to the first-order factors, (vi) the lower limit for the contribution of the variables' variances to the first-order factors, (vii) the upper limit for the contribution of the variables' variances to the second-order factors, (viii) the lower limit for the contribution of the variables' variances to the second-order factors [Parameters (vii) and (viii) control the obliquity of the first-order factors since the variables in a second-order analysis are the first-order factors. If parameters (vii) and (viii) are set near or at zero, the first-order solution will have uncorrelated factors since the sum of the squares of the rows of the second-order factor loading matrix will be near-zero.], (ix) the critical value for which lower loadings are set to zero for first-order factors, (x) the critical value for which lower loadings are set to zero for second-order factors, (xi) minimum number of low loadings on each first-order factor, (xii) minimum number of low loadings on each second-order factor, (xiii) minimum number of high loadings on one factor which are low on the second factor in the first-order, (xiv) minimum number of high loadings on one factor which are low on the second factor in the second-order, (xv) lower limit of the least eigenvalue, (xvi) number of individuals in samples, (xvii) standard deviation of the noise added to the factor loadings, (xviii) initial random number generator 'seed', (xix) coefficients to 'order' the factors to ensure distinct factors and (xx) the total number of replications.

It is obvious that there are so many potential combinations of parameters that it would be an extremely complex project to examine all (or even a large subset of all) of the possible combinations. As such, only the number of variables and number of factors will be altered for a fixed amount of noise (standard deviation of noise = 0·050 which is considered a moderate amount). The results will be computed for analyses with the significant loadings being of one sign only (one-sided solution) as well as analyses with both positive and negative significant loadings (two-sided solution) with the remainder of the parameters set at the program's constants. Strong, moderate and weak simple structure conditions will be examined. It is important that both one-sided and two-sided solutions be computed as different meteorological data sets fall into each category as illustrated under the discussion of the DAPPFR rotation.

9.4. *Matching factors or components*

In the previous section it was shown how the unrotated, rotated and randomly spun solutions would be compared to the initial structure. In order to quantitatively assess the goodness-of-match, the coefficient of congruence will be used. The coefficient of congruence (g) between any factors

[components] A and B is given by (Harman, 1976)

$$g_{AB} = \frac{\sum_{j=1}^{n} b_{jA} b_{jB}}{\left[\left(\sum_{j=1}^{n} b_{jA}^2\right)\left(\sum_{j=1}^{n} b_{jB}^2\right)\right]^{\frac{1}{2}}} \tag{38}$$

where b are the component/factor loadings or pattern coefficients. The form of the coefficient is similar to a correlation coefficient, as the possible values range from +1·0 for total agreement through 0 for no relationship to −1·0 for total inverse agreement, and it corresponds to the cosine of the angular separation between the pairs of loadings. However, the coefficient of congruence uses loadings (b's) which are *not* deviations from their respective means. This is important as the mean of a component loading vector is a fixed property of that component, and translating the origin (which is implied by expressing a value as a deviation) is not warranted. Specifically, the *magnitude* of the loadings is an important aspect of component *identification*. An example of this is presented by Richman and Lamb (1985) where they compare two sets of PCs with PC A having the loadings [0·95, 0·65, 0·30, 0, −0·03, −0·06, −0·10] and PC B having the loadings [0·20, −0·10, −0·45, −0·75, −0·78, −0·81, −0·85]. These two PCs are definitely different, as many of the variables which load highly on PC A are low on PC B, which is characteristic of a simple structure solution. Accordingly, the coefficient of congruence between A and B reflects this with a value of 0·07. The question arises: why not simply correlate the two PCs? The reason why the correlation coefficient is not a good matching coefficient for PCs is that it removes the mean of each vector. When these two PCs were correlated, the correlation coefficient was 1·00, which would indicate a perfect match. Hence, the correlation coefficient is not the optimal similarity coefficient for matching components. Korth and Tucker (1975) used a Monte Carlo approach to obtain normative data on the distribution of the congruence coefficient, with a view to establishing a baseline for evaluating its significance. This led Tucker (1983) to attach the following goodness-of-match names to specific ranges of absolute congruence coefficients: 0·98 to 1·00 (excellent match); 0·92 to 0·98 (good match); 0·82 to 0·92 (borderline match); 0·68 to 0·82 (poor match); less than 0·68 (terrible match). This was done since the congruence coefficient is biased towards a higher value than the corresponding correlation coefficient. Therefore, the meaning of a congruence coefficient between two sets of PC loadings of 0·8 is not nearly as strong as would be intuitively expected with a correlation coefficient of 0·8. Any coefficient of less than approximately |0·70| is *not* any better than would be expected by randomly spinning the PC axes prior to matching the sets of PC loadings.

10. RESULTS OF THE MONTE CARLO SIMULATION

The rotations which were compared were largely available on the four mathematical/statistical packages in Table VII. Varimax was the sole orthogonal rotation available on the packages used, for the reason mentioned in Section 8.2 (which was confirmed in an earlier Monte Carlo study). However, the orthogonal Procrustes experimental transformation is also included to attempt to assess whether it is more accurate than Varimax. The oblique rotations tested are also available on the packages (with the exception of the one- and two-sided DAPPFR solutions) and represent the second-generation (i.e. mid-sixties) types of linear transformations. Some of the original oblique rotations were also tested (e.g. Binormamin and Oblimax) [not shown herein] an the results were clearly inferior to those rotations tested. Both the Harris–Kaiser Case II IC and Harris–Kaiser Case II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ rotations were tested along with Direct oblimin with $\delta = 0$ (default on packages) and both Promax $k = 2$ and Promax $k = 4$. The unrotated solution is also included along with a random rotation in which the factor axes are spun by a randomly generated Θ angle. As stated in section 9.3, the standard deviation of the random noise was set at 0·05, which represented a middle value of those tested (0·025 to 0·075). This value would represent a moderate amount of noise in an actual analysis as 95 per cent of the loadings would be expected to be perturbed by values $\leq$0·10 from their ideal values. There were 200 replications of

each rotation type within each cell to obtain a representative distribution of results. This is in contrast to a pilot study (Richman, 1983b) in which only 50 replications were drawn. The results will be subdivided into four categories: (i) one-sided twenty-five variables/five factors retained, (ii) one-sided fifty variables/ten factors retained, (iii) two-sided twenty-five variables/five factors retained, and (iv) two-sided fifty variables/ten factors retained. Additionally, a few replications were run at one hundred variables/twenty factors with results consistent with the fifty variables/ten factors except for a moderate reduction in accuracy (not shown). The cost of running 200 replications for 9 solutions for 3 strengths of simple structure on both one- and two-sided solutions at 100 variables/20 factors retained was prohibitive and only the results for five and ten factors will be shown. The five and ten factor runs are in the range of the majority of meteorological studies; therefore the results should be applicable to most research designs.

