

The effect of various methodological options on the detection of leading modes of sea level pressure variability

By RADAN HUTH, *Institute of Atmospheric Physics, Prague, Czech Republic*

(Manuscript received 23 May 2005; in final form 2 September 2005)

ABSTRACT

The effects of several methodological options in the application of principal component analysis (PCA) to gridded data are examined for monthly sea level pressure anomalies over the Northern Hemisphere in the winter half-year. The options include two grid-related ones, viz., the density of the grid and whether and how the uneven areal distribution of gridpoints is compensated for, and two PCA-related ones, viz., the selection of similarity matrix and rotation. The compensation by a cosine weighting and by the use of a quasi-equal-area (QEA) grid has an almost identical effect if covariance matrix is used; the principal components (PCs) based on a regular latitude/longitude grid differ considerably from those based on the QEA grid and cosine-weighted data. The effect of a grid density is small, but recognizable for correlation-based PCs. The PCs derived from a correlation and covariance matrix differ from each other, the difference being considerably smaller for rotated solutions. Unrotated and rotated solutions differ from each other in the degree of simple structure they possess and in their correspondence to correlation maps. The rotated PCs exhibit much stronger simple structure and are more similar to the correlation maps, which suggests that in the interpretation of modes of variability, rotated solutions should be preferred. This implies that in the description of the leading mode of the Northern Hemisphere circulation variability, the sectorial view of the North Atlantic Oscillation should be preferred to the hemispheric view of the Arctic Oscillation.

1. Introduction

One of the most frequent applications of principal component analysis (PCA) in the atmospheric sciences is a detection of modes of atmospheric variability, in particular on intraseasonal and interannual timescales. When using PCA, one is faced with making several methodological decisions, namely, the selection of the similarity matrix (correlation or covariance in most cases), the use of unrotated or rotated principal components (PCs), and the number of PCs to rotate. Some of the methodological options have been broadly discussed in the literature and their effects are relatively well understood, which is the case of the PC rotation issue (e.g. Richman and Lamb, 1985; Richman, 1986; Barnston and Livezey, 1987; O'Lenic and Livezey, 1988; White et al., 1991; Richman et al., 1992; Serrano et al., 1999) and, to a lesser extent, the selection of a similarity matrix (Brinkmann, 1999; Dommenget and Latif, 2002). In many studies, however, little attention seems to be paid to these methodological options, and the decisions taken there are not justified or even not mentioned. This is particularly true for the selection of a similarity matrix

in analyses of the Arctic Oscillation (AO; e.g. Thompson and Wallace, 1998, 2000), where covariance is used without any reasoning or justification and without mentioning that the use of a correlation matrix leads to a dramatically different output (Dommenget and Latif, 2002).

In addition to it, several other options are related to gridded data, in particular, the horizontal resolution, i.e. grid spacing, and whether and how to compensate for an uneven areal distribution of gridpoints on a regular latitude/longitude grid, on which geopotential and sea level pressure (SLP) data are usually defined. Different ways of compensation have been used, but only Araneo and Compagnucci (2004) provide a comparison of some of them.

The goal of this study is to describe the effects that different methodological options related to PCA and grid characteristics have on detected modes of variability in the SLP field on the Northern Hemisphere (NH).

2. Data and methodological options

The database consists of monthly mean SLP fields on a regular 5° by 5° grid extending from 20°N northwards for the winter half-year (November to April) in period 1948–1999. The SLP

Corresponding author.
e-mail: huth@ufa.cas.cz

data are taken from the NCAR (National Center for Atmospheric Research) daily dataset (Trenberth and Paolino, 1980; updated). To quantify a similarity between the loading patterns and of the loading patterns with one-point correlation maps, the correlation coefficient is used, although Richman and Lamb (1985) and Richman (1986) advocate the congruence coefficient (uncentred correlation) in this context. The reasons for our choice is that the values of the correlation and congruence coefficients are very close in this study, the implications are identical regardless of which coefficient is used, and the interpretation of correlations is more straightforward.

The following options relating to the grid and PCA methodology are examined: (i) grid type and weighting, (ii) grid spacing, (iii) similarity matrix and (iv) rotation of PCs.

2.1. Grid type and weighting

It is a well-known fact that the uneven spatial distribution of data points may bias the PCA outputs by exaggerating the influence of data-rich areas (Karl et al., 1982). In hemispheric analyses on a regular latitude/longitude grid, this problem becomes effective in polar areas where meridians converge. To circumvent this, three approaches have been used: (i) the transformation of data onto an equal-area grid, that is, a grid in which each gridpoint represents (at least approximately) the same area (e.g. Clinet and Martin, 1992; Araneo and Compagnucci, 2004), (ii) a rather arbitrary deletion of different numbers of gridpoints at different latitudes to attain their more even distribution (e.g. Barnston and Livezey, 1987) and (iii) a weighting by an appropriate function of cosine of the latitude, which compensates for a decrease of gridbox area towards the pole (e.g. Livezey and Smith, 1999; Thompson and Wallace, 2000). Here we compare the results of PCA on the original regular (latitude-longitude) grid without weighting with cosine-weighted data and a quasi-equal area grid. In cosine weighting, the data are weighted by the square root of cosine of the latitude, which compensates for an unequal area of grid boxes if second-order statistics are computed. A possible disadvantage of cosine weighting is that it neglects the gridpoint at the pole. We decided not to create a true equal-area grid because this would require some kind of spatial interpolation, which could introduce a bias into the data; instead, a quasi-equal-area (QEA) grid was constructed by retaining or deleting individual points of the regular grid. The basic idea of the QEA grid is that each its point should represent as closely as possible the same area as the gridpoint at the pole. At each latitude, we calculate the longitudinal distance between points needed for the area of a gridbox to be equal to that at the pole and the closest corresponding integer number of gridpoints is determined. Table 1 shows it for the equal area grids, corresponding to latitude separations by 5° and 10° . To create a QEA grid, the gridpoints to be retained and to be deleted from the original grid are distributed along the latitude circles as uniformly as possible, and the array of retained and deleted gridpoints is rotated for each latitude separately by a

