North Atlantic climate variability from a self-organizing map perspective

David B. Reusch,[1] Richard B. Alley,[2] and Bruce C. Hewitson[3]

Received 2 May 2006; revised 31 July 2006; accepted 21 September 2006; published 20 January 2007.

[1] North Atlantic variability in general, and the North Atlantic Oscillation (NAO) in particular, is a long-studied, very important but still not well-understood problem in climatology. The recent trend to a higher wintertime NAO index was accompanied by an additional increase in the Azores High not coupled to changes in the Icelandic Low, as shown by a self-organizing maps (SOMs) analysis of monthly mean DJF mean sea level pressure data from 1957 to 2002. SOMs are a nonlinear tool to optimally extract a user-specified number of patterns or icons from an input data set and to uniquely relate any input data field to an icon, allowing analyses of occurrence frequencies and transitions complementary to principal component analysis (PCA). SOMs analysis of ERA-40 data finds a North Atlantic ''monopole'' roughly colocated with the mean position of the Azores High, as well as the well-known NAO dipole involving the Icelandic Low and the subtropical high. Little trend is shown in December, but the Azores High increased along with the NAO in January and February over the study interval, with implications for storminess in northwestern Europe. In short, our SOM-based analyses of winter MSLP have both confirmed prior knowledge and expanded it through the relative ease of use and power with nonlinear systems of the SOM-based approach to climatological analysis.

Citation: Reusch, D. B., R. B. Alley, and B. C. Hewitson (2007), North Atlantic climate variability from a self-organizing map perspective, *J. Geophys. Res.*, *112*, D02104, doi:10.1029/2006JD007460.

1. Introduction

[2] Synoptic climatology has developed many techniques useful for extracting patterns from large data sets such as daily mean sea level pressures in some region [e.g., *Barry and Carleton*, 2001]. It may prove useful and informative, for example, to note that the data field for a particular day has strong loadings on the first two principal components of the data set. A human observer wishing to summarize such a data set might also notice similarities between particular fields and historical events. That is, it may be instructive to note that a pattern is very similar to the (hypothetical) ''Great Blizzard of '95,'' which serves as an icon to which other data fields can be compared. However, particular data fields are unlikely to be unique enough to be an optimal icon (extreme events are by definition exceptional). Self-organizing map (SOM [*Kohonen*, 2001]) analysis provides an objective way to optimally extract a user-specified number of icons or SOM states from an input data set, such that each element of the data set is uniquely most similar to one of the SOM states, and the total mismatch between elements of the data set and the SOM states is minimized. The analysis returns the icons or SOM states in a grid or ''map,'' with similar states near each other and the most extreme states at the corners. Often, the states at the ends of one diagonal are similar to the positive and negative phases of the first principal component of the input data, with the second principal component correspondingly at the ends of the other diagonal; however, this is not required and rarely exact. Because each input data element maps uniquely to one SOM state, SOM analysis easily allows characterization of time trends in frequency of occurrence, preferred transitions, and hysteresis of preferred patterns.

[3] North Atlantic climate variability arises from diverse sources over broad spatial and temporal scales [*Marshall et al.*, 2001] and has been a rich field of study for many years. Despite decades of research and identification of many important properties and behaviors, additional important questions remain unanswered. Some of these questions arise from the limitations of the linear tools generally used in climate studies. The nonlinear aspects known to exist in the climate system are thus often only approximated, or sometimes ignored altogether, in the purely linear approach, to the detriment of our understanding. New tools with the ability to handle nonlinear behavior are thus potentially of great value to our study of climate.

[4] The North Atlantic Oscillation (NAO) is an area of particular interest to the study of North Atlantic variability [e.g., *Wallace and Gutzler*, 1981; *Rogers*, 1984; *Hurrell*, 1995; *Appenzeller et al.*, 1998; *Cullen et al.*, 2001; *Ogi et*

[1]Earth and Environmental Systems Institute, College of Earth and Mineral Sciences, Pennsylvania State University, University Park, Pennsylvania, USA.
[2]Department of Geosciences and Earth and Environmental Systems Institute, College of Earth and Mineral Sciences, Pennsylvania State University, University Park, Pennsylvania, USA.
[3]Department of Environmental and Geographical Sciences, University of Cape Town, Rondebosch, South Africa.

0148-0227/07/2006JD007460

al., 2003; *Cassou et al.*, 2004]. As the only year-round teleconnection pattern in the northern hemisphere [*Barnston and Livezey*, 1987], the NAO has a widespread climatic influence from eastern North America to western Europe (and beyond) and has been widely studied in recent decades. The large-scale alternation of atmospheric mass between surface pressure centers near the Azores (high) and Iceland (low) alters circulation patterns in the North Atlantic on multiple timescales. Modified storm tracks and temperature and precipitation patterns result throughout the region (along with numerous other effects on the climate system [e.g., *Wanner et al.*, 2001]). Given the nonlinearities known to exist in the climate system in general [e.g., *Stocker*, 1999], and the NAO in particular [*Lamb and Peppler*, 1987], it is essential to explore nonlinear analysis techniques so as to better understand these aspects.

[5] Self-organizing maps [*Kohonen*, 2001] are a tool from the field of artificial neural networks with characteristics well suited to such problems. SOMs provide an alternative to more traditional techniques, such as principal component analysis (PCA), that is less complex, more robust and less subjective while also accommodating nonlinear relationships in the data. They also provide a powerful visualization approach for studying structure in large, complex data sets. A SOM analysis produces a set of generalized patterns from the input data that, taken together, describe the multidimensional distribution function of the data [*Kohonen*, 2001; *Hewitson and Crane*, 2002]. In the case of atmospheric circulation data, such as in this study, these patterns capture the full range of synoptic conditions while also treating the data as a continuum, unlike, for example, cluster analysis. SOM-based analysis has been applied to such problems as climate downscaling [*Crane and Hewitson*, 1998], climate classification [*Cavazos*, 1999, 2000; *Cavazos et al.*, 2002; *Cassano et al.*, 2006a], climate changes over time [*Hewitson and Crane*, 2002; *Cassano et al.*, 2006b; *Lynch et al.*, 2006], area-average regridding of precipitation data [*Hewitson and Crane*, 2005], physical oceanography [*Liu and Weisberg*, 2005; *Liu et al.*, 2006a], and reconstruction of past climate from ice core-based climate proxies [*Reusch et al.*, 2005b]. As discussed further below (and by *Reusch et al.* [2005a] and *Liu et al.* [2006b]), SOM-based analysis differs from more traditional linear analysis in a number of ways that provide additional power over nonlinear data sets. SOM-based analysis thus complements linear techniques without replacing them.

[6] Given the long history of NAO studies, we are clearly not the first to study this problem or to apply novel analysis techniques to relevant data sets. For example, *Hurrell et al.* [2003], *Thompson et al.* [2003] and *Wanner et al.* [2001] each provide thorough reviews of prior work and the open questions. Among the latter is whether the NAO is truly distinct from the Arctic Oscillation (AO, also known as the Northern Hemisphere Annular Mode, or NAM) or whether it is an artifact of the analysis used. Numerous studies have addressed this particular aspect [e.g., *Deser*, 2000; *Thompson and Wallace*, 2000, 2001; *Wallace*, 2000] but without unambiguously defining the answer. This ambiguity is just one aspect of the larger problem of identifying stable weather regimes and circulation patterns. It is beyond the scope of this work to comprehensively review prior work on these topics and we refer the reader to the literature for more details including other (nonlinear) approaches [e.g., *Michelangeli et al.*, 1995; *Cassou et al.*, 2004] to studying Northern Hemisphere variability.

[7] The next two sections describe our data and methodologies and present results from the SOM and PCA analyses. We follow with a discussion of temporal variability in the generalized patterns, the additional benefits of the SOM patterns versus just using an NAO index, and a brief comparison of SOM-based versus PCA-based NAO analysis.

2. Data and Methodology

2.1. Gridded Meteorological Data

[8] The European Centre for Medium-Range Weather Forecasts' (ECMWF) reanalysis ERA-40 is our source for GCM-scale meteorological data at 2.5° horizontal resolution for the period 1957–2002 [*Gibson et al.*, 1999]. The spatial domain (20–85°N, 85°W to 25°E, Figure 1) was selected to capture both action centers of the NAO and adjacent areas known to feel its effects (e.g., Scandinavia). Mean sea level pressure (MSLP) was extracted for our spatial domain, regridded, averaged and standardized for the SOM-based analyses reported here. Ongoing and future work includes other surface and upper air variables (e.g., 500 mbar geopotential height) to examine further aspects of atmospheric variability in this region. Resampling the ERA-40 2.5° data to an equal area, 819 point grid (a 250 km version of the National Snow and Ice Data Center EASE-Grid [*Armstrong and Brodzik*, 1995]) was done to provide uniform spatial sampling density over the full domain. Monthly means and standard deviations were calculated from the 6-hourly regridded data. As the last step before SOM analysis, the data were normalized to the 1971–2000 baseline mean and standard deviation. This step is not essential when using just one variable, but such normalization aids future comparison with other variables that may have widely different mean values and variances (and puts all data into units of standard deviations, which simplifies visual tests of statistical significance). It is also useful to use a baseline widely used in the research community (i.e., 1971–2000). To highlight differences from the mean state in the SOM results, anomalies were calculated at each grid point by subtracting the grid point mean over the full series. Because the NAO has its strongest expression during boreal winter, we have focused

Figure 1. Generalized SOM patterns for (a) monthly mean of MSLP and (b) monthly standard deviation of MSLP. To the upper left of each pattern is the pattern index number. To the upper right of each pattern in Figure 1a is the frequency of occurrence (%) of the pattern in the full (DJF) 135 month data set (1957–2002). This number is in bold (italic) if it is at least 50% more (4.95%) or less (1.65%) than the expected value (3.3%) from a uniform distribution of the 135 months on the 30 SOM patterns. Map boundaries are 20 to 85°N and 80°W to 25°E. Grid spacing is 15° north-south and 30° east-west. Here and in subsequent figures with SOM patterns, the values are shown as grid point anomalies (the mean over the 30 patterns has been subtracted from each grid point) to better show the differences between patterns. Data were initially standardized to the 1971–2000 baseline mean and standard deviation.