The first result is for the one-sided 25 variable/5 factor design with strong simple structure (Table IX(a)). The analysis in this cell reveals that the oblique rotations tested could all reproduce the input structure almost perfectly. Varimax did not do quite as well, with 93 per cent of its replications having a matching congruence coefficient between 0·92 and 0·98 (good range) and the orthogonal Procrustes was slightly more accurate. It is noteworthy that the unrotated results did poorly in recovering the input structure, with random rotation being the worst. The results suggest that a small study with strong simple structure of all positive (or all negative) loadings can be well described by any of the oblique methods. Although Varimax was not as good, it may be accurate enough for many applications if the pairwise plots are examined. With the simple structure at the moderate level and all other parameters, except those listed in Table VIII held constant, the analysis (Table IX(b)) still indicates mostly good to excellent matches for the various rotations. DAPPFR (one-sided solution) is now the most accurate rotation, and Promax $k = 2$, Direct oblimin, and Harris–Kaiser Case II $\mathbf{B}^{T}\mathbf{B}$ also do well. Any of these three would generally yield good results. Promax $k = 4$ and Harris–Kaiser Case II IC did not do as well, and Varimax was also below these, which is probably due to the orthogonality constraint, as the orthogonal Procrustes was only slightly better. The unrotated and random rotation were the worst, as before. The simple structure was next dropped to the weak level, indicating very little clustering and

Table IX. Results of 200 replications of the Monte Carlo simulation for 25 variables, 5 factors and 1-sided correlations. Degree of simple structure varies from strong (a), to moderate (b) and to weak (c)

		Root mean square congruence coefficient distribution									
Subjective match	Congruence coefficient	Unrotated	Normal Varimax	Direct oblimin	Harris–Kaiser II IC	Harris–Kaiser II $\mathbf{B}^{T}\mathbf{B}$	Promax $k = 4$	Promax $k = 2$	DAPPFR	Orthogonal Procrustes	Random transformation
(a) Strong simple structure											
Excellent	$g \geq 0{\cdot}98$		6	194	195	195	194	195	195	96	
Good	$g \geq 0{\cdot}92$		186	6	5	5	6	5	5	101	
Borderline	$g \geq 0{\cdot}82$	5	8							3	
Poor	$g \geq 0{\cdot}68$	192									90
Terrible	$g \leq 0{\cdot}67$	3									110
(b) *Moderate simple structure*											
Excellent	$g \geq 0{\cdot}98$		1	157	27	145	56	158	185	29	
Good	$g \geq 0{\cdot}92$		180	33	156	51	132	31	15	157	
Borderline	$g \geq 0{\cdot}82$	1	19	10	17	4	12	11		14	
Poor	$g \geq 0{\cdot}68$	198									141
Terrible	$g \leq 0{\cdot}67$	1									59
(c) Weak simple structure											
Excellent	$g \geq 0{\cdot}98$			30	1	15	2	53	99	2	
Good	$g \geq 0{\cdot}92$		131	100	56	93	120	101	75	138	
Borderline	$g \geq 0{\cdot}82$	1	61	67	109	85	67	46	26	59	10
Poor	$g \geq 0{\cdot}68$	194	8	3	25	7	10			1	158
Terrible	$g \leq 0{\cdot}67$	5			9		1				32

fewer variables in the hyperplane, and all other parameters except those listed in Table VIII were held constant (Table IX(c)). The results point out that DAPPFR (one-sided solution) does best with 87 per cent of the replications in the good to excellent range, Promax $k = 2$ was the best package rotation (77 per cent good to excellent), followed by direct oblimin (65 per cent good to excellent). The remaining rotations were a little less accurate, although they fared better than either the unrotated solutions or a random rotation. These results point to the importance of choosing one of the more accurate rotations for these types of condition, as there are a wide range of matches. This result is confirmed by the example shown in Figure 1 for the three flow types. The pairwise graphical plots indicated a very weak simple structure in a positive manifold for the 36 variable/3 PC experiment and illustrated how the DAPPFR rotation was more accurate than Varimax under these conditions.

Next, the number of variables was increased to fifty and the number of factors retained to ten. This is a much more difficult situation to achieve good matches, since, even if one factor out of ten is poorly matched, it will drag the resulting RMS congruence coefficients down considerably. (Some runs were also performed at fifty variables and five factors retained, yet the matches were almost as good as twenty-five variables and five factors, indicating that the number of factors retained was the most critical aspect in degrading the matches). The ten factor runs are in the range of some of the larger meteorological studies (e.g. Walsh *et al.* (1982) retained 9 PCs, Karl and Koscielny (1982) retained 9 PCs, Lamb and Richman (1983a, b) retained 10 PCs). The simple structure was set as one-sided strong (Table X(a)) with the best rotation being the one-sided DAPPFR with 100 per cent replications having $g \geq 0{\cdot}92$. Of the available package rotations, Promax $k = 2$ was best with almost 97 per cent of the replications in the good to excellent range with the other oblique solutions being slightly less accurate. Orthogonal Procrustes was more accurate than Varimax, which did not do well, with almost half of its replications in the borderline range (0·82 to 0·92). The unrotated results were poorer than any of the rotations, with the random spin coming in last. With simple structure at the moderate level, and all other parameters except those listed in Table VIII held constant (Table X(b)), the results indicate a large range in the ability of the solutions to recover the input data. The one-sided DAPPFR rotation was clearly best, with 57 per cent of the replications in the good to excellent range, and there was a

Table X. Results of 200 replications of the Monte Carlo simulation for 50 variables, 10 factors and 1-sided correlations. Degree of simple structure varies from strong (a), to moderate (b) and to weak (c)

		Root mean square congruence coefficient distribution									
Subjective match	Congruence coefficient	Unrotated	Normal Varimax	Direct oblimin	Harris–Kaiser II IC	Harris–Kaiser II $\mathbf{B}^T\mathbf{B}$	Promax $k = 4$	Promax $k = 2$	DAPPFR	Orthogonal Procrustes	Random transformation
(a) Strong simple structure											
Excellent	$g \geq 0{\cdot}98$		1	33	1	38	2	83	148	17	
Good	$g \geq 0{\cdot}92$		100	142	127	141	180	110	52	155	
Borderline	$g \geq 0{\cdot}82$	1	99	25	72	21	18	7		28	
Poor	$g \geq 0{\cdot}68$	182									6
Terrible	$g \leq 0{\cdot}67$	17									194
(b) *Moderate simple structure*											
Excellent	$g \geq 0{\cdot}98$								42		
Good	$g \geq 0{\cdot}92$		4	6	1	3	3	18	72	5	
Borderline	$g \geq 0{\cdot}82$		132	153	25	102	104	145	85	143	
Poor	$g \geq 0{\cdot}68$	198	64	41	165	95	93	37	1	52	26
Terrible	$g \leq 0{\cdot}67$	2			9						174
(c) Weak simple structure											
Excellent	$g \geq 0{\cdot}98$										
Good	$g \geq 0{\cdot}92$										
Borderline	$g \geq 0{\cdot}82$		38			1	6	29	35	34	
Poor	$g \geq 0{\cdot}68$	186	156	185	50	191	170	169	163	162	43
Terrible	$g \leq 0{\cdot}67$	14	6	15	150	8	24	2	2	4	157

large drop-off to the next best solution, Promax $k = 2$, with Direct oblimin slightly below this along with the two orthogonal rotations. An important point to note is that although Promax $k = 2$ and Direct oblimin had most replications in the borderline range, they are still better than the unrotated results. The simple structure was dropped to the weak level (one-sided) with all other parameters held constant (Table X(c)). The results are not particularly encouraging, as even the best rotations are only marginally better than the unrotated results and are in the poor range. Varimax was best (19 per cent of the replications in the borderline range), DAPPFR (one-sided) was only slightly behind (18 per cent) followed by orthogonal Procrustes (17 per cent) and Promax $k = 2$ (15 per cent) with the remainder of the solutions being in the same range as the unrotated results (poor). This points to a problem in dealing with many factors/PCs and weak simple structure for all solutions; hence *caution should be applied prior to physically interpreting such analyses without additional corroborating evidence.*