Table 1. The longitudinal grid spacing required for an equal-area grid and the number of gridpoints at individual latitudes corresponding to it for the 5° and 10° latitudinal gridsteps. The asterisks indicate where the number of gridpoints required is larger than the number of gridpoints available in the original regular grid; the number of gridpoints in the quasi-equal-area grid is set to 72 there

Latitude	5° separation		10° separation	
	Spacing (°)	Number	Spacing (°)	Number
85	45.1	8	—	—
80	22.6	16	45.3	8
75	15.2	24	—	—
70	11.5	31	23.0	16
65	9.3	39	—	—
60	7.9	46	15.7	23
55	6.8	53	—	—
50	6.1	59	12.2	29
45	5.5	65	—	—
40	5.1	70	10.3	35
35	4.8	75*	—	—
30	4.5	79*	9.1	40
25	4.3	83*	—	—
20	4.2	86*	8.4	43

randomly determined number of longitudinal segments, in order to avoid areas with retained and/or deleted points concentrated along meridians. For the QEA grid with a 5° latitudinal spacing, the number of points required exceeds the number of points available in the original grid (72) south of 40°N ; therefore at these latitudes, the number of points in the QEA grid is set to 72 and their latitudinal spacing to 5° . Several realizations of the QEA grid were created by different random rotations of arrays at individual latitudes; the results for them are almost identical, so we present just one realization, shown in Fig. 1.

2.2. Grid density

Results are compared for grids with two different resolutions, 5° and 10° . For the QEA grids, the resolution applies to the latitudinal gridstep; the numbers of gridpoints along longitudinal circles are given in Table 1.

2.3. Similarity matrix

PCAs based on the correlation and covariance matrix lead to different results if individual variables (gridpoint values in our case) have unequal variances. The resulting PCs may differ considerably between the correlation and covariance matrices and even their correspondence to real variability modes may be widely different (Dommenget and Latif, 2002). For SLP, geopotential heights, and similar variables, variance in the mid- and high-latitudes is much larger than in the subtropics and tropics; low latitudes therefore have relatively less weight in covariance-based

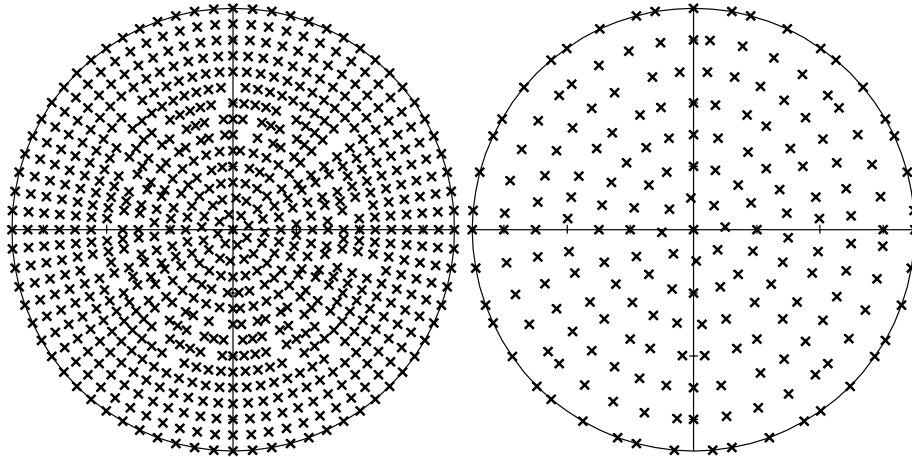


Fig. 1. Quasi-equal-area (QEA) grids with a 5° (left) and 10° (right) spacing used in the study. The points retained in the QEA grid are displayed by crosses.

PCA (Barnston and Livezey, 1987). Here we compare results from the correlation and covariance matrix.

2.4. Rotation

A better physical interpretation of PCs can frequently be achieved by their rotation, either orthogonal or oblique. Many studies of low-frequency variability modes utilize rotated PCA because it leads to more regionalized loading patterns (e.g. Barnston and Livezey, 1987; O'Lenic and Livezey, 1988). On the other hand, the concept of the AO is based on an unrotated analysis. We compare unrotated solutions with those rotated orthogonally by the Varimax criterion. The oblique rotation was also examined, but it did not bring any improvement over the orthogonal rotation in terms of simple structure (see below), and therefore is not discussed here. The numbers of PCs to rotate were determined from the eigenvalue versus PC number plots, where the appropriate choices are those at the ends of shelves (i.e. parts with a small slope, followed by a pronounced drop; O'Lenic and Livezey, 1988). For each PCA, rotations for several numbers of PCs, complying with this criterion, were performed.

The PCs are displayed as maps of their projections onto the regular 5° × 5° latitude–longitude grid; that is, for the correlation/covariance matrix, the correlations/covariances of the PCs with the time series at each gridpoint are mapped. The display of projections is preferred to that of PC loadings because the former allows a direct comparison between different methodological options.

3. Effects of grid type, weighting and density

3.1. Grid type and weighting

First, the effect of compensation for unequal gridbox size is examined: the results for cosine-weighted data and the QEA

grid are compared with the original, unweighted data on a regular latitude/longitude grid. The comparison is carried out for the grid with finer spacing, corresponding to the 5° resolution. For PCA of correlation matrix, any weighting of data has naturally no effect as correlation removes scale from the variables. Nevertheless, Livezey and Smith (1999) propose and apply a procedure allowing a weighting in correlation-based PCA: data are first standardized, then weighted and finally subjected to covariance-based PCA. The effects of grid type and weighting on the PC loading patterns are illustrated in Figs. 2 and 3 for the unrotated and rotated analysis, respectively, of the covariance matrix.