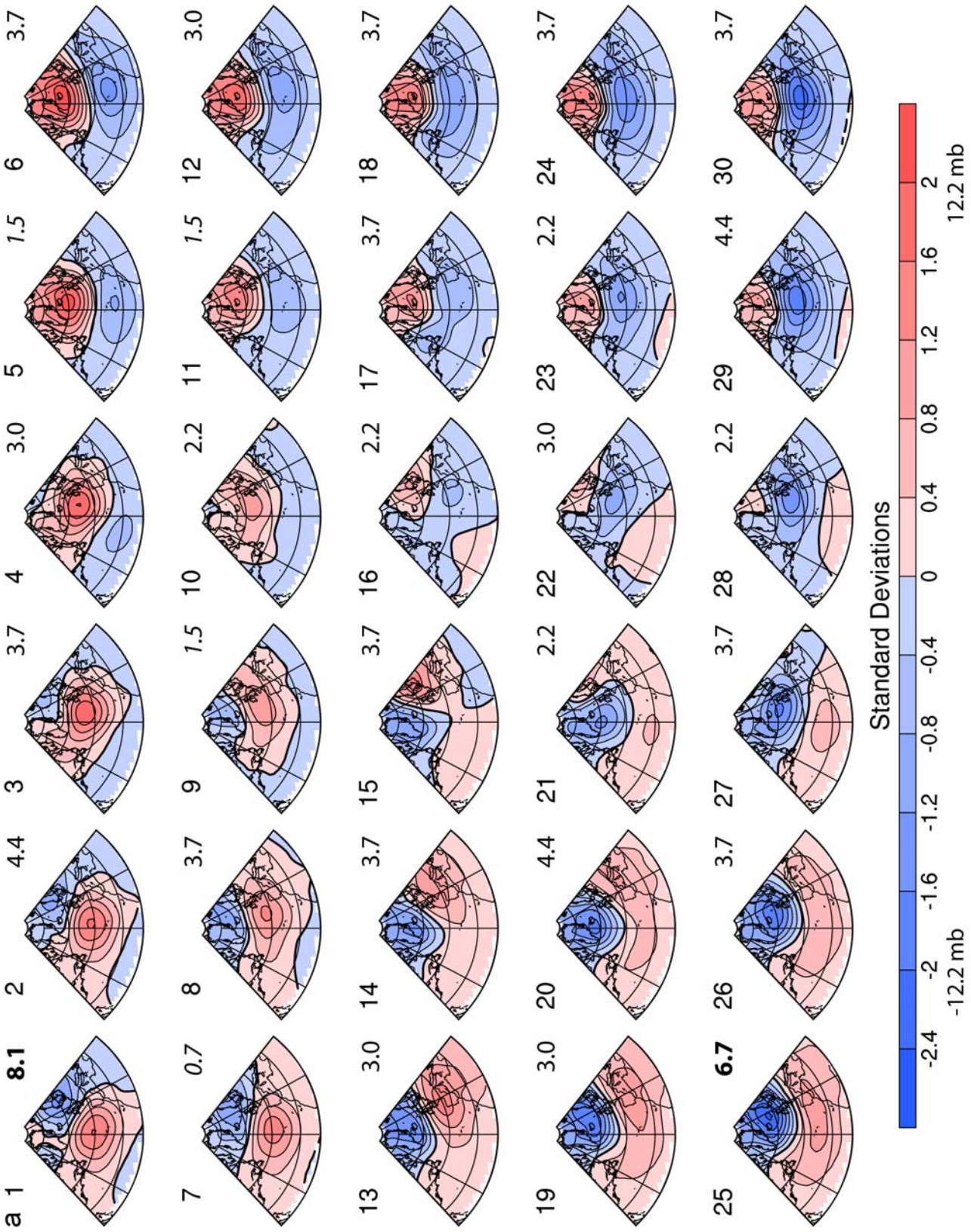
a 1
8.1
2
4.4
3
3.7
4
3.0
5
1.5
6
3.7
7
0.7
8
3.7
9
1.5
10
2.2
11
1.5
12
3.0
13
3.0
14
3.7
15
3.7
16
2.2
17
3.7
18
3.7
19
3.0
20
4.4
21
2.2
22
3.0
23
2.2
24
3.7
25
6.7
26
3.7
27
3.7
28
2.2
29
4.4
30
3.7
Standard Deviations
-2.4
-2
-1.6
-1.2
-0.8
-0.4
0
0.4
0.8
1.2
1.6
2
-12.2 mb
12.2 mb

Figure 1

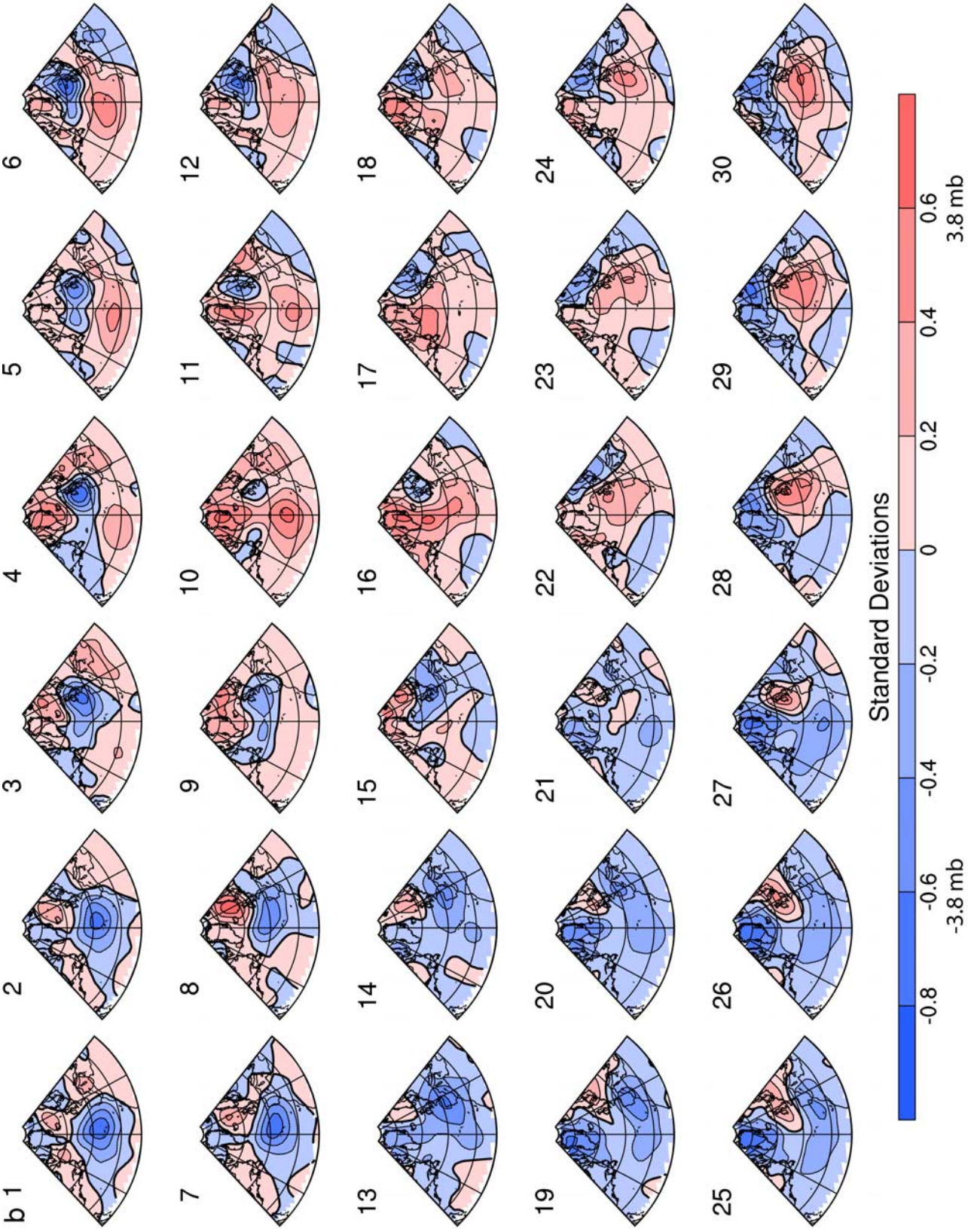

Figure 1. (continued)

exclusively on the months of December, January and February (DJF). As such, the MSLP data consist of a total of 135 months (each defined at 819 grid points) from 1957 to 2002.

2.2. Self-Organizing Maps

[9] Self-organizing maps [*Kohonen*, 2001] are an analysis tool from the field of artificial neural networks. SOMs support analysis of variability in large, multivariate and/or multidimensional data sets through the creation of a spatially organized set of generalized patterns of variability from the data. The technique provides a complementary, nonlinear alternative to more frequently used but linear tools such as principal component analysis (PCA). (See *Reusch et al.* [2005a] and *Liu et al.* [2006b] for detailed comparisons of PCA and SOMs.) SOMs have a number of advantages including readily handling nonlinear behavior and robust interpolation into areas of the input space not present in the available training input. SOMs also have the benefit of being a completely independent uniformitarian analysis pathway and thus provide independent results for comparison with more traditional techniques. (The sense of ''uniformitarian'' used here relates to the practice of calibration of a proxy to a (sometimes short) observational record, followed by extrapolation deeper into time (or space) using the proxy alone. *Alley and Cuffey* [2001] discuss this approach further.) SOMs have previously been used quite successfully in studies of synoptic-scale circulation in temperate latitudes [e.g., *Hewitson and Crane*, 2002; *Tennant*, 2004] and the polar regions [e.g., *Reusch et al.*, 2005b; *Cassano et al.*, 2006b; *Lynch et al.*, 2006]. SOMs support development of synoptic climatologies with an arbitrary number of smoothly transitioning climate states, in contrast to traditional synoptic classification techniques.