The second phase of the simulation incorporated two-sided rotations into the analysis, as both significant positive and negative loadings were allowed. The results for the 25 variable/5 factor solution (Table XI(a)) are similar to the corresponding one-sided case, as all of the oblique rotations do very well with over 90 per cent of the replications in the excellent range. As before, the Varimax rotation has 65 per cent of its cases in the good range and 35 per cent in the excellent, which is lower in accuracy than any of the oblique methods, but still may be sufficient for most research purposes, whereas orthogonal Procrustes fares slightly better (42 per cent excellent). The unrotated solution does poorly in recovering the input structure and the random spin does worst. When the amount of simple structure was reduced to the moderate level for two-sided data (Table XI(b)), the results indicated a wide range in the ability of the various solutions to recover the input structures. The two-sided DAPPFR rotation is clearly superior to the rest with 83 per cent of the replications having congruence coefficients $\geq 0{\cdot}92$. Promax $k = 2$, Harris–Kaiser Case II $\mathbf{B}^{T}\mathbf{B}$ and Direct oblimin are well behind DAPPFR, with orthogonal Procrustes and Varimax falling slightly behind these and the unrotated solution doing poorest. As the amount of simple structure was again dropped to the weak level for the two-sided solution, the results (Table XI(c)) indicate a further degradation of the quality of match as most of the

Table XI. Results of 200 replications of the Monte Carlo simulation for 25 variables, 5 factors and 2-sided correlations. Degree of simple structure varies from strong (a), to moderate (b) and to weak (c)

		Root mean square congruence coefficient distribution									
Subjective match	Congruence coefficient	Unrotated	Normal Varimax	Direct oblimin	Harris–Kaiser II IC	Harris–Kaiser II $\mathbf{B}^{T}\mathbf{B}$	Promax $k = 4$	Promax $k = 2$	DAPPFR	Orthogonal Procrustes	Random transformation
(a) Strong simple structure											
Excellent	$g \geq 0{\cdot}98$		70	188	181	189	187	188	191	83	
Good	$g \geq 0{\cdot}92$		128	12	18	11	12	12	9	116	
Borderline	$g \geq 0{\cdot}82$	1	2		1		1			1	
Poor	$g \geq 0{\cdot}68$	199									93
Terrible	$g \leq 0{\cdot}67$										107
(b) *Moderate simple structure*											
Excellent	$g \geq 0{\cdot}98$		14	50	30	51	42	59	134	26	
Good	$g \geq 0{\cdot}92$		142	115	80	115	115	103	32	136	
Borderline	$g \geq 0{\cdot}82$	3	34	33	86	31	40	33	30	33	
Poor	$g \geq 0{\cdot}68$	197	10	2	4	3	3	5	4	5	86
Terrible	$g \leq 0{\cdot}67$										114
(c) Weak simple structure											
Excellent	$g \geq 0{\cdot}98$										
Good	$g \geq 0{\cdot}92$		9	10	9	12	11	12	20	10	
Borderline	$g \geq 0{\cdot}82$	5	88	109	80	102	98	98	93	95	
Poor	$g \geq 0{\cdot}68$	192	101	81	106	85	90	89	87	94	90
Terrible	$g \leq 0{\cdot}67$	3	2		5	1	1	1		1	110

rotations are in the poor to borderline range. The two-sided DAPPFR rotation was the best with 10 per cent of the replications in the good range, the Harris–Kaiser Case II $\mathbf{B}^{T}\mathbf{B}$ was the best package rotation (6 per cent in the good range), and the unrotated solution was clearly lower than any of the rotations, vet the distance between the two was closer than in previous two-sided runs.

Next, the number of variables was increased to fifty and the number of factors was increased to ten for two-sided data with simple structure set as strong, and the results (Table XII(a)) indicate that DAPPFR (two-sided) had the most accuracy by a wide margin with 92 per cent of the replications having congruence coefficients $\geq 0{\cdot}92$. Promax $k = 2$ and Harris–Kaiser Case II $\mathbf{B}^{T}\mathbf{B}$ were the most accurate package rotations with between 68 and 70 per cent of the cases in the good to excellent range. Of the orthogonal rotations, Procrustes has 54 per cent of the replications in the good to excellent range with Varimax having 45 per cent in the good range; the unrotated was poorest. When the simple structure was reduced to the moderate level for two-sided solutions, the resulting accuracy for all rotated solutions dropped considerably (Table XII(b)) as all rotated solutions, with the exception of the two-sided DAPPFR, were very close to the unrotated accuracy. This is a decline from the one-sided results in Table X(b) and points to the difficulty in accurately reproducing modes with large positive and negative features. DAPPFR does somewhat better with 5 per cent of the cases in the good range and over 25 per cent having congruence coefficients $\geq 0{\cdot}82$ (borderline or better). The only package rotations slightly better than the unrotated solution are Promax $k = 2$ and, possibly, Harris–Kaiser Case II $\mathbf{B}^{T}\mathbf{B}$. One possibility for a data set with these conditions would be to apply the DAPPFR or Promax $k = 2$ rotation to roughly place the PC axes and then use a computerized visually guided graphical plot program such as ROTOPLOT (Cattell and Foster, 1963) or Tucker's (1955) semi-analytic procedure to define the best simple structure position of the PC axes with the maximum number of variables in the hyperplanes. Finally, when the simple structure was reduced to the weak level, indicating very little clustering and few variables in the hyperplane, the results (Table XII (c)) indicate that this is the most difficult situation for *any* solution to accurately capture, and further indicate that simple structure rotation does not work well when the two-sided data exhibit little simple structure and many factors are retained. The most obvious result is that the unrotated solution does best in this one

Table XII. Results of 200 replications of the Monte Carlo simulation for 50 variables, 10 factors and 2-sided correlations. Degree of simple structure varies from strong (a), to moderate (b) and to weak (c)