The QEA grid and the cosine-weighted data yield almost identical PC loading patterns. The pattern correlations between the corresponding PC projections for the 15 leading unrotated PCs all exceed 0.976. For rotated solutions, the pattern correlations are only slightly lower, e.g. for nine rotated PCs, they are between 0.939 and 0.996. Only tiny and unimportant differences can be recognized visually (compare two upper rows in Fig. 3). This indicates that for PCA based on covariance matrix, the weighting by the square root of cosine of the latitude has virtually the same effect as a transformation of data onto a QEA grid, and that the neglect of the polar gridpoint in the former does not affect the results.

The PCs for the QEA and regular grid differ much more from each other in several ways (see Fig. 2 and lower two rows in Fig. 3). First of all, the plausible numbers of PCs to retain for rotation are different. Looking at the scree plot in Fig. 4, one can identify several numbers of PCs to rotate: 6, 9, 11 and 14 for the QEA grid, while 4, 8, 10 and 14 for the regular grid. In addition to this difference, the eight leading PCs for the regular grid explain slightly more variance (78.4%) than the nine leading PCs for the QEA grid (76.1%). The difference in the plausible numbers of PCs to rotate makes it impossible to compare solutions with the same number of PCs. In the following, we therefore discuss the

COVARIANCE - UNROTATED

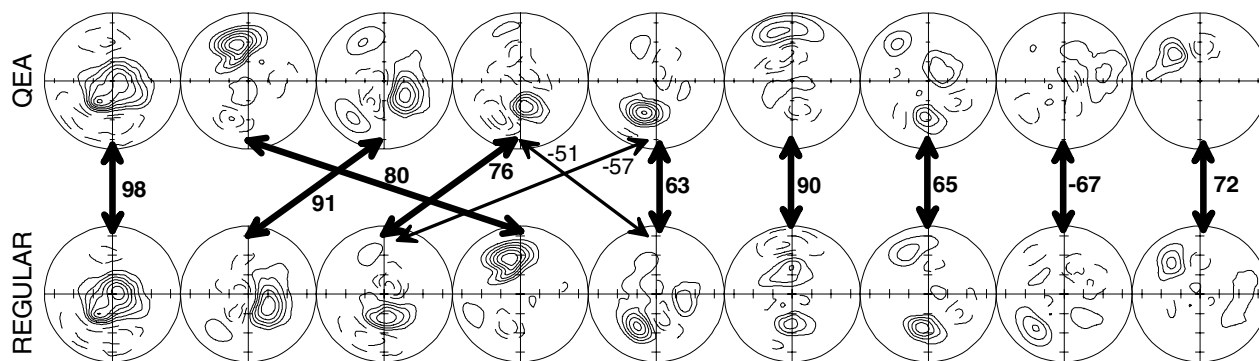


Fig. 2. Projections of loadings for nine leading PCs for unrotated PCA of the covariance matrix: for the QEA grid (top) and regular grid (bottom). Contour interval is 1 hPa, zero line is omitted and negative contours are dashed. On the maps, the Greenwich meridian points downwards, 90°E to the right, etc. The pattern correlations exceeding 0.5 are denoted by arrows with the correlation (100×) shown; the bold arrows indicate a one-to-one correspondence between the patterns, irrespective of the magnitude of correlation.

COVARIANCE - ROTATED

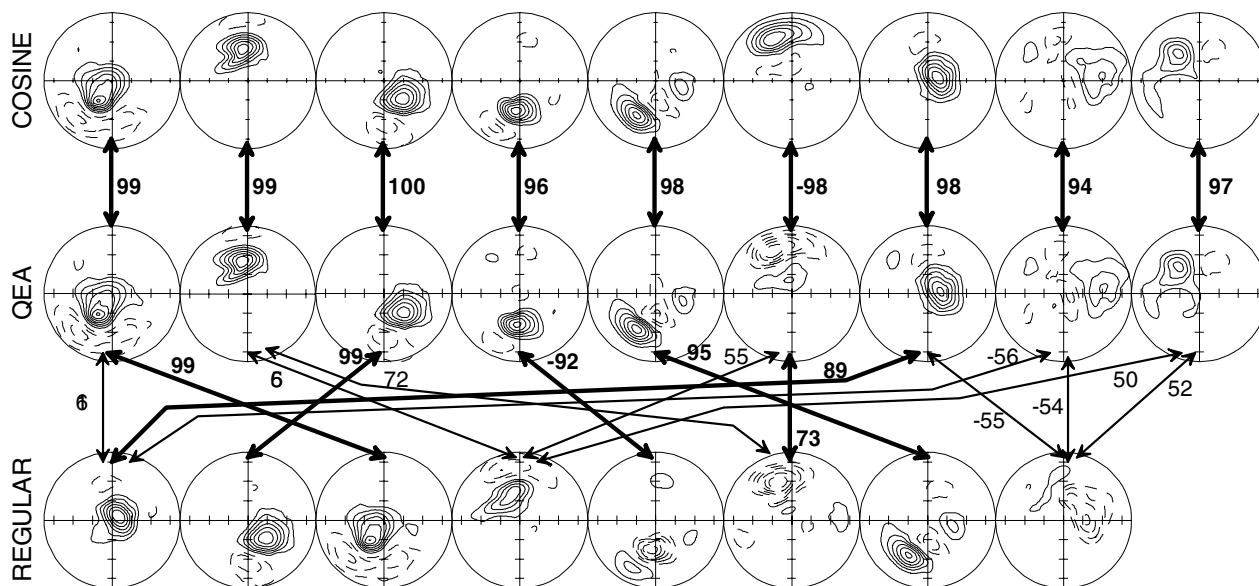


Fig. 3. Projections of loadings for rotated PCA of the covariance matrix: cosine weighting, nine rotated PCs (top); QEA grid, nine rotated PCs (middle); regular grid, eight rotated PCs (bottom). For the explanation of arrows, see Fig. 2.

rotated solution with nine PCs for the QEA grid and that with eight PCs for the regular grid.