[10] When applied to atmospheric data, a SOM analysis yields a nonlinear classification of the continuum of atmospheric conditions. Training of the SOM (the analysis) effectively projects multidimensional data onto a two-dimensional array of nodes (representative states), each of which retains the full dimensionality of the original data through an associated reference vector representing the state's generalized pattern. The size of the grid (number of states) directly influences the amount of generalization: smaller (larger) node arrays have fewer (more) available states to characterize the n-dimensional data space, so the final states developed during training will tend to do more (less) generalization of the input. For mathematical reasons, namely stability during the learning process, the SOM grid normally has asymmetric dimensions rather than being square. Ideally, the grid dimensions match well with the shape of the two-dimensional projection of the input data set's probability density function (PDF [*Kohonen*, 2001]), but this is not always known in advance and is not normally a significant issue.

[11] The generalized patterns in a SOM grid are normally identified either by an (x, y) coordinate or a sequence number within the two-dimensional array of nodes. Using the row,column coordinate (e.g., 2,3 is row 2, column 3) identifies patterns consistently across different grid sizes. Sequence numbers (counting by row from 1 in the upper left to the maximum value, e.g., 12 for a 4×3 SOM, in the lower right) are easier to compare but not consistent across grid sizes. We use sequence numbers here for notational simplicity. We also follow the convention of showing the generalized patterns on a regularly spaced grid of nodes. In reality, the relative distance between nodes in data space is variable as a function of the information content and distribution of the raw data. Sammon maps [*Sammon*, 1969] are often used to visualize these relative distances (e.g., Figure 2). Sammon mapping projects multidimensional vectors into a 2-D space thereby simplifying visualization of the relationships between the vectors. Here the reference vectors for each generalized SOM pattern are projected allowing the SOM nodes to be plotted on the basis of relative neighbor-to-neighbor similarity, rather than the simpler, but less informative regularly spaced grid format normally used. However, for the purposes of displaying relative frequencies and related attributes of each node, it is adequate to treat the nodes as a regular grid.

[12] Details on SOM training are readily available in the literature [e.g., *Kohonen*, 2001; *Hewitson and Crane*, 2002; *Reusch et al.*, 2005a, 2005b; *Cassano et al.*, 2006b]. (However, because the application of this technique is still relatively new to climatology, we recommend reviewing multiple sources for their perspectives on best practices.) Briefly, an iterative and unsupervised process is used to adjust the reference vectors representing each SOM node on the basis of differences between the reference vectors and each input record. A learning rate (positive, typically less than 0.1) determines how much of the difference is applied in each adjustment. The node reference vectors are first initialized with values based on a two-dimensional subspace defined by the first two eigenvectors of the training data (rather than randomly) being most efficient (the eigenvectors serve as a first approximation of the data set) [*Kohonen*, 2001]. Training then consists of multiple iterations of reference vector adjustment until stable values are reached. In each iteration, the best matching reference vector is found for each input record (on the basis of the Euclidean distance between the input and the reference vectors) and updated to be more like the input data. Neighboring reference vectors of each best match are also updated (to a lesser degree) resulting in adjacent nodes having the strongest similarity and similarity decreasing with increasing internode distance. This training process ultimately produces reference vectors that each represent a distinct portion, or subspace, of the multidimensional input space.

[13] Once the SOM is trained, groups of similar input records can be created by matching to the final reference vectors, thereby producing a classification of the input data. Input records with patterns in common will map to the same generalized patterns of the SOM. (Note, however, that since SOM patterns are intended to be generalizations of the input data, the input data samples are not expected to match exactly with generalized patterns [*Kohonen*, 2001].) This basic classification procedure is the basis of an important additional analysis tool: frequency maps. Instead of looking at the individual data records in each group, a frequency map simply tracks how often a generalized pattern appears in a given subset of the input data (which may include the entire data set). Subsets may be temporal (e.g., by season), based on an external index (e.g., the NAO), or some other criteria found to be useful. Frequency maps based on such data stratification help to identify temporal changes and

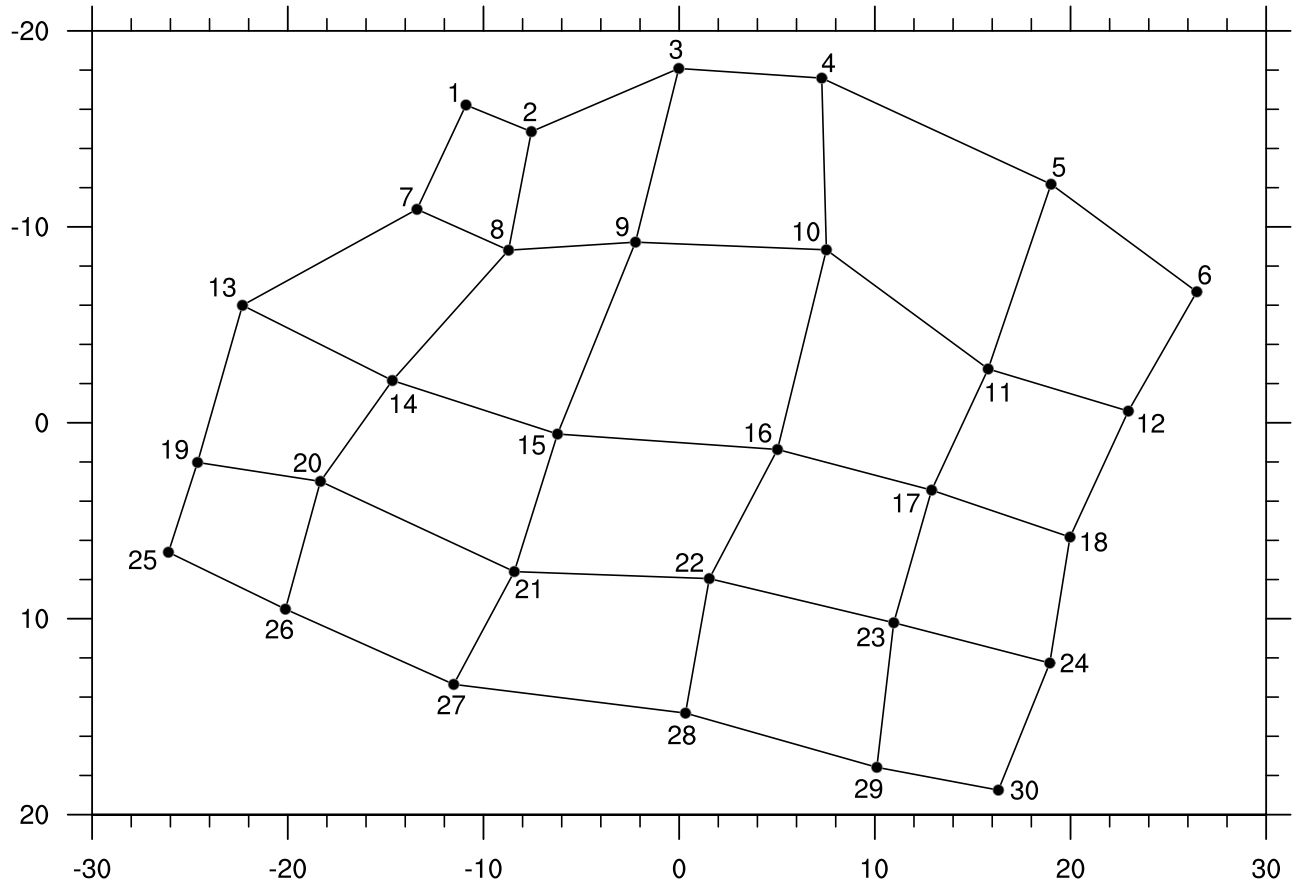

Figure 2. A Sammon map for the SOM in Figure 1 showing the relative position of each SOM pattern in the 2-D projection of the SOM's reference vectors (i.e., an estimate of the data's probability density function). Axis values are arbitrary (but based on distances between SOM node reference vectors).

relationships with other aspects of the climate system. For example (see below), we have used the NAO index to identify which SOM patterns are associated with high/low values of the index and to determine the statistics of the NAO index for each generalized pattern.

[14] As mentioned above, the size of the SOM directly affects the amount of generalization of the input in the final SOM patterns. Since the SOM grid size (the total number of patterns) is a subjective choice, it is standard practice to do the analyses using multiple grid sizes. Grid sizes normally range from 3×2 up to a value consistent with the amount of data available for training (having more patterns than input may not always be useful). For our 135-record monthly data set, we tested 3×2, 4×3, 5×3, 5×4 and 6×5 SOMs. We will be focusing on results from the largest grid but will also compare results between grids. All of our analyses were run on the complete data set (135 months x 819 grid points) using the freely available SOM-PAK software [*Kohonen*, 2001].