		Root mean square congruence coefficient distribution									
Subjective match	Congruence coefficient	Unrotated	Normal Varimax	Direct oblimin	Harris–Kaiser II IC	Harris–Kaiser II $\mathbf{B}^{T}\mathbf{B}$	Promax $k = 4$	Promax $k = 2$	DAPPFR	Orthogonal Procrustes	Random transformation
(a) Strong simple structure											
Excellent	$g \geq 0{\cdot}98$			1		4	1	13	114	10	
Good	$g \geq 0{\cdot}92$		90	80	69	135	118	123	70	97	
Borderline	$g \geq 0{\cdot}82$		101	109	122	57	72	58	16	89	
Poor	$g \geq 0{\cdot}68$	196	9	10	9	4	9	6		4	1
Terrible	$g \leq 0{\cdot}67$	4									199
(b) *Moderate simple structure*											
Excellent	$g \geq 0{\cdot}98$										
Good	$g \geq 0{\cdot}92$								10		
Borderline	$g \geq 0{\cdot}82$	1	5	2	3	17	8	17	41	15	
Poor	$g \geq 0{\cdot}68$	183	166	156	135	158	161	167	139	163	2
Terrible	$g \leq 0{\cdot}67$	16	29	42	62	25	31	16	10	22	198
(c) Weak simple structure											
Excellent	$g \geq 0{\cdot}98$										
Good	$g \geq 0{\cdot}92$										
Borderline	$g \geq 0{\cdot}82$										
Poor	$g \geq 0{\cdot}68$	189	7	13	9	8	8	9	9	7	3
Terrible	$g \leq 0{\cdot}67$	11	193	187	191	192	192	191	191	193	197

case, although the best replications are in the poor range, indicating a questionable correspondence to the input data. Every simple structure rotation is even worse than the unrotated solution, being near or at the random spin accuracy level. Data in this configuration might benefit from a Procrustes transformation if there is sufficient prior knowledge concerning the nature of the modes of variation. If this is not the case, another type of analysis technique may be warranted.

11. SUMMARY AND DISCUSSION

This study has outlined some pitfalls of the conventional use of unrotated principal components, examined the theory behind simple structure and compared the accuracy of both unrotated and rotated solutions. Owing to the exploratory nature of meteorological applications of principal component analysis, most research in the literature has not examined how well the unrotated solutions fare in capturing the modes of variation of data sets. Within the past decade, Buell (1975, 1979) has illustrated the domain shape dependence of unrotated EOF/PCs where the topography of the patterns was shown to be predictable and primarily a function of the geometric shape of the domain and not the covariation of the data. Richman and Lamb (1985) illustrated the lack of stability in the modes of variation when subdomains are analysed separately. The significance of this is that the unrotated solutions confound the analysis by yielding two different sets of modes of decomposition. The third disadvantage of unrotated solutions occurs when the analysis has closely spaced eigenvalues. North *et al.* (1982) and Kendall (1980) point out that the sampling errors become very large if neighbouring eigenvalues are very close and the EOF patterns can become intermixed. The effects of nearly equal eigenvalues presented herein further support North *et al.*'s claims, as sample sizes as large as 10,000 could not capture the correct unrotated population patterns. This problem can be so severe that Storch and Hannoschöck (1985) have recommended renouncing a physical interpretation of unrotated EOFs in many instances. A final reason why unrotated solutions may not provide the most accurate decomposition of the input data into modes of variation was shown by Richman and Lamb (1985) who compared the set of unrotated PC loadings of their rainfall data to the actual set of variations embedded within the correlation matrix teleconnection patterns. The two sets of patterns were visually inspected and quantitatively matched with the correlation and congruence coefficients. The results of their study indicated that, for the rainfall data set, the modes of variation were all of one sign (i.e. all rainfall stations were either positively correlated to each other or had near-zero correlations). The corresponding unrotated PCs had flip-flops of PC loadings from PC2 to PC10 which had *no* significant physical basis but were simply artefacts of the technique. Horel (1984, p. 1661) neatly sums up the state of the unrotated solution when he remarks that 'principal component analysis was used for many years before its inherent limitations were fully realized'.

The four basic disadvantages of unrotated solutions have been shown to be reduced by using the rotated solutions. Domain shape dependence does not appear to be a problem, subdomains have virtually identical patterns to the full domain, sampling errors are greatly reduced even if the neighbouring eigenvalues are almost equal, *as long as there is structure within the input data* (e.g. see Tables IX(a), X(a), XI(a) and XII(a)) and, finally, the rotated solutions frequently do a better job of agreeing with the physical structure of the data as embedded within the correlation or covariance input matrix. These four advantages of simple structure rotations make them good candidates to accurately identify modes of variation in climatological data sets. A number of studies have recently applied rotation to the unrotated CFs/PCs/EOFs with successful results, including those of Horel (1981, 1984), Richman (1981), Walsh and Richman (1981), Balling and Lawson (1982), Karl and Koscielny (1982), Karl *et al.* (1982), Lamb and Richman (1983a, b), Walsh *et al.* (1982), Cohen (1983), Hayden (1983), Ashbaugh *et al.* (1984, 1985), Ronberg and Wang (1985), Barnston and Livezey (1985), Englehart and Douglas (1985) and Thurston and Spengler (1985). These studies involved meteorological fields such as temperature, rainfall, sea-surface temperature, cyclone frequency, chemical ions and geopotential height, yet there have been many other unrotated EOF/PC studies involving parameters (e.g. vertical velocity (Le Drew, 1980); IR radiation (Weickmann, 1983); vertical temperature profiles (Uddstrom

and Wark, 1985) etc.) which have never been rotated. Application of rotation to such fields would be interesting to assess the amount of simple structure inherent in such data.

The results concerning the six ways to present a data set to maximize its physical interpretation are important when rotation is applied to the analysis. In the past, with unrotated solutions, the researcher typically set the variables to the smaller dimension of the data matrix and the cases to the larger dimension since the results are equivalent either way (Hirose and Kutzbach, 1969). For example, if a data matrix had 180 observations of long-wave radiation for 654 grid points, the typical unrotated analysis would set the observations to variables and the grid points to cases (in a T-mode analysis) for its cost-effectiveness (Weickmann, 1983). In a rotated solution each of the six modes of decomposition yields a *unique clustering* of the variables in an analysis; therefore some meteorological insight should be applied *prior* to the analysis to determine which mode makes the most physical sense theoretically. An example is presented in section 6 which illustrates which parameters each mode clusters along with some suggestions. There may even be situations where none of the six modes are suitable and, if there is sufficient prior knowledge concerning potential modes of variation, a Procrustes transformation can be applied. In this situation, any set of loadings can be specified, *a priori*, and they do not have to be in the form of an unrotated or a simple structure solution.