In comparisons of two sets of PCs, it is useful to introduce the concept of a 'one-to-one correspondence'. The one-to-one correspondence occurs if the maximum correlation for each PC from one set is also a maximum for any PC from the other set, and vice versa. In general, there is a one-to-one correspondence for all PC loading patterns in the unrotated solution and for their majority in the rotated solution (this holds for the correlation matrix as well – not shown). Correlations for some pairs of patterns are high, exceeding 0.9 and indicating an excellent match (e.g. PC 1/1 in the unrotated solution, and PCs 1/3, 3/2, 4/5 and 5/7 in the

rotated solution), but the similarity of other pairs is rather limited, which corresponds to relatively weak correlations, as low as 0.63. There are several correlations 'across' pairs of matching PCs that exceed 0.5, which indicates blending between the PCs. This means that a part of a signal contained in one PC (e.g. unrotated QEA PC 5) does not appear in its counterpart (unrotated regular PC 5) but in a different pattern (in this example, in unrotated regular PC 3). The 'across pairs' correlations are more numerous in the rotated solution. There are two reasons for that: the number of PCs is not equal, so at least one PC cannot be paired, and the rotated PC loadings in a single solution can be mutually correlated by definition.

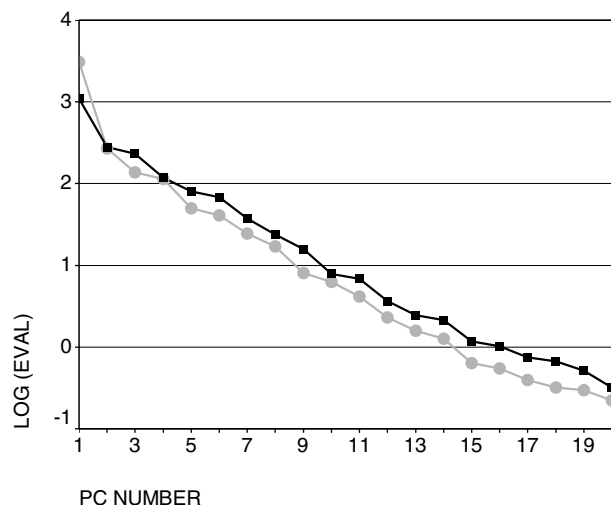


Fig. 4. Dependence of the logarithm of eigenvalue on the PC number for the QEA grid (black squares) and regular grid (grey circles), all for the covariance matrix.

Another notable fact is that the order of some PCs is changed, which means that some patterns tend to explain different amounts of variance in the two grids. This concerns particularly the patterns with action centres in high latitudes and around the pole, which tend to be relatively more important on the regular grid because of an excessive density of gridpoints there, and vice versa. This is the case, e.g. of QEA PCs 3 and 4 in the unrotated solution and, more pronounced, of QEA PC 7 in the rotated solution, which becomes the leading mode for the regular grid. This effect has recently been described by Araneo and Compagnucci (2004) in their analysis of stratospheric temperatures.

Because of the differences between the QEA and regular grids, and since the use of the QEA grid is more appropriate as it eliminates the undesirable convergence of gridpoints towards the pole, the further discussion is confined to the QEA grid.

Table 3. Percentage of variance explained by the leading 12 PCs for the unrotated solution of the correlation-based analysis on the coarse and fine grid

PC	Fine	Coarse
1	15.30	15.28
2	9.79	10.16
3	8.07	7.56
4	7.27	6.97
5	5.11	4.91
6	4.74	4.68
7	4.29	4.24
8	4.12	4.01
9	3.78	3.76
10	3.26	3.17
11	2.89	2.92
12	2.31	2.26

3.2. Grid density

Now, let us proceed to a comparison between the QEA grids corresponding to the 5° and 10° resolution along a longitude circle. In Table 2, correlation coefficients are shown for the nine leading PCs of the unrotated and rotated (with nine PCs retained) solutions for the correlation and covariance matrix. There is almost an excellent match for the covariance-based PCs, the correlation coefficients between the patterns both on the fine and coarse grid all exceeding 0.98. The alternation of the order between the PC 4 and PC 5 in the rotated solution is of no relevance given the virtual identity of the corresponding patterns.

The correspondence between the PCs on the fine and coarse grid is somewhat worse for the correlation-based analysis. In the unrotated solution, PCs 7 and 8 are blended. This seems to be mainly a consequence of the relevant eigenvalues being fairly close to each other (see Table 3), which implies that the PCs are

Table 2. Pattern correlations (100×) between the fine and coarse grid for different PCA solutions. Shown are only correlations exceeding 0.5; the number of the corresponding PC on a coarse grid is shown in parentheses

PC number on fine grid	Correlation, unrotated	Correlation, rotated	Covariance, unrotated	covariance, rotated
1	100 (1)	99 (1)	100 (1)	100 (1)
2	94 (2)	97 (2)	100 (2)	98 (2)
3	91 (3)	98 (3)	100 (3)	100 (3)
4	91 (4)	97 (4)	100 (4)	−100 (5)
5	99 (5)	95 (5)	99 (5)	99 (4)
6	97 (6)	97 (8)	98 (6)	100 (6)
7	−77 (7); 58 (8)	−98 (7)	100 (7)	99 (7)
8	−61 (7); −79 (8)	82 (6)	100 (8)	99 (8)
9	96 (9)	88 (9)	100 (9)	99 (9)