2.3. Principal Component Analysis

[15] Principal component analysis (PCA) is a well-known, well-documented statistical technique widely used in climatology to identify dominant patterns of variability and/or reduce the dimensionality of climate data [e.g., *Lorenz*, 1956; *Kutzbach*, 1967; *Walsh*, 1978; *Cohen*, 1983; *Smith et al.*, 1996]. Briefly, standard PCA creates a set of new orthogonal (hence uncorrelated) variables that most closely and efficiently (in a statistical sense) represent the variance in a data set [e.g., *Jolliffe*, 2002]. In theory, the new variables provide a simpler and more easily interpretable version of the original data, in part through dimensionality reduction since generally only the first few components are significant. Used in this manner, PCA has been widely and successfully applied to help understand, interpret, and reconstruct large, multivariate climate data sets. We refer the reader to the literature for further details [e.g., *Jolliffe*, 2002] and to *Reusch et al.* [2005a] and *Liu et al.* [2006b] for comparisons with SOMs in pattern identification. Because of mathematical requirements (a valid PCA requires more rows than columns in the input), the input grid was subsampled (every third row/column) to a 91-point grid for the PCA analysis. We have no reason to believe PCA results would change significantly using the full grid. The 135×91 data matrix was then the input for an unrotated S-mode [*Cattell*, 1952] PCA using Matlab. Time series for each PC were created by scaling the principal component coefficients into data space using the standardized input data.

3. Results

[16] In this section we describe results from the SOM- and PCA-based analyses of the MSLP monthly mean and standard deviation data.

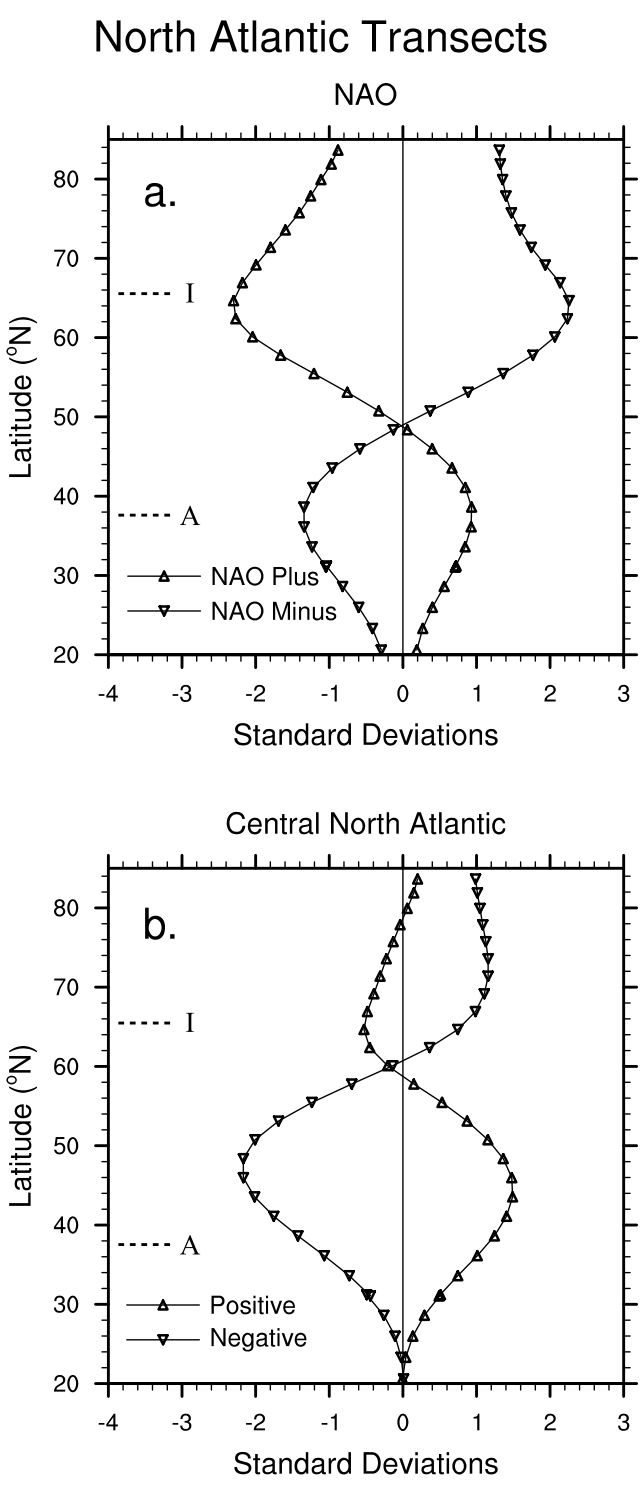

Figure 3. Transects in the North Atlantic basin. Grid points closest to a great circle through Ponta Delgada, Azores (37.7°N, 25.7°W), and Akureyri, Iceland (65.7°N, 18.1°W), are shown for the four corner patterns of Figure 1a: (a) high and low NAO patterns from lower left and upper right and (b) high and low central North Atlantic patterns from upper left and lower right. These locations were used in the NAO index of *Lamb and Peppler* [1987] and are indicated by the dashed lines with labels A and I, respectively. Values are grid point anomalies as in Figure 1.

3.1. SOM-Based Analysis of Monthly MSLP

[17] The generalized patterns (Figure 1) from the 6 × 5 SOM analysis of monthly MSLP mean and standard deviation, analyzed jointly, capture multiple aspects of the variability of this climate field. Complexity is readily seen in the variety of locations, extents and magnitudes of anomalies present in the patterns of mean MSLP (Figure 1a). Note that while two variables (mean and standard deviation of MSLP) were analyzed, only one SOM grid is produced and each SOM node contains both variables. The mean and standard deviation patterns are shown separately for clarity. Figure 1a includes results from a frequency mapping of the complete data set, i.e., the values (% frequency) to the upper right of each pattern. Given a uniform distribution of 135 months over 30 available patterns, the expected frequency of occurrence of any given pattern is 3.3%. Thresholds of 50% more (4.95%) or less (1.65%) may be used to (subjectively) identify patterns that occur more or less frequently than expected. With these criteria, Figure 1a shows a total of six patterns occurring more (patterns 1 and 25) or less (patterns 5, 7, 9 and 11) frequently than might be expected. We will return to pattern frequency as a measure of temporal variability below. Overall, the two main features of the generalized patterns are seen at each end of the two diagonals (i.e., the sets of opposite corners): a north-south dipole suggestive of the NAO (patterns 25 and 6) and a mid-North Atlantic "center of action" (patterns 1 and 30).

[18] Before discussing the patterns further, we briefly explore the similarity relationships of the MSLP patterns via the Sammon map of the SOM reference vectors (Figure 2), which shows only modest distortion of the overall SOM grid. The lower left/upper right diagonal is moderately longer than the upper left/lower right diagonal (∼5:4 ratio of lengths), indicating relatively less similarity between the end patterns of the former diagonal. Patterns with relatively greater similarity can be seen in the upper and lower left corners (nodes 1, 2, 7 and 8, and nodes 19, 20, 25 and 26, respectively). There are also some areas of relatively larger differences between patterns, e.g., 9, 10, 15 and 16. It also appears that the lower left patterns are more closely related to each other than their counterparts in the opposite corner.

3.2. Leading Patterns of Variability: Evidence for the NAO

[19] Patterns characteristic of the variability of MSLP in the NAO [*Hurrell et al.*, 2003] appear in the lower left and upper right corners of Figure 1. Negative anomalies in the Icelandic Low and positive anomalies in the Azores High (Figure 1a, lower left corner patterns) are associated with higher standard deviations across western Europe and Scandinavia (Figure 1b), a reflection of the increased storm activity associated with positive-mode NAO periods. Conversely, positive Icelandic Low anomalies and negative Azores High anomalies (Figure 1a, upper right corner patterns) are associated with an increased storminess in Greenland and a reduction in Scandinavia (Figure 1b). Thus the patterns on the lower left–upper right diagonal capture both the meridional seesaw of pressure changes in the Icelandic Low and Azores High and the horizontal changes over time in the centers and extent of these two pressure

Table 1. Extreme Values of the CRU Monthly NAO Index [*Jones et al.*, 1997] With Associated Month of Occurrence, in Chronological Order From 1957 to 2002, and Best Matching SOM Pattern

Highest ($x \geq 4$)			High ($3 \leq x < 4$)			Low ($-4 < x \leq -3$)			Lowest ($x \leq -4$)		
Date	NAO	SOM	Date	NAO	SOM	Date	NAO	SOM	Date	NAO	SOM
Feb 1961	4.06	20	Jan 1974	3.75	26	Dec 1963	−3.09	17	Jan 1963	−4.09	6
Jan 1983	4.82	19	Dec 1974	3.06	19	Feb 1965	−3.03	5	Feb 1986	−4.02	12
Feb 1990	5.11	25	Dec 1986	3.42	20	Feb 1969	−3.16	6	Dec 1996	−4.70	6
Feb 1997	5.26	25	Jan 1989	3.53	13	Dec 1976	−3.63	17			
Feb 2000	4.37	13	Feb 1989	3.61	25	Jan 1979	−3.22	6			
			Jan 1990	3.50	26	Dec 1995	−3.33	12			
			Feb 1992	3.18	13	Jan 1996	−3.27	16			
			Jan 1993	3.91	19						
			Feb 1995	3.13	25						
			Feb 2002	3.01	25						

systems. A transect through two NAO index locations (Ponta Delgada, Azores and Akureyri, Iceland, from the index of *Lamb and Peppler* [1987]) provides an alternative, one-dimensional view of the changes between these two patterns (Figure 3a). The SOM patterns clearly capture the changes in the Azores High and Icelandic Low between high and low NAO index conditions. Minima and maxima in the transect are also very close to the NAO index latitudes, indicating that the two-dimensional SOM generalization has incorporated the one-dimensional NAO index metric.

[20] Quantitative evidence (and confirmation) that these patterns represent the NAO comes from stratification of the input data using high and low values (absolute values greater than 3) of the CRU NAO index [*Jones et al.*, 1997]. The 15 highest NAO index months (Table 1) map to five patterns at the lower left (13, 19, 20, 25 and 26). Similarly, the ten lowest NAO index months (Table 1) map primarily to the upper right corner patterns (5, 6, 12 and 17) (one month, January 1996, matches a more intermediate pattern (16)). These high (low) index patterns account for ~21% (~14%) of the input data. Note that storminess (positive anomalies in the standard deviation of MSLP, Figure 1b) over Scandinavia is not exclusively associated with the NAO extreme periods. Positive (7, 8, 14) and negative (18, 23, 24, 30) anomalies of the standard deviation are also seen in patterns with less extreme NAO index values (Figure 1b).