A suggested analysis 'plan of attack' for a data set would first include choosing the most physically meaningful mode of decomposition (e.g. P, S, T, etc.) from Figure 9, relating the data with one of the dispersion matrices, choosing the most appropriate eigenmodel and deciding on the number of EOFs/PCs/factors to retain, followed by application of one of the more accurate simple structure rotations under a wide variety of conditions. This work points to two: DAPPFR or Promax $k = 2$ if a mathematical/statistical package is available. By using one of these two, the chances of arriving at the optimal simple structure solution are maximized prior to interpreting the results. The construction of the graphical pairwise plots of the rotated PC loadings illustrated in section 5 is a very important step in an analysis. By inspecting the $\frac{1}{2}r\ (r-1)$ plots and comparing these to Figure 7, the overall degree of simple structure within the data set can be assessed. The next step involves scanning the input dispersion matrix for the magnitudes and signs of the relationships and rotated loadings (primary pattern matrix in oblique cases) to detect whether salient or significant loadings should be all of one sign (one-sided) or contain both large positive and negative values (two-sided). At this point, information on the number of variables, number of rotated PCs retained, signs of the loadings and overall degree of simple structure can be assembled and compared to the results of the Monte Carlo simulation in Tables IX–XII to find what solution (unrotated, rotated) should be the most accurate for the data under investigation. If the initial guess was not one of the better solutions, the principal components can be re-rotated (i.e. to another criterion) if necessary and the pairwise plots of PC loadings inspected to ensure a reasonable simple structure. Analysis can then be made directly from the primary pattern matrix (under oblique rotation) or from the PC loading matrix (orthogonal rotation or unrotated solutions).

A final note concerns the results of the Monte Carlo simulation. The newly created DAPPFR rotation was the most accurate solution in 10 of 12 situations (Tables IX–XII) which is particularly noteworthy as it is the only analytic transformation which fully attempts to fulfil Thurstone's concept of simple structure (rather than simply maximizing the numbers of both low and high loadings). Promax $k = 2$ appears to be the most accurate widely available solution, tying with DAPPFR in one case and being placed second in seven others. It is interesting to note how the accuracy of the Promax rotation dropped considerably at the default value of $k = 4$. The Harris–Kaiser case II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ and Direct oblimin rotations also did fairly well, scoring in the top three solutions between five and seven times. The one rotation which has been most frequently applied to meteorological data, Varimax, does not fare as well, scoring highest only in Table X(c). This does not necessarily mean that Varimax is a poor choice, but rather that it should be compared to one of the more accurate oblique methods (i.e. DAPPFR, Promax $k = 2$, Harris–Kaiser Case II $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ or Direct oblimin) to ensure that the orthogonality constraint has not degraded the solution substantially (see Richman and Lamb, 1985, for example). There may be circumstances when the need for an orthogonal solution outweighs the disadvantages incurred; for

example, if the PC scores are to be used in a regression equation on independent data. In these cases one alternative approach would be to initially apply the orthogonal Procrustes transformation based on a DAPPFR or Promax $k = 2$ target matrix to obtain uncorrelated PC scores which are input into a regression equation. The IMSL package offers orthogonal Procrustes under program 'OFSCHN'. The initial results using this option appear positive, as it was slightly to moderately more accurate than Varimax in 10 of 12 cases. Further applications on meteorological data (e.g. Lamb and Richman, 1986) are needed to establish the utility of this procedure. The traditional use of unrotated PCs was found to yield low congruence coefficients in all 12 cases yet was superior to randomly spinning the components prior to matching. However, it was the most accurate match in the case of weak simple structure which had numerous PCs retained with both negative and positive correlations; consequently, unrotated solutions might have utility under these conditions (e.g. Richman and Walsh, 1985). In the other 11 situations the use of one of the various rotations more accurately captured the modes of variation. Tables IX–XII provide a means to extract this information so that the choice of a final solution no longer has to be a blind decision. No one solution (unrotated, rotated) or specific criterion will *always* yield the most accurate results; however, specific ones are shown to work well for large classes of data. Consequently, unyielding adherence to any one solution within a particular eigenmodel for all types of data will ultimately lead to disappointing results (e.g. Legates and Willmott (1983) and Daultrey (1976) posit that PCs should never be rotated under any circumstances based on Mather's (1971, 1972) comments to Davies (1971a, b; 1972)). If any eigenmodel may be less appropriate for a number of analytic rotation schemes, it is EOFs (due to the unit-length normalization of the eigenvectors), as some package algorithms (e.g. Harris–Kaiser) are designed for the input of eigenvectors whose coefficients have been scaled by the square-root of the corresponding eigenvalue (i.e. PC/CF loadings), while others (e.g. Promax) raise the coefficients to values as high as the 4th power. This can lower salient EOF coefficients to insignificantly small values when a large number of variables are analyzed (due to the normalization). The investigator must ultimately use his knowledge of the substance and literature in his domain, combined with judicious application of the optimal experimental design and solution (unrotated, simple structure rotation, Procrustes), to arrive at the most meaningful analysis.

ACKNOWLEDGEMENTS

This work was partly supported by NSF Grant ATM 81-00158. The author would like to thank Peter Lamb, John Walsh, John Horel, Bob Livezey, Rudy Preisendorfer, John Bistline and two anonymous reviewers for their beneficial suggestions and comments during various stages of this work. Additional thanks go to Rebecca Runge for typing the manuscript and to John Brother for drafting the Figures. The author extends particular gratitude to Ledyard Tucker for sharing his keen insight into the nature of simple structure and for supplying a source program of the simulation package for modification without which this work would not have been possible.

REFERENCES

Ashbaugh, L. L., Myrup, L. O. and Flocchini, R. G. 1984. 'A principal component analysis of sulfer concentrations in the western United States', *Atmos. Environ.*, **18,** 783.

Ashbaugh, L. L., Myrup, L. O. and Flocchini, R. G. 1985. 'Authors' reply to discussion on A principal component analysis of sulfer concentrations in the western United States', *Atmos. Environ.*, **19,** 1973.

Balling, R. C. and Lawson, M. P. 1982. 'Twentieth century changes in winter climatic regions', *Climatic Change,* **4,** 57.

Barnston, A. G. and Livezey, R. E. 1985. 'High resolution rotated empirical orthogonal function analysis of Northern Hemisphere 700 mb heights for predictive purposes', *Preprints Ninth Conf. on Prob. and Stat. in Atmos. Sci.,* Virginia Beach, VA, Amer. Meteor. Soc., 290.

Bedi, H. S. and Bindra, M. M. S. 1980. 'principal components of monsoon rainfall', *Tellus,* **32,** 296.

BMDP. 1981. *Biomedical Computer Programs P-Series,* University of California Press, Berkeley, CA 94720.

Brier, G. W. and Meltesen, G. T. 1976. 'Eigenvector analysis for prediction of time series', *J. Appl. Meteor.*, **15,** 1307.

Brinkmann, W. A. R. 1981. 'Sea-level pressure patterns over eastern North America', *Mon. Wea. Rev.*, **109,** 1305.

Buell, C. E. 1975. 'The topography of empirical orthogonal functions', *Preprints Fourth Conf. on Prob. and Stats, in Atmos. Sci.,* Tallahassee, FL, Amer. Meteor. Soc. 188.