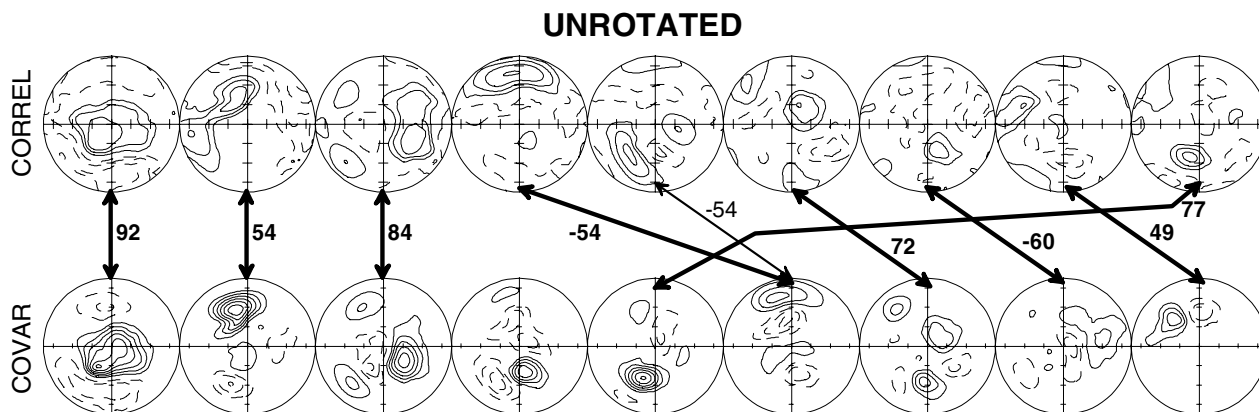


Fig. 5. Projections of loadings for nine leading PCs for unrotated PCA on the fine QEA grid: correlation matrix (top) and covariance matrix (bottom). For the explanation of arrows, see Fig. 2. For the correlation matrix, contour interval is 0.2, otherwise as for the covariance matrix.

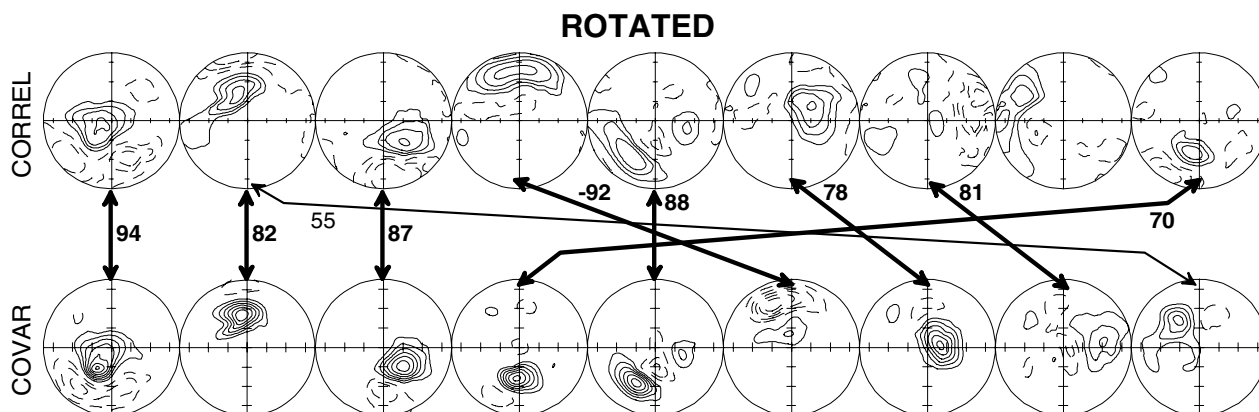


Fig. 6. As in Fig. 5, except for rotated PCA (nine PCs retained for rotation).

not well separated and the signal may be mixed in them in an arbitrary way (North et al., 1982). In the rotated solution, no mixing between successive PCs occurs (which is one of the reasons why rotation is performed, indeed). Nevertheless, the correspondence for two pairs of patterns drops below 0.9 of the correlation coefficient, and the order of PCs 6 and 8 is interchanged. This indicates that the selection of the grid density may have a small, but non-negligible effect on the correlation-based PCs, whereas the covariance-based PCs seem to be almost unaffected by this choice. However, this is likely to be dependent on the dataset used and cannot be generalized. For the sake of simplicity, further discussions are confined to the fine grid.

4. Effects of PCA-related options

4.1. Similarity matrix

For the unrotated solutions, the similarity between the PCs of the correlation and covariance matrix is fairly weak (Fig. 5). With the exception of PC1, the correlations between the corresponding patterns are all below 0.8. For the PC pair 8/9, the correlation is as low as 0.49, the patterns exhibiting very little resemblance.

Two correlation-based modes (PCs 4 and 5) are blended to form covariance-based PC 6, which however has a spatial structure rather different from either of its counterparts. One covariance-based PC (PC 4) lacks its correlation-based counterpart. The one-to-one correspondence almost disappears for higher order PCs (not shown).

The rotated solutions (Fig. 6 shows projections of loading patterns for nine rotated PCs) display a better correspondence between the correlation and covariance-based PCs, although this correspondence is far from perfect. The correlations are higher in general, although a one-to-one correspondence is not achieved either: correlation-based PC 8 is not correlated with any covariance-based PC by more than 0.4 and covariance-based PC 9 is linked to correlation-based PC 2 (by correlation coefficient of 0.55), which is, however, strongly coupled to covariance-based PC 2. The correspondence between the correlation and covariance-based PCs is even slightly better for 11 rotated PCs (only one PC of each set of patterns lacking its counterpart, and the correlation coefficients between corresponding patterns ranging from 0.83 to 0.94; not shown). For all reasonable numbers of PCs retained (up to 15), the rotated patterns exhibit a better correspondence than the unrotated ones (not shown).