[21] Figure 4 graphically summarizes statistics for the CRU NAO index versus the SOM patterns, i.e., match count, the mean NAO index and its standard deviation is shown for each SOM node. Within each subplot, each representing one SOM pattern, the mean and standard deviation are plotted as a cross on an axis centered on each position of the SOM grid. This cross is located at the mean value of the NAO index. The horizontal arms of the cross are scaled to ±1 standard deviation of the NAO index values of the months mapping to a given pattern, a measure of variability of the mean value. This format readily shows that the highest (lowest) average index values are found among the lower left (upper right) patterns, as seen in the index-based stratification of the MSLP data. Also, roughly 1/3 of the patterns have mean NAO index values (possibly) indistinguishable from zero (i.e., the range of the mean ± the standard deviation includes zero). Considering average index values between ±1 as neutral (13 patterns), there are ten positive (7, 9, 13–14, 19–21, 25–27) and seven negative (5–6, 12, 16–18, 30) patterns, collectively accounting for approximately 33% and 21% of the input data, respectively. These frequency values are in line with the relative frequency of positive and negative values in the original index series (51 and 35 months, or 38 and 26%, respectively). Apart from a few relatively low values, standard deviations of the NAO index values are comparable between patterns. Further aspects of the NAO as seen through the SOM analysis will be discussed below.

3.3. Evidence for Other North Atlantic Variability

[22] As with the NAO and North Atlantic climate variability, the NAO-related patterns represent only a portion of the generalized patterns extracted by the SOM. In fact, the most frequent single pattern in the SOM, pattern one (8.1%), is unrelated to the NAO; that is, it only involves one of the centers of action associated with standard NAO indices. This pattern, and its approximate complement (pattern 30), represent extremes of a northward shifted Azores High (or, more generally, a central North Atlantic anomaly) associated with approximately neutral NAO index values. In the positive mode (patterns 1, 2, 7 and 8), the subpolar low has migrated well away from the climatological position near Iceland and is generally weak. Higher pressure extends northward over Greenland in the most frequent of the positive mode patterns, although this feature may be an artifact of reducing the high-elevation surface pressure over the ice sheet to sea level (areas of central Greenland exceed 2300 m asl). The four positive patterns account for nearly 17% of the input data. Patterns from the negative mode (24, 28–30) have a below normal Azores High with, again, a near neutral/slightly positive anomaly subpolar low (closer to its normal position near Iceland) and account for 14% of the input data. The NAO index-based transect through the corner patterns (1 and 30, Figure 3b) shows that the region to the north of the Azores is the main area of change between the two extremes with variability nearly twice that seen in the NAO transects (Figure 3a). Also of note is the greatly reduced variability in the highest latitudes (“Icelandic”) versus the NAO transects.

[23] In short, the climate states extracted from MSLP mean and standard deviation partition the input into NAO positive (32%; 9, 13–14, 19–21, 25–27), NAO negative (18%; 5–6, 12, 16–18), above normal Azores High (17%; 1–2, 7–8) and below normal Azores High (14%; 24, 28–30) modes. The balance of the data (19%) is found among the suite of intermediate (transitional) patterns, which may

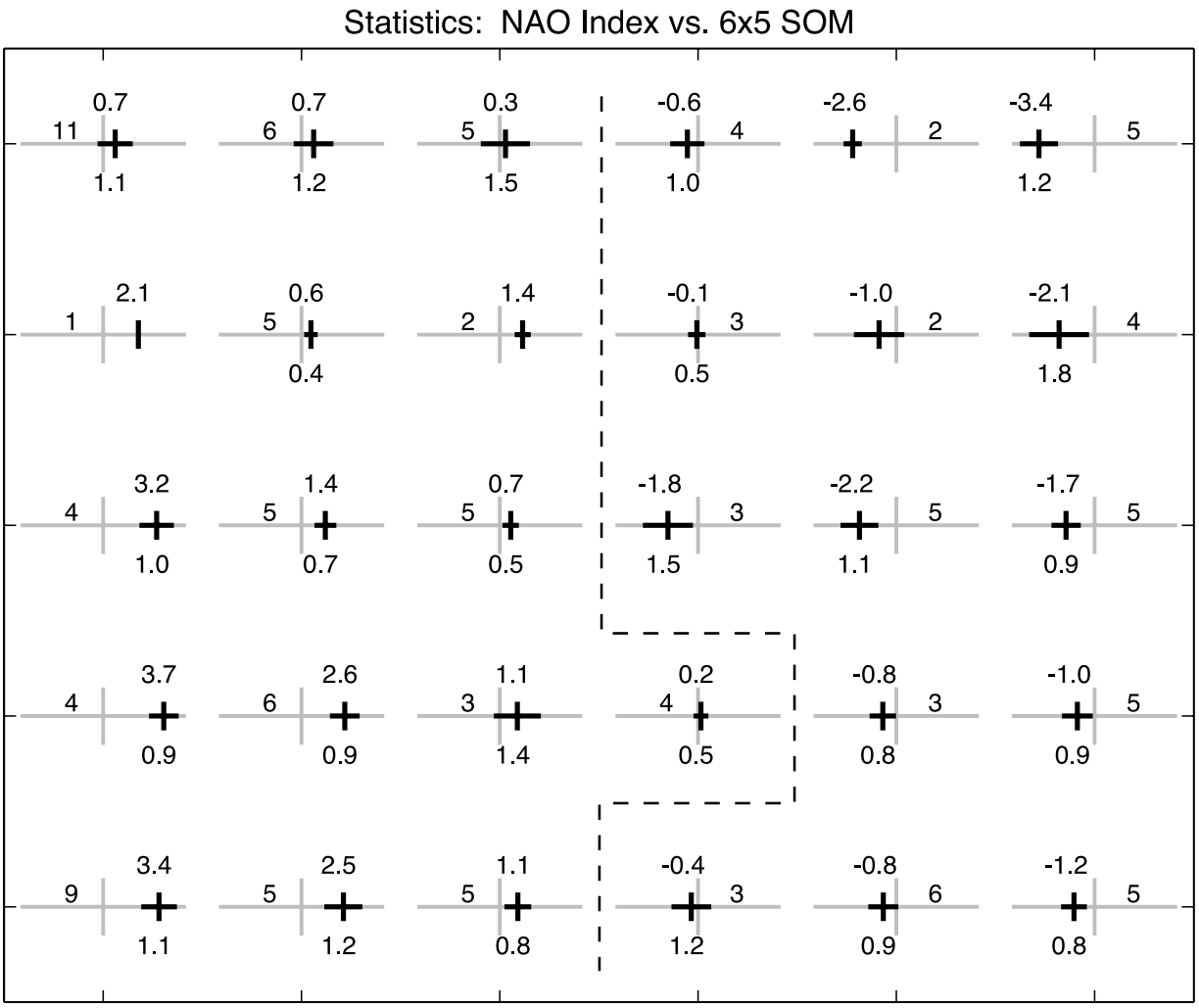

Figure 4. Statistics of the CRU monthly NAO index [*Jones et al.*, 1997] versus the SOM, shown as plots of mean and standard deviation centered on each position of the SOM grid. For each pattern plot, a solid cross is located at the mean of the NAO index values of the months mapping to the pattern (shown above each plot). The horizontal arms of the cross are scaled to ±1 standard deviation of the NAO index values (shown below each plot, but omitted for patterns with three or fewer matches), a measure of variability of the mean value. Where the range of mean ± standard deviation includes zero (the *y* axis of the light shaded cross), the mean may not strictly be distinguishable from zero. The number of months matching each pattern is shown at the left or right of each plot. The dashed line separates positive and negative index values. Note that height is not significant in these plots.

represent nonlinear combinations of these modes (and potentially other, less common behaviors). Note that the Sammon map of the SOM (Figure 2) indicates that the NAO is the best, simple (i.e., first-order) way to classify the data since the NAO patterns fall on the longer diagonal (lower left/upper right). The relative lengths of the two diagonals (∼5:4) suggests that the central North Atlantic/Azores High patterns should also explain substantial variability (as they do). Thus the overlapping spatial extent of the leading modes of variability in the data has not prevented the SOM from successfully extracting patterns that capture this complex variability.

3.4. SOM Grid Size and Pattern Generalization

[24] SOM analyses were run on a number of different grid sizes (complete pattern sets not shown) to examine how generalization varied as the number of available patterns increased. Here we focus on the 3×2, 4×3, 5×4 and 6×5

Figure 5. Comparison of high NAO index patterns from different size SOM grids, ordered by frequency count: (a) 3×2 grid, 6 total patterns; (b) 4×3 grid, 12 total patterns; (c) 5×4 grid, 20 total patterns; and (d) 6×5 grid, 30 total patterns. To the upper left of each pattern is its grid index (not directly comparable between grid sizes because of changing number of patterns). To the upper right of each pattern is the frequency count (number of high NAO index months mapping to the pattern). There are a total of 15 months with index > 3 (Table 1).