Buell, C. E. 1979. 'On the physical interpretation of empirical orthogonal functions', *Preprints Sixth Conf. on Prob. and Stats. in Atmos. Sci.*, Banff, Alta., Amer. Meteor. Soc., 112.
Cahalan, R. F. 1983. 'EOF spectral estimation in climate analysis', *Preprints Second International Conf. on Stat. Climat.*, Lisbon, Portugal, National Institute of Meteor. and Geophysics, 4.5.1.
Carroll, J. B. 1953. 'An analytic solution for approximating simple structure in factor analysis', *Psychometrika*, **18**, 23.
Carroll, J. B. 1957. 'Biquartimin criterion for rotating to oblique simple structure in factor analysis', *Science*, **126**, 1114.
Cattell, R. B. 1952. *Factor Analysis*, Harper and Row, New York, NY.
Cattell, R. B. 1966a. 'The scree test for the number of factors', *Multivar. Behav. Res.*, **1**, 245.
Cattell, R. B. 1966b. *Handbook of Multivariate Experimental Psychology*, Rand McNally, Chicago, IL.
Cattell, R. B. and Foster, M. J. 1963. 'The rotoplot program for multiple, single-plane, visually-guided rotation', *Behav. Sci.*, **8**, 156.
Cattell, R. B. and Muerle, J. L. 1960. 'The "maxplane" program for factor rotation to oblique simple structure', *Educ. and Psych. Measur.*, **20**, 269.
Christenson, W. I. and Bryson, R. A. 1966. 'An investigation of the potential of component analysis for weather classification', *Mon. Wea. Rev.*, **94**, 697.
Cliff, N. and Hamburger, C. D. 1967. 'A study of sampling errors in factor analysis by means of artificial experiments', *Psych. Bull.*, **68**, 430.
Cohen, S. J. 1983. 'Classification of 500 mb height anomalies using obliquely rotated principal components', *J. Clim. Appl. Meteor.*, **22**, 1975.
Craddock, J. M. 1965. 'A meteorological application of factor analysis'. *The Statistician*, **15**, 143.
Craddock, J. M. and Flood, C. R. 1969. 'Eigenvectors for representing the 500-mb geopotential surface over the Northern Hemisphere', *Quart. J. Roy. Meteor. Soc.*, **95**, 576.
Crawford, C. B. 1967. 'A general method of rotation for factor analysis', Spring Meeting of the Psychometric Society, Madison, WI.
Daultrey, S. 1976. '*Principal Components Analysis*', Geo Abstracts Ltd, Norwich, Great Britain.
Davies, W. K. D. 1971a. 'Varimax and the destruction of generality. A methodological note', *Area*, **3**, 112.
Davies, W. K. D. 1971b. 'Varimax and generality: A reply', *Area*, **3**, 254.
Davies, W. K. D. 1972.'Varimax and generality: A second reply', *Area*, **4**, 207.
Davis, R. E. 1976. 'Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean', *J. Phys. Oceanogr.*, **6**, 249.
Diaz, H. F. 1981. 'Eigenvector analysis of seasonal temperature, precipitation and synoptic-scale system frequency over the contiguous United States. Part II: Spring, summer, fall and annual', *Mon. Wea. Rev.*, **109**, 1285.
Diaz, H. F. and Fulbright, D. C. 1981. 'Eigenvector analysis of seasonal temperature, precipitation and synoptic-scale system frequency over the contiguous United States. Part I: Winter', *Mon. Wea. Rev.*, **109**, 1267.
Dyer, T. G. J. 1975. 'The assignment of rainfall stations into homogeneous groups: An application of principal component analysis', *Quart. J. Roy. Meteor. Soc.*, **101**, 1005.
Dyer, T. G. J. 1979. 'Rainfall along the east coast of southern Africa, the southern oscillation and the latitude of the subtropical high pressure belt', *Quart. J. Roy. Meteor. Soc.*, **105**, 445.
Eckart, C. and Young, G. 1936. 'The approximation of one matrix by another of lower rank', *Psychometrika*, **1**, 211.
Englehart, P. J. and Douglas, A. V. 1985. 'A statistical analysis of precipitation frequency in the conterminous United States, including comparisons with precipitation totals', *J. Clim. Appl. Meteor.*, **24**, 350.
Fukuoka, A. 1951. 'A study of 10-day forecast (a synthetic report)', *Geophys. Mag.*, **22**, 177.
Gatz, D. F. 1978. 'Identification of aerosol sources in the St. Louis area using factor analysis', *J. Appl. Meteor.*, **17**, 600.
Gilman, D. L. 1957. *Empirical Orthogonal Functions Applied to Thirty-Day Forecasting*. M.I.T. Dept. of Meteorology Science Report, No. 1. Contract AF19(604)–1283.
Glahn, H. R. 1965. 'Objective weather forecasting by statistical methods', *The Statistician*, **15**, 111.
Gray, B. M. 1981. 'On the stability of temperature eigenvector patterns', *J. Climatol.*, **1**, 273.
Gregory, S. 1975. 'On the delineation of regional patterns of recent climatic fluctuations', *Weather*, **30**, 276.
Hakstian, A. R. 1974. 'The development of a class of oblique factor solutions', *Brit. J. Math. Stat. Psych.*, **27**, 100.
Hakstian, A. R. and Abell, R. A. 1974. 'A further comparison of oblique factor transformation methods', *Psychometrika*, **39**, 429.
Hakstian, A. R., Rogers, W. T. and Cattell, R. B. 1982. 'The behavior of number-of-factors rules with simulated data', *Multivar. Behav. Res.*, **17**, 193.
Harman, H. H. 1976. *Modern Factor Analysis*. The University of Chicago Press, Chicago, IL.
Harris, C. W. and Kaiser, H. F. 1964. 'Oblique factor analytic solutions by orthogonal transformations', *Psychometrika*, **29**, 347.
Hastenrath, S. and Wendland, W. M. 1979. 'On the secular variation of storms in the tropical North Atlantic and Eastern Pacific', *Tellus*, **31**, 28.
Hayden, B. P. 1983. 'Prediction of seasonal temperatures based on persistences in cyclone frequency EOFs', *Preprints Second International Conf. on Stat. Climat.*, Lisbon, Portugal, National Institute of Meteor. and Geophysics, 13.3.1.
Hendrickson, A. E. and White, P. O. 1964. 'Promax: a quick method to oblique simple structure', *Brit. J. Stat. Psych.*, **17**, 65.
Hirose, M. and Kutzbach, J. E. 1969. 'An alternate method for eigenvector computation', *J. Appl. Meteor.*, **8**, 701.
Horel, J. D. 1981. 'A rotated principal component analysis of the interannual variability of the northern hemisphere 500 mb height field', *Mon. Wea. Rev.*, **109**, 2080.
Horel, J. D. 1984. 'Complex principal component analysis: theory and examples', *J. Clim. Appl. Meteor.*, **23**, 1660.
Horn, J. L. 1965. 'A rationale and test for the number of factors in factor analysis', *Psychometrika*, **30**, 179.
Hotelling, H. 1933. 'Analysis of a complex of statistical variables into principal components,' *J. Educ. Psych.*, **24**, 417, 498.
Hsuing, J. and Newell, R. E. 1983. 'The principal nonseasonal modes of variation of global sea surface temperature', *J. Phys. Oceanogr.*, **13**, 1957.