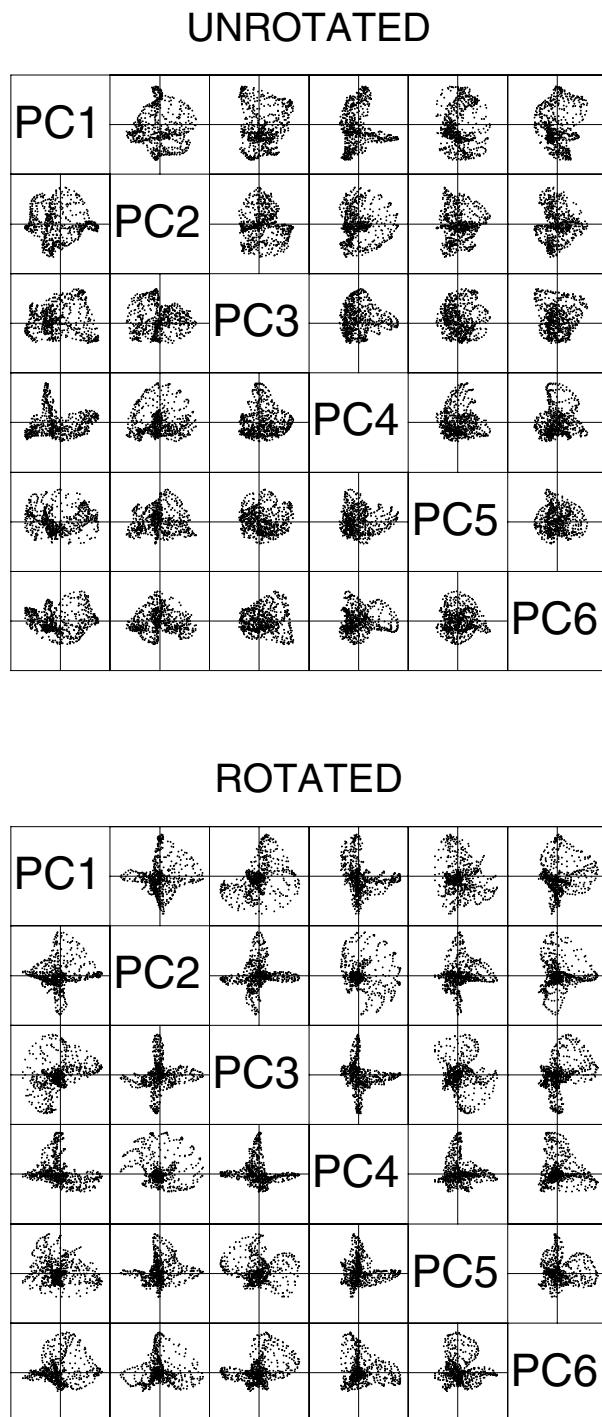


Fig. 7. Scatterplots for six leading PCs for the correlation matrix: unrotated (top) and orthogonally rotated (bottom) solution. The axes in diagrams range from -1 to 1 , with zero indicated by a line.

4.2. Rotation of PCs

In this subsection, the PC loadings are compared between the unrotated and rotated solutions. We apply two criteria helping to decide whether the unrotated or rotated solution should be

preferred in interpretations of atmospheric circulation variability. These criteria are the simple structure and comparisons of PC loadings with correlation maps.

The 'simple structure' is the key concept in the interpretation of PCs. The PCs can be reasonably interpreted in terms of the original variables, only if the information from each variable is concentrated into (i.e. each variable is loaded on) as few PCs as possible. Such a state is called a 'simple structure'. In other words, simple structure requires that the majority of PC loadings be near zero. The degree of simple structure can be estimated from the pairwise PC scatterplots: in the presence of a strong simple structure, data points are aligned closely along the PC axes (e.g. Lyons and Bonell, 1994). More information about the concept of simple structure can be found in, for example, Richman (1986) and in textbooks on PCA and factor analysis, e.g. Reymont and Jöreskog (1996).

The scatterplots for the six leading PCs of the correlation matrix, both for the unrotated and orthogonally rotated (with nine PCs retained) solutions, are shown in Fig. 7. The scatterplots for the covariance matrix yield analogous implications and, therefore, are not presented here. There is only a weak simple structure in the unrotated PCs: the points are spread far from the axes and only little alignment along them can be observed. Moreover, several plots (e.g. PC1 vs. PC4 and PC1 vs. PC5) show signatures of alignment along lines tilted to the PC axes, thus suggesting a need for a rotation. On the other hand, the simple structure in the orthogonally rotated PCs appears to be considerably stronger. For some pairs of PCs, the alignment along the axes is almost perfect (e.g. PC2 vs. PC3, PC3 vs. PC4, PC4 vs. PC5), and also for the others, it is without doubt better than for the unrotated solution. The simple structure may be quantified by counting the numbers of data points whose loadings are distant from the axes in the scatterplots. The mean number of data points in a scatterplot with both loadings being greater than 0.2 is twice as large for the unrotated than rotated PCs: there are on average 145 (72) such data points in one scatterplot for the unrotated (rotated) solution. This result is insensitive to the correlation threshold. It is natural that the rotated solution possesses a stronger simple structure because rotation procedures are designed so as to increase it relative to the original unrotated solution. The well-known fact that rotated PC loadings are more regionalized, which can be easily seen if one compares corresponding maps in Figs. 5 and 6, is a consequence of a better simple structure present in the rotated PCs.