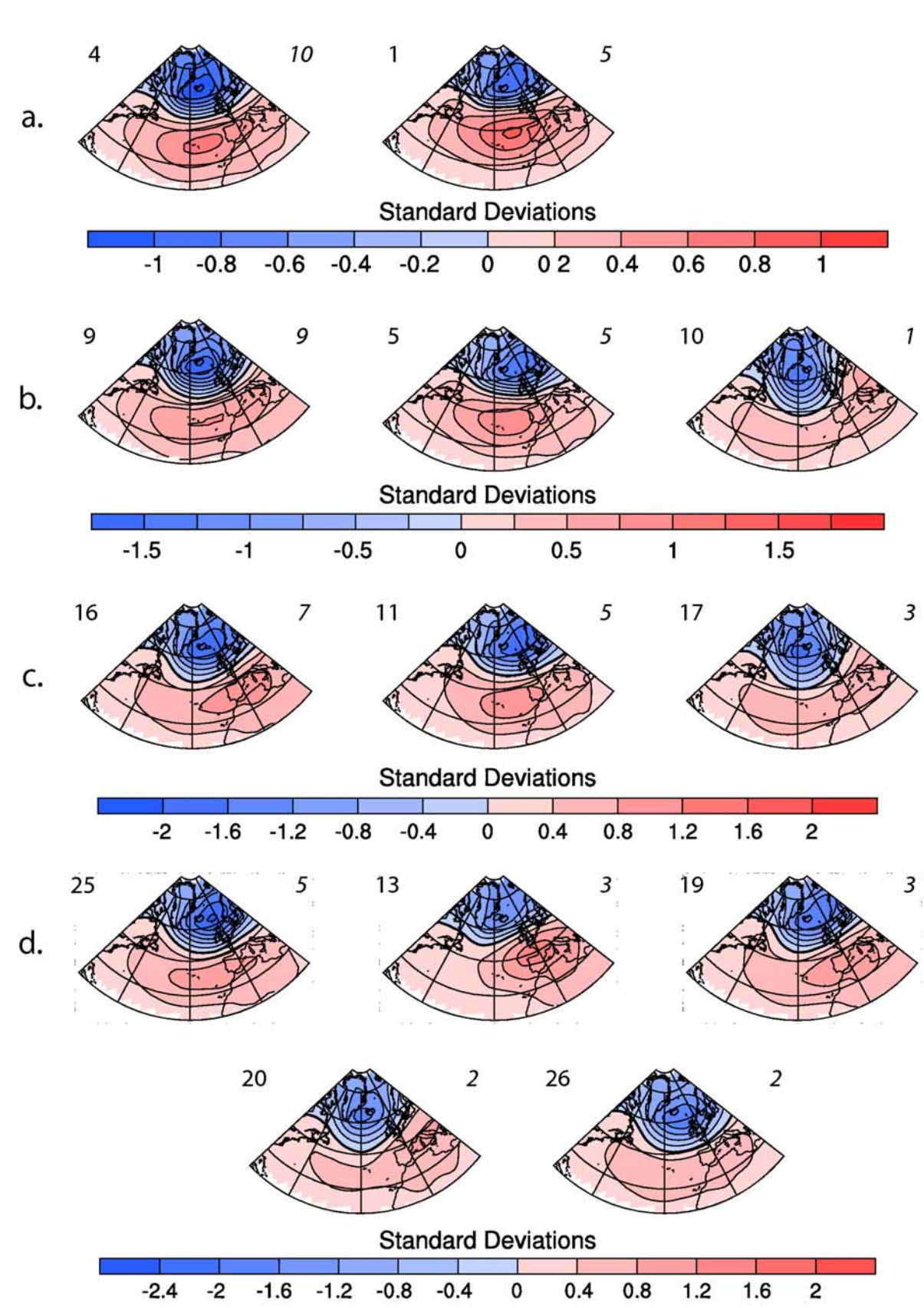
a.
4
10
1
5
Standard Deviations
-1 -0.8 -0.6 -0.4 -0.2 0 0 2 0.4 0.6 0.8 1
b.
9
9
5
5
10
1
Standard Deviations
-1.5 -1 -0.5 0 0.5 1 1.5
c.
16
7
11
5
17
3
Standard Deviations
-2 -1.6 -1.2 -0.8 -0.4 0 0.4 0.8 1.2 1.6 2
d.
25
5
13
3
19
3
20
2
26
2
Standard Deviations
-2.4 -2 -1.6 -1.2 -0.8 -0.4 0 0.4 0.8 1.2 1.6 2

Figure 5

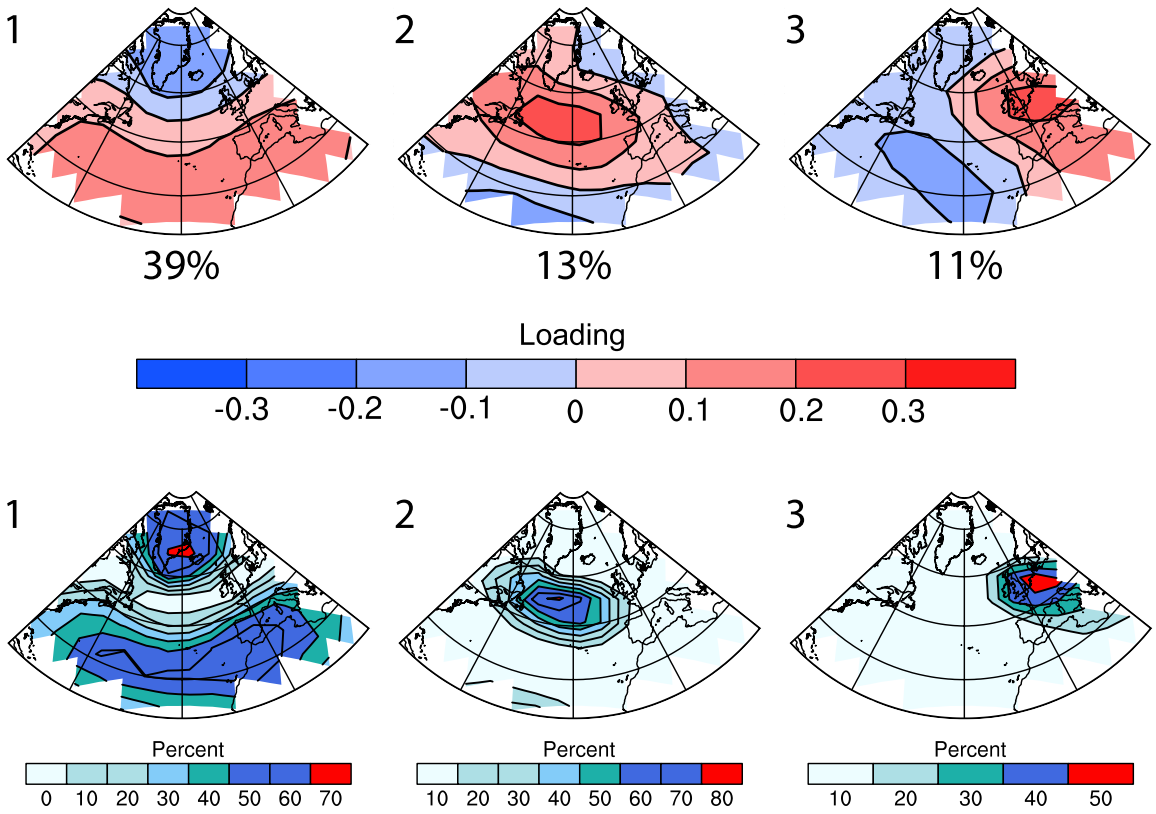

Figure 6. (top) Leading patterns from a principal component analysis of monthly mean MSLP with (bottom) corresponding variance. Patterns are shown as the loading of each grid point on the principal component.

grids and just the patterns associated with the 15 high NAO index (CRU index > 3) months (Figure 5). Generalization by the SOM can be seen in a number of characteristics. First, as seen in the color bar scales, the full SOM is capturing more of the variability in the data as the number of patterns increases, roughly doubling from the smaller to the larger grids. This reflects the denser coverage of the input data's PDF by the SOM through the growing number of available patterns, which lets each individual pattern get closer to the original data, but also reduces generalization. Second, all of the grids split the 15 high-index months across at least two patterns with more subgroups appearing as more patterns become available (patterns are ordered by frequency in Figure 5).

[25] The smallest grid (Figure 5a) has split the input into two groups based primarily on the strength of the Azores High. This is not, however, a direct stratification of the high (3 < index < 4) and highest (index > 4) months as there are months from both index ranges associated with each pattern (not shown). The two intermediate size grids (Figures 5b and 5c) each split the high NAO index months over three patterns, two of which resemble those of the 3 × 2 grid. Comparing the two leading patterns from each grid, we see moderate refinement of the shape, position and depth of the centers of action as the grid size increases.

[26] At the largest grid size, the 15 months are distributed nearly uniformly (5, 3, 3, 2, 2) across the patterns in contrast to the smallest grid which was quite asymmetric (10, 5). The five patterns of the 6 × 5 grid (Figure 5d) include those seen in the smaller grids plus two transitional (intermediate) patterns not seen before (13 and 20). The character of these new patterns shows that they are capturing more of the spatial variability to be found in high NAO index conditions. In fact, the fourth and fifth highest index months (February 2000 and March 1961, each with index greater than four) both map to these new patterns. This is an excellent example of the additional insight the SOM approach brings to understanding the NAO. It is well known that using a two-points-in-space NAO index does not readily reveal the high spatial variability of the system. In combination with the SOM approach, however, the full range of variability is easily seen.