Humphreys, L. G. and Montanelli, R. G. 1975. 'An investigation of the parallel analysis criterion for determining the number of common factors', *Multivar. Behav. Res.*, **10,** 193.
Hurley, J. R. and Cattell, R. B. 1962. 'The Procrustes program: producing direct rotation to test a hypothesized factor structure', *Behav. Sci.*, **7,** 258.
IMSL. 1979. *International Mathematical and Statistical Laboratory, Edition 7, Volume 3,* Chapter 0, IMSL Publ., Houston, TX.
Jennrich, R. I. and Sampson, P. F. 1966. 'Rotation for simple loadings', *Psychometrika,* **31,** 313.
Johnston, R. J. 1981. 'Regarding the delimitation of regions according to climatic fluctuations', *Arch. Meteor. Geophys. Biokl., Ser. B.*, **29,** 215.
Kaiser, H. F. 1958. 'The Varimax criterion for analytic rotation in factor analysis', *Psychometrika,* **23,** 187.
Kaiser, H. F. 1959. 'Computer program for Varimax rotation in factor analysis', *Educ. Psych. Meas.*, **19,** 413.
Kaiser, H. F. 1974. 'A note on the Equamax criterion', *Multivar. Behav. Res.*, **9,** 501.
Kaiser, H. F. and Cerny, B. A. 1978. 'Casey's method for fitting hyperplanes from an intermediate Orthomax solution', *Multivar. Behav. Res.*, **13,** 395.
Kaiser, H. F. and Dickman, K. W. 1959. 'Analytic determination of common factors', *Amer. Psych.*, **14,** 425.
Kaiser, H. F. and Dickman, K. W. 1977. 'Some properties of Binormamin', *Psychometrika,* **42,** 307.
Karl, T. R. and Koscielny, A. J. 1982. 'Drought in the United States', *J. Climatol.*, **2,** 313.
Karl, T. R., Koscielny, A. J. and Diaz, H. F. 1982. 'Potential errors in the application of principal component (eigenvector) analysis to geophysical data', *J. Appl. Meteor.*, **21,** 1183.
Katz, J. O. and Rohlf, R. J. 1974. 'Functionplane—a new approach to simple structure rotation', *Psychometrika,* **39,** 37.
Kendall, M. G. 1980. *Multivariate Analysis,* Charles Griffin, High Wycombe.
Klaus, D. 1978. 'Spatial distribution and periodicity of mean annual precipitation south of the Sahara', *Arch. Meteor. Geophys. Biokl., Ser. B,* **26,** 17.
Korth, B. A. and Tucker, L. R. 1975. 'The distribution of chance congruence coefficients from simulated data', *Psychometrika,* **40,** 361.
Kutzbach, J. E. 1967. 'Empirical eigenvectors of sea-level pressure, surface temperature, and precipitation complexes over North America', *J. Appl. Meteor.*, **6,** 791.
Kutzbach, J. E. 1969. 'Large-scale features of monthly mean northern hemisphere anomaly maps of sea-level pressure', *Mon. Wea. Rev.*, **98,** 708.
Lamb, P. J. and Richman, M. B. 1983a. 'Regionalization of central United States for short-period summer rainfall', *Proc. Seventh Climate Diagnostics Workshop,* Boulder, CO, 180.
Lamb, P. J. and Richman, M. B. 1983b. 'An analysis of the space and time variation of growing season rainfall in the central United States', *Preprints Eighth Conf. on Prob. and Stats. in Atmos. Sci.*, Hot Springs, AR, Amer. Meteor. Soc., 49.
Lamb, P. J. and Richman, M. B. 1986. 'On the modes of variation of growing season rainfall in the central United States', to be submitted to *Mon. Wea. Rev.*
Le Drew, E. F. 1980. 'Eigenvector analysis of the vertical velocity fields over the eastern Canadian arctic', *Mon. Wea. Rev.*, **108,** 1992.
Legates, D. R. and Willmott, C. J. 1983. 'A comparative evaluation of principal components-based and information theory methods of precipitation regionalization', *Archiv. Meteor. Geophys. Biokl, Ser. B,* **32,** 381.
Lins, H. F. 1985. 'Streamflow variability in the United States: 1931–1978', *J. Clim. Appl. Meteor.*, **24,** 463.
Lorenz, E. N. 1956. 'Empirical orthogonal functions and statistical weather prediction', *Sci. Rep.* 1, Statistical Forecasting Project, Dept. of Meteor., Mass. Institute of Technology [NTIS AD 110268].
Mather, P. M. 1971. 'Varimax and generality: A comment', *Area,* **3,** 252.
Mather, P. M. 1972. 'Varimax and generality: A second comment', *Area,* **4,** 27.
McGuirk, J. P. 1982. 'A century of precipitation variability along the Pacific coast of North America and its impact', *Climatic Change,* **4,** 41.
Molteni, F., Bonelli, P. and Bacci, P. 1983. 'Precipitation over Northern Italy: A description by means of principal component analysis', *J. Clim. Appl. Meteor.*, **22,** 1738.
Morin, G., Fortin, J. P., Sochanska, W. and Lardeau, J. P. 1979. 'Use of principal component analysis to identity homogeneous precipitation stations for optimal interpolation', *Water Resources Res.*, **15,** 1841.
Mulaik, S. A. 1972. *The Foundations of Factor Analysis,* McGraw Hill, New York, NY.
Neuhaus, J. O. and Wrigley, C. 1954. 'The Quartimax method: an analytical approach to simple structure', *Brit. J. Stat. Psych.*, **7,** 81.
North, G. R. 1984. 'Empirical orthogonal functions and normal modes,' *J. Atmos. Sci.*, **41,** 879.
North, G. R., Bell, T. L., Cahalan, R. F. and Moeng, F. J. 1982. 'Sampling errors in the estimation of empirical orthogonal functions', *Mon. Wea. Rev.*, **110,** 699.
Overland, J. E. and Preisendorfer, R. W. 1982. 'A significance test for principal components applied to cyclone climatology', *Mon. Wea. Rev.*, **110,** 1.
Preisendorfer, R. W. and Barnett, T. P. 1977. 'Significance tests for empirical orthogonal functions', *Preprints Fifth Conf. on Prob. and Stats. in Atmos. Sci.*, Las Vegas, NV, Amer. Meteor. Soc., 169.
Preisendorfer, R. W., Zwiers, F. W. and Barnett, T. P. 1981. *Foundations of Principal Component Selection Rules,* SIO Reference Series 81-4, Scripps Institute of Oceanography, La Jolla, CA.
Resio, D. T. and Hayden, B. P. 1975. 'Recent secular variations in mid-Atlantic winter extratropical storm climate', *J. Appl. Meteor.*, **14,** 1223.
Richman, M. B. 1981. 'Obliquely rotated principal components: An improved meteorological map typing technique?', *J. Appl. Meteor.*, **20,** 1145.
Richman, M. B. 1983a. 'Specification of complex modes of circulation with T-mode factor analysis', *Preprints Second International Conf. on Stat. Climate.*, Lisbon, Portugal, National Institute of Meteor. and Geophysics, 5.1.1.
Richman, M. B. 1983b. 'Rotation of principal components in climatological research. Part 2: how the various analytic simple