If PCs represent a real, physically interpretable signal, their loading maps should resemble the correlation/covariance structure they were derived from. Therefore, correlation maps were calculated for the base points located at the centres of both the unrotated and rotated loading patterns. Figure 8 in its top two rows displays a set of five leading loading patterns for the unrotated and rotated solutions of the correlation matrix. Below them, correlation maps for some of their action centres are displayed to which the correlation coefficients with the unrotated

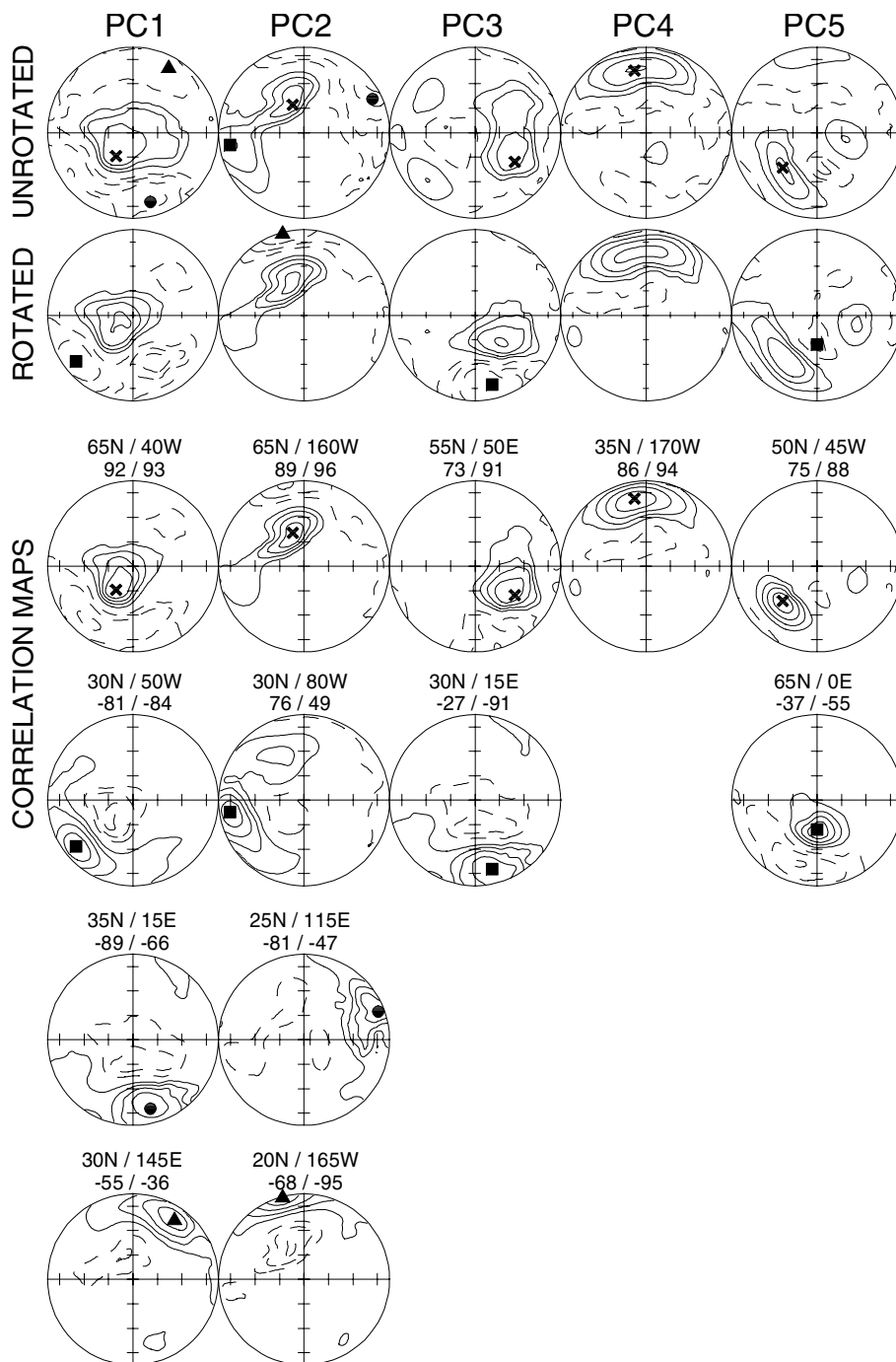


Fig. 8. Top two rows: unrotated and rotated loadings for five leading PCs of the correlation matrix. Below: one-point correlation maps for the base-points located at the action centres of the loading patterns. The base points are identified by their geographical coordinates and by a symbol (cross, square, circle, triangle) plotted also on the corresponding loading map. The numbers attached to the correlation maps are pattern correlation coefficients (100 $\times$) between the correlation map, and the unrotated (left) and rotated (right) loadings.

and rotated loadings are attached. The general impression is that the correlation maps more resemble the rotated than the unrotated loadings. This is in agreement with Barnston and Livezey (1987), who conducted a similar analysis. For example, in PC1

[corresponding to the North Atlantic Oscillation (NAO) and AO, respectively], the mid-latitude centres of the unrotated loading (denoted by a circle and triangle) are in fact mutually uncorrelated. The unrotated PC2 loading possesses two distant centres

of the same sign over North America (cross and square), which are, however, also almost uncorrelated. Its negative centre over East Asia (circle) has only a very weak connection with them. On the other hand, the subtropical Pacific centre of the rotated pattern (which is only weakly pronounced in the unrotated pattern; triangle) exhibits a strong correlation with the North American centre of PC2. In PC4, the correlations with the Pacific centre do not extend to the European and Asian sectors, as is suggested by the unrotated loadings, but are limited to the North Pacific and North America in accordance with the rotated loading pattern.