3.5. PCA of Monthly Mean MSLP

[27] Figure 6 shows the three leading principal components (PCs) and associated variance patterns from a PCA of monthly mean MSLP. (Unlike the SOM analysis, standard deviation was not included in the PCA). For simplicity, we did not attempt any rotations of the PCs, although this extra step is often found in the literature [e.g., *Richman*, 1981, 1986]. PC1 represents the traditional, PCA-based version of the NAO and resembles NAO patterns found throughout the literature (although it does not necessarily exactly reproduce any single one). It accounts for 39% of the variance of the monthly MSLP data. Differences from previously published studies of the NAO [e.g., *Hurrell et al.*, 2003] are attributed to minor differences in methodology, e.g., our specific spatial domain and time period used. Note that both the high and low NAO index states are present in PC1 because

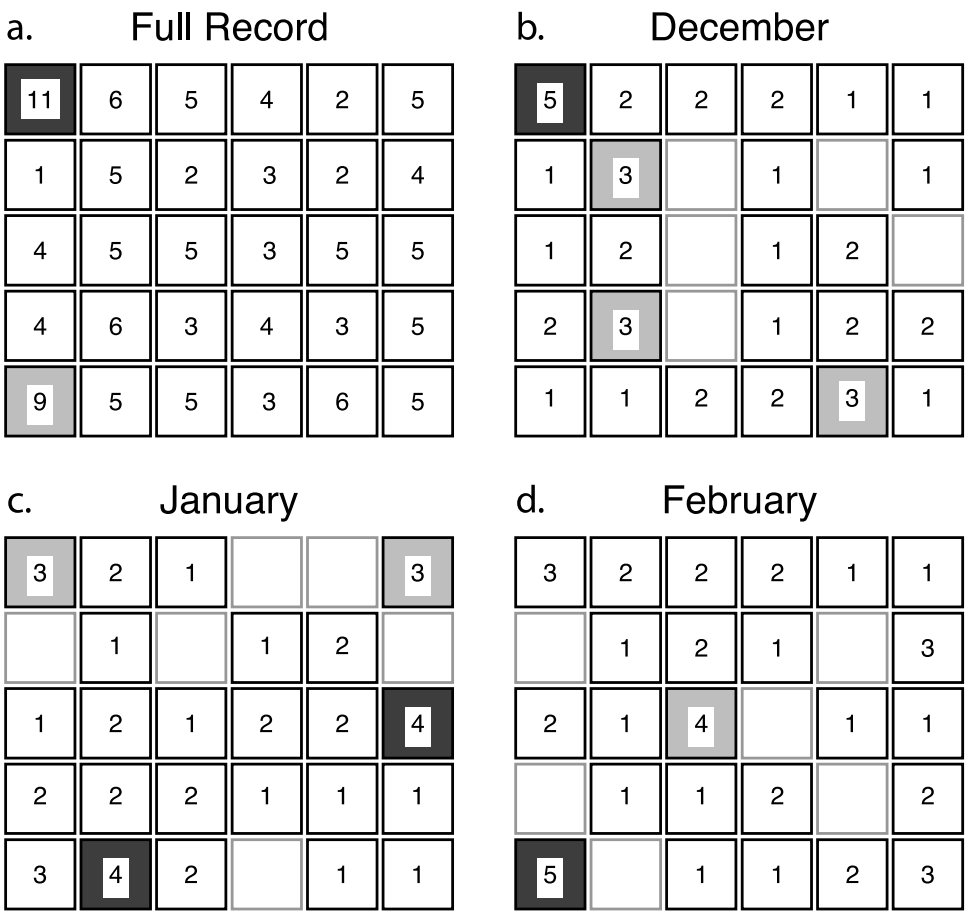

Figure 7. Frequency map for the SOM patterns of Figure 1 based on (a) the complete record and (b) December, (c) January, and (d) February. Numbers in each square are the number of times that pattern is the best match for an input record. The highest count has a near-black background and the second-highest count has a gray background. With 135 and 45 months per interval, for the full record and monthly plots, respectively, and 30 SOM patterns per plot, the expected average frequency counts are 4.5 and 1.5 per pattern (or 3.3% in relative terms).

of the symmetry of PCA analysis. The center of action in the central North Atlantic seen in PC2 is suggestive of the variability seen in patterns 1 and 30 of the SOM analysis. PC3 represents localized variability only over portions of northwest Europe.

4. Discussion

[28] Temporal variability in the generalized SOM patterns, the additional benefits of the SOM patterns versus just using an NAO index, and a brief comparison of SOM-based versus PCA-based NAO analysis are covered in this section.

4.1. Temporal Variability

[29] Frequency maps for the full record and each month are shown in Figure 7. As noted previously (Figure 1), the most frequent patterns in the full record reflect a high NAO index (lower left) and a high central North Atlantic anomaly (upper left). Monthly frequency maps show how the frequency of occurrence of each SOM pattern changes from month to month (although the sample size (45) is relatively small compared to the number of possible states (30)). Comparing each month (Figures 7b–7d) with the full record (Figure 7a) shows how each month contributes to the full set of patterns. The monthly maps show some notable differences between the three months (e.g., the upper left corner peak in December and the lower left corner peak in February). Patterns are not distributed uniformly and each month has five or six patterns that do not occur during that month over the 45 year record. However, no pattern exceeds ~11% and each month is characterized by multiple peaks in frequency. December tends to favor more neutral NAO index conditions with the strongest peak in the upper left corner (high central North Atlantic mean anomaly). However, high NAO index patterns (13, 19, 20, 25 and 26) also occur eight times. January also has a secondary peak (3) in the upper left corner but includes more high (12) and low (nine; patterns 5, 6, 12, 16 and 17) NAO index patterns. February has a relatively strong peak (five) on the highest NAO index pattern (26) but includes a secondary peak on pattern 15. The latter shows high (northwest Europe) and low (Greenland) mean subpolar anomalies with close to neutral values in the subtropics.

[30] To simplify visual comparisons and better highlight month-to-month differences, frequency counts can be

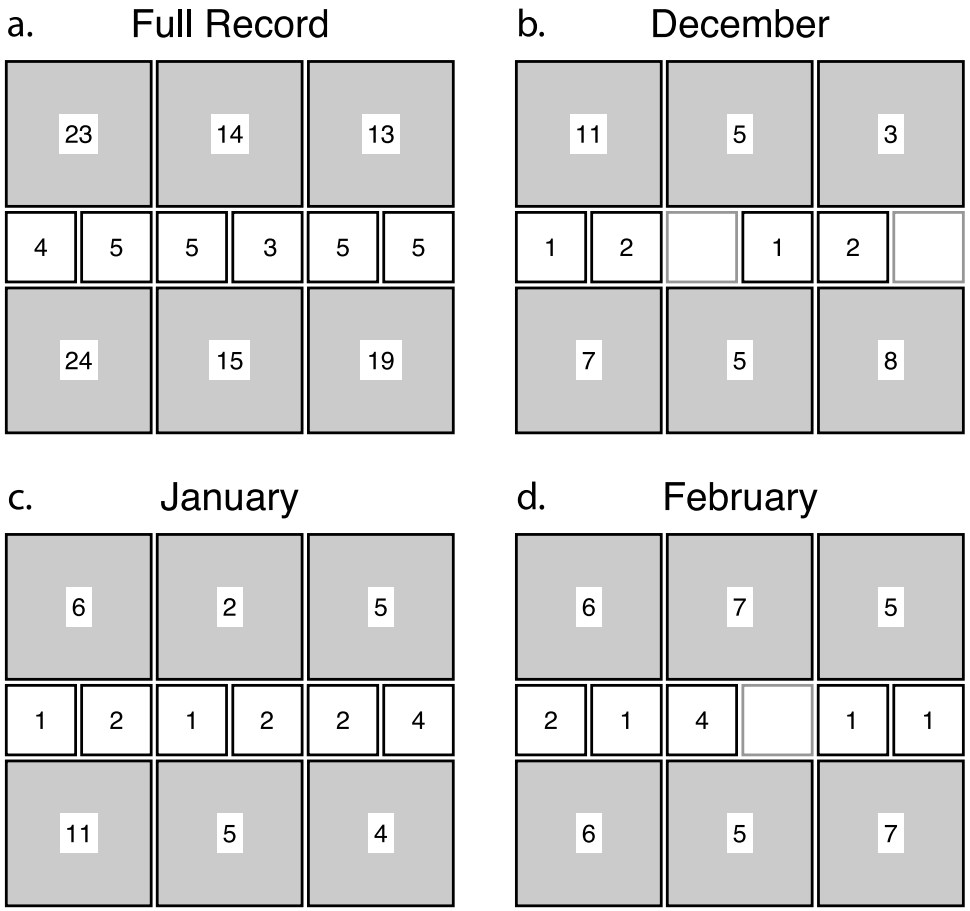

Figure 8. Frequency maps by month with summation of 2 × 2 blocks in corners and along top and bottom edges. Differences between months are more distinguishable in this format (e.g., upper/lower left differences between December and January).

grouped to show larger patterns in the SOM grid. Here we group the corners and upper/lower rows in 2 × 2 sections (Figure 8). The upper/lower left differences between December and January are now more distinct. December also has a more obvious peak in the lower right corner, indicating that the central North Atlantic anomaly ''monopole'' appears to be the strongest influence in this month (along with the positive NAO mode). January's variability is more obviously centered on the positive NAO mode. In contrast, February does not appear to have any particularly dominant pattern (or set of patterns).

[31] Further subdividing the time series into 15-year periods (reducing each frequency map to 15 input records) shows how the pattern frequencies have changed over time within each month (Figure 9) and highlights temporal variability. February in particular shows a trend from having no dominant pattern at all in the early (1957–1972) record to a more clear structure with one NAO high pattern (26) most common in the late (1987–2002) record. The latter is an expression of the generally higher frequency of high NAO index values that began in the 1970s, which also appears (to lesser degree) in the other months. Changes in January show a shift from negative NAO and central North Atlantic anomalies to a focus on positive NAO patterns, similar to February but more spread out. To better understand the temporal changes, Figure 9 includes the mean x,y position for the time slice being mapped in each frequency map and a summary of how the mean position changes over the three time intervals (the trajectories). (It is important to remember that these coordinates are further abstractions of the SOM results that do not represent particular SOM patterns. One could conceivably weight the SOM patterns according to the coordinate values to create a ''mean pattern,'' but it is unclear whether this brings any value.) On the basis of this further summarization of the data, December has changed relatively little over the ERA-40 period while the remaining months have both shifted more toward patterns reflecting a mix of positive NAO and a high central North Atlantic anomaly conditions. From this perspective, January has shifted more strongly toward positive NAO than February. This shift occurred earlier in January than February, starting from approximately the same position. These results show that the increase in high NAO index months has not been a simple matter of across-the-board change but that there has also been subseasonal variability (including changes in occurrence of the central North Atlantic patterns).