structure rotations on the major statistical packages compare on different types of data', *Preprints Eighth Conf. on Prob. and Stat. in Atmos. Sci.*, Hot Springs, AR, Amer. Meteor. Soc., 115.
Richman, M. B. 1985. 'The eigentechniques, pseudo-random distributions and the interpretation of spurious signal', *Preprints Ninth Conf. on Prob. and Stats. in Atmos. Sci.*, Virginia Beach, VA, Amer. Meteor. Soc., 262.
Richman, M. B. and Lamb, P. J. 1985. 'Climatic pattern analysis of 3- and 7-day summer rainfall in the central United States: Some methodological considerations and a regionalization', *J. Clim. Appl. Meteor.*, **24,** 1325.
Richman, M. B. and Walsh, J. E. 1985. 'Hemispheric pattern analysis of weekly 500 mb data', *Preprints Ninth Conf. on Prob. and Stats. in Atmos. Sci.*, Virginia Beach, VA, Amer. Meteor. Soc., 276.
Rogers, J. C. 1976. 'Sea surface temperature anomalies in the eastern North Pacific and associated wintertime atmospheric fluctuations over North America, 1960–1973', *Mon. Wea. Rev.*, **104,** 985.
Ronberg, B. and Wang, W-C. 1985. 'Climate regions derived from Chinese long-term precipitation records', *Preprints Ninth Conf. on Prob. and Stats. in Atmos. Sci.*, Virginia Beach, VA, Amer. Meteor. Soc., 301.
Rummel, R. J. 1970. *Applied Factor Analysis,* Northwestern University Press, Evanston, IL.
Salinger, M. J. 1980. 'New Zealand climate: I. Precipitation patterns', *Mon. Wea. Rev.*, **108,** 1892.
SAS. 1982. *SAS Institute Inc. User's Guide: Statistics,* SAS Institute Publ., Cary, NC.
Saunders, D. R. 1961. 'The rationale for an "Oblimax" method of transformation in factor analysis', *Psychometrika,* **26,** 317.
Saunders, D. R. 1962. 'Trans-Varimax: Some properties of the Ratiomax and Equamax criteria for blind orthogonal rotation', *Amer. Psych.*, **17,** 395.
Schönemann, P. H. 1966. 'A generalized solution of the orthogonal Procrustes program', *Psychometrika,* **31,** 1.
Sellers, W. D. 1957. 'A statistical-dynamic approach to numerical weather prediction', *M.I.T. Department of Meteorology, Scientific Report No. 2,* Statistical Forecasting Project, Contract AF19(604)-1566.
SPSS. 1979. *Statistical Package for the Social Sciences,* Third Edition, McGraw Hill Book Co., New York, NY.
Storch, H. v. and Hannoschöck, G. 1985. 'Statistical aspects of estimated principal vectors (EOFs) based on small sample sizes', *J. Clim. Appl. Meteor.*, **24,** 716.
Tabony, R. C. 1981. 'A principal component and spectral analysis of European rainfall', *J. Climatol.* **1,** 283.
Thurston, G. D. and Spengler, J. D. 1985. 'A multivariate assessment of meteorological influences on inhalable particle source impact', *J. Clim. Appl. Meteor.*, **24,** 1245.
Thurstone, L. L. 1947. *Multiple Factor Analysis,* The University of Chicago Press, Chicago, IL.
Tucker, L. R. 1940. 'A rotational method based upon the mean principal axis of a subgroup of tests', *Psych. Bull.*, **37,** 578.
Tucker, L. R. 1944. 'A semi-analytical method of factorial rotation to simple structure', *Psychometrika,* **9,** 43.
Tucker, L. R. 1955. 'The objective definition of simple structure in linear factor analysis', *Psychometrika,* **20,** 209.
Tucker, L. R. 1983. Personal communication.
Tucker, L. R. and Finkbeiner, C. T. 1982. 'Transformation of factors by artificial personal probability functions', *ETS research report 81–58, test and measurement no. TM 820429* [Available from ERIC clearinghouse, Educational Testing Service, Princeton, NJ 08541, U.S.A.].
Tucker, L. R. Koopman, R. F. and Linn, R. L. 1969. 'Evaluation of factor analytic procedures by means of simulated correlation matrices', *Psychometrika,* **34,** 421.
Uddstrom, M. J. and Wark, D. Q. 1985. 'A classification scheme for satellite temperature retrievals', *J. Clim. Appl. Meteor.*, **24,** 16.
Vargas, W. M. and Compagnucci, R. H. 1983. 'Methodological aspects of principal component analysis in meteorological fields', *Preprints Second International Conf. on Stat. Climat.*, Lisbon, Portugal, National Institute of Meteor. and Geophysics, 5.3.1.
Veitch, L. G. 1965. 'The description of Australian pressure fields by principal components', *Quart. J. Roy. Meteor. Soc.*, **91,** 184.
Walsh, J. E. and Mostek, A. 1980. 'A quantitative analysis of meteorological anomaly patterns over the United States, 1900–1977', *Mon. Wea. Rev.*, **108,** 615.
Walsh, J. E. and Richman, M. B. 1981. 'Seasonality in the associations between surface temperature over the United States and the North Pacific Ocean'. *Mon. Wea. Rev.*, **109,** 767.
Walsh, J. E., Richman, M. B. and Allen, D. W. 1982. 'Spatial coherence of monthly precipitation in the United States', *Mon. Wea. Rev.*, **110,** 272.
Weickmann, K. M. 1983. 'Intraseasonal circulation and outgoing longwave radiation modes during Northern Hemisphere winter', *Mon. Wea. Rev.*, **111,** 1838.
Wigley, T. M. L., Lough, J. M. and Jones, P. D. 1984. 'Spatial patterns of precipiattion in England and Wales and a revised, homogeneous England and Wales precipitation series', *J. climatol.*, **4,** 1.
Willmott, C. J. 1977. 'A component analytic approach to precipitation regionalization in California', *Archiv. Meteor. Geophys. Biokl., Ser. B,* **24,** 269.
Willmott, C. J. 1978. 'P-mode principal component analysis, grouping and precipitation regions in California', *Archiv. Meteor. Geophys. Biokl., Ser. B,* **26,** 277.
Zwick, W. P. and Velicer, W. F. 1982. 'Factors influencing four rules for determining the number of components to retain', *Multivar. Behav. Res.*, **17,** 253.