We demonstrate that the rotated PCs possess much stronger a simple structure, and their loadings better resemble the correlation structures they were derived from. There are two more reasons for preferring rotated to unrotated PCA, which are of a more general applicability: rotated PCs possess (i) a smaller sensitivity to the selection of the analysis domain (Richman, 1986) and (ii) reduced sampling error (Barnston and Livezey, 1987). Therefore, the rotated PCs should be given preference to the unrotated ones in the interpretation of results of PCA.

5. Conclusions

We have examined the effects the methodological options have on the outputs of PCA of monthly SLP anomalies over the NH, with the following results.

(1) If PCA is calculated from the covariance matrix, there is virtually no difference between the two ways of compensation for a decreasing area of gridboxes in a latitude/longitude grid towards the pole, namely, between the cosine weighting and the QEA grid.

(2) There is a considerable difference between the regular latitude/longitude grid and the one compensating for an unequal gridbox area. Mixing of PCs and a lack of one-to-one correspondence between the PC sets are observed. The differences are more pronounced for unrotated solutions.

(3) The grid density has a small, but non-negligible effect on the correlation-based PCs, whereas the covariance-based PCs seem to be almost unaffected by it.

(4) The PCs differ considerably between the covariance and correlation matrices for the unrotated solutions; the differences are much smaller for the rotated solutions.

(5) The unrotated PCs possess much weaker a simple structure; in contrast, the simple structure in the rotated PCs is fairly strong. The correlation maps with the base points located at the action centres of the loading patterns correspond much better to the rotated than unrotated loadings. This implies that the interpretation of unrotated PCs is doubtful and the interpretation of PCA results in terms of modes of variability should be based on rotated PCs.

The latter finding has implications for the debate on whether the AO or NAO should be considered the leading modes of the NH SLP variability (e.g. Wallace, 2000; Ambaum et al., 2001;

Wallace and Thompson, 2002). Since the AO appears in the unrotated solutions only, whereas the rotated solutions result in the NAO, it is the NAO that should be preferred to the AO in interpreting the NH circulation variability. This study gives a warning that the AO may be an artefact of the analysis method rather than a real variability mode.

6. Acknowledgments

The study was supported by the Grant Agency of the Czech Republic, project 205/05/2282.

References

- Ambaum, M. H. P., Hoskins, B. J. and Stephenson, D. B. 2001. Arctic Oscillation or North Atlantic Oscillation? *J. Clim.* **14**, 3495–3507.
- Araneo, D. C. and Compagnucci, R. H. 2004. Removal of systematic biases in S-mode principal components arising from unequal grid spacing. *J. Clim.* **17**, 394–400.
- Barnston, A. G. and Livezey, R. E. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Weather Rev.* **115**, 1083–1126.
- Brinkmann, W. A. R. 1999. Application of non-hierarchically clustered circulation components to surface weather conditions: Lake Superior basin winter temperatures. *Theor. Appl. Climatol.* **63**, 41–56.
- Clinet, S. and Martin, S. 1992. 700-hPa geopotential height anomalies from a statistical analysis of the French Hemis data set. *Int. J. Climatol.* **12**, 229–256.
- Dommenget, D. and Latif, M. 2002. A cautionary note on the interpretation of EOFs. *J. Clim.* **15**, 216–225.
- 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. Meteorol.* **21**, 1183–1186.
- Livezey, R. E. and Smith, T. M. 1999. Considerations for use of the Barnett and Preisendorfer (1987) algorithm for canonical correlation analysis of climate variations. *J. Clim.* **12**, 303–305.
- Lyons, W. F. and Bonell, M. 1994. Regionalization of daily mesoscale rainfall in the tropical wet/dry climate of the Townsville area of north-east Queensland during the 1988–1989 wet season. *Int. J. Climatol.* **14**, 135–163.
- North, G. R., Bell, T. C., Cahalan, R. F. and Moeng, F. J. 1982. Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.* **110**, 699–706.
- O'Lenic, E. A. and Livezey, R. E. 1988. Practical considerations in the use of rotated principal component analysis (RPCA) in diagnostic studies of upper-air height fields. *Mon. Weather Rev.* **116**, 1682–1689.
- Reyment, R. A. and Jöreskog, K. G. 1996. *Applied Factor Analysis in the Natural Sciences*. Cambridge University Press, Cambridge, 371.
- Richman, M. B. 1986. Rotation of principal components. *J. Climatol.* **6**, 293–335.
- Richman, M. B. and Lamb, P. J. 1985. Climatic pattern analysis of three- and seven-day summer rainfall in the Central United States: Some methodological considerations and a regionalization. *J. Clim. Appl. Meteorol.* **24**, 1325–1343.

- Richman, M. B., Angel, J. R. and Gong, X. 1992. Determination of dimensionality in eigenanalysis. In: *Fifth Int. Meeting on Statist. Climatol.* Toronto, Canada, 229–235.
- Serrano, A., García, J. A., Mateos, V. L., Cancillo, M. L. and Garrido, J. 1999. Monthly modes of variation of precipitation over the Iberian Peninsula. *J. Climate* **12**, 2894–2919.
- Thompson, D. W. J. and Wallace, J. M. 1998. The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.* **25**, 1297–1300.
- Thompson, D. W. J. and Wallace, J. M. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Clim.* **13**, 1000–1016.
- Trenberth, K. E. and Paolino, D. A. jr. 1980. The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Mon. Weather Rev.* **108**, 855–872.
- Wallace, J. M. 2000. North Atlantic Oscillation/annular mode: Two paradigms – one phenomenon. *Q. J. R. Meteorol. Soc.* **126**, 791–805.
- Wallace, J. M. and Thompson, D. W. J. 2002. The Pacific center of action of the Northern Hemisphere annular mode: Real or artifact? *J. Clim.* **15**, 1987–1991.
- White, D., Richman, M. and Yarnal, B. 1991. Climate regionalization and rotation of principal components. *Int. J. Climatol.* **11**, 1–25.