[32] Altogether, these frequency maps show that while the winter months are known to carry the strongest expression of the NAO, the spatial characteristics of that expres-

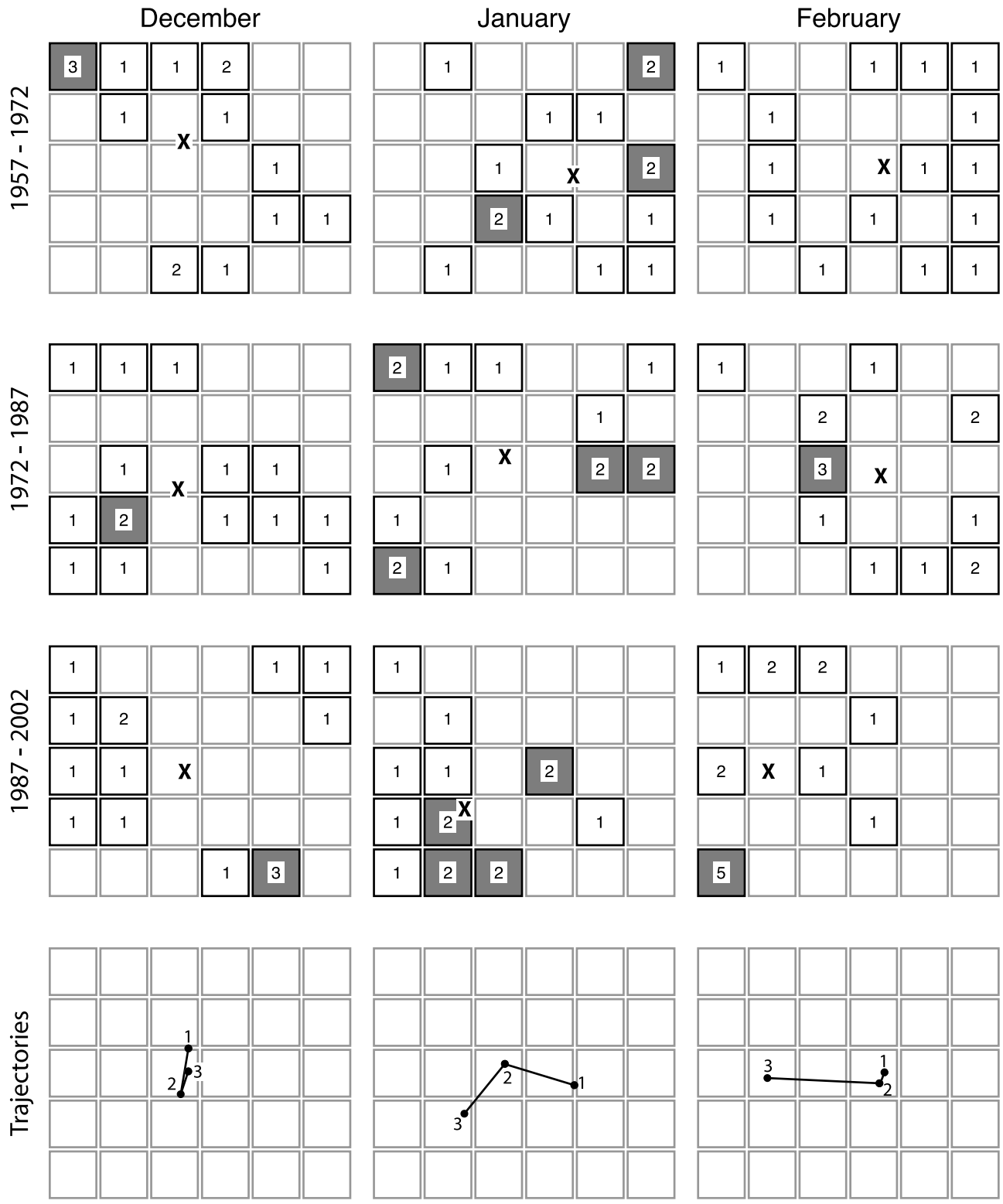

Figure 9. Frequency maps (first, second, and third rows) and trajectories (fourth row) by 15-year period and month. With 15 months per interval and 30 SOM patterns per plot, the expected average frequency count is 0.5 per pattern. Frequency maps include the mean x,y coordinate for the subset plotted (indicated by the cross). Highest-frequency patterns have a dark background. Trajectories summarize the change in position of the mean coordinate for each month.

Table 2. SOM Patterns With Associated High/Low Index Months From the CRU Monthly NAO Index [*Jones et al.*, 1997]

Pattern	High NAO Index Months	Pattern	Low NAO Index Months
13	Jan 1989, Feb 1992, Feb 2000	5	Feb 1965
19	Dec 1974, Jan 1983, Jan 1993	6	Jan 1963, Feb 1969, Jan 1979, Dec 1996
20	Feb 1961, Dec 1986	12	Feb 1986, Dec 1995
25	Feb 1989, Feb 1990, Feb 1995, Feb 1997, Feb 2002	16	Jan 1996
26	Jan 1974, Jan 1990	17	Dec 1963, Dec 1976

sion are not the same month-to-month. Further, the SOM captures the full range of variability: NAO patterns do not particularly dominate in frequency of occurrence.

4.2. SOM Patterns Versus the NAO Index

[33] As discussed above, stratification of the input data using high/low values of a monthly NAO index [*Jones et al.*, 1997] shows that similar extremes in the index can be associated with quite different MSLP patterns from the SOM analysis (Figure 4). This is not altogether surprising given the linear, two-points-in-space nature of the NAO index, but the SOM does have significant potential for insights to these spatial characteristics. (As an historical aside, the index used by Walker and Bliss in 1932 included pressure and temperature records from nine sites around the North Atlantic basin [*Stephenson et al.*, 2003].)

[34] Table 2 shows the months for the highest (>3) and lowest (<−3) values of the CRU NAO index [*Hurrell*, 1995; *Jones et al.*, 1997] with their associated SOM patterns (on the basis of matching each month's data to the generalized SOM patterns). Of the 135 months in the analysis, 15 (10) were in the highest (lowest) category, or approximately 11% (7%) of the 1957–2001 record. (Overall, approximately 1/3 of the months had absolute NAO index values > 2.)

[35] For the high patterns (Figure 1), the Icelandic Low has moderate variability in depth, location and extent and does not wander far from the Greenland/Iceland region. The variability of the Azores High is much greater for all three parameters as the anomaly sometimes stretches across the whole North Atlantic basin. The storm tracks, as reflected in MSLP standard deviation, are centered over Scandinavia and northwest Europe primarily with reduced activity over Greenland and southwest Europe.

[36] Four of the five MSLP mean patterns (excluding 16) for the low NAO index values show the expected filling of the Icelandic Low and reduced Azores High. The corresponding MSLP standard deviation patterns generally show the expected decrease of storminess in Scandinavia with modest increases to the north and south. Notably, storminess for pattern 17 is increased in a broad region centered south of Greenland rather than in the eastern north Atlantic. The remaining MSLP mean pattern (16) indicates a state closer to neutral across the north Atlantic but with increased storminess in a north-south band near 30°W. The two-points-in-space nature of the NAO still produces a low index value under these conditions. The SOM patterns thus readily highlight the benefits of a full spatial analysis.

4.3. SOM and PCA Results Versus the Data

[37] Figure 10 compares the original monthly MSLP data to the corresponding SOM- and PCA-based patterns for a small number of months with high and low NAO index values. Note that this subset is not meant to be representative of the overall performance of either technique and none were selected because of particularly good or bad results. Also, as mentioned above, minor features over Greenland should be treated with caution because of the high elevations there. Different grid spacings (250 km for the SOMs, 750 km for PCA) may also produce apparent differences between results due to differences in contouring, for example.

[38] To a first approximation, the SOM and PCA patterns do reasonably well in capturing the main features of the high NAO index months and the first low NAO index month (December 1995). Both the SOM and the PCA use almost the same pattern for the two high NAO index months, although there are more (small) differences between the SOM patterns than the PCA patterns. For the lowest NAO index month, the SOM pattern, while not as good a match as the others, is distinctly better than the PCA pattern. The latter uses much the same pattern as for the previous low NAO index month and does not reflect the changes of the spatial pattern between the two months.

[39] As an alternative appraisal of technique performance, Figure 11 shows the residual difference between the data and its corresponding SOM and PCA patterns. Residuals will be negative (positive) in areas where the data are negative (positive) and the SOM or PCA pattern underestimates the data (and conversely in areas of overestimation). There are a number of examples where both techniques "miss," e.g., the small minima/maxima in the residuals for the two low NAO index months. Overall, however, there tend to be more differences than similarities in how each technique performed.

[40] Root-mean-square (RMS) values are often used to quantify predictive skill, e.g., when comparing model forecast results to actual observations. This metric is not ideal for spatial data, since it does not account for the spatial variability of the residuals, but it still provides a useful perspective for comparisons. Table 3 has summary statistics for all the patterns that, overall, indicate that SOM residuals are closer to the original data than the patterns produced with PCA. (Note that we do not expect SOM patterns or PCs to match exactly since they are generalizations and orthogonal projections of the data, respectively.) The RMS values for specific months shown in Figure 11 include cases where the PC1 patterns do perform better than their SOM counterparts (as suggested by the overlapping ranges of the RMS values in Table 3), but these are the exception, not the rule.

4.4. PCA Results Versus the SOM

[41] Figure 12 summarizes PC1 values for each pattern of the 6 × 5 SOM in the same style as our comparison of the CRU NAO index in Figure 4. Comparison of Figures 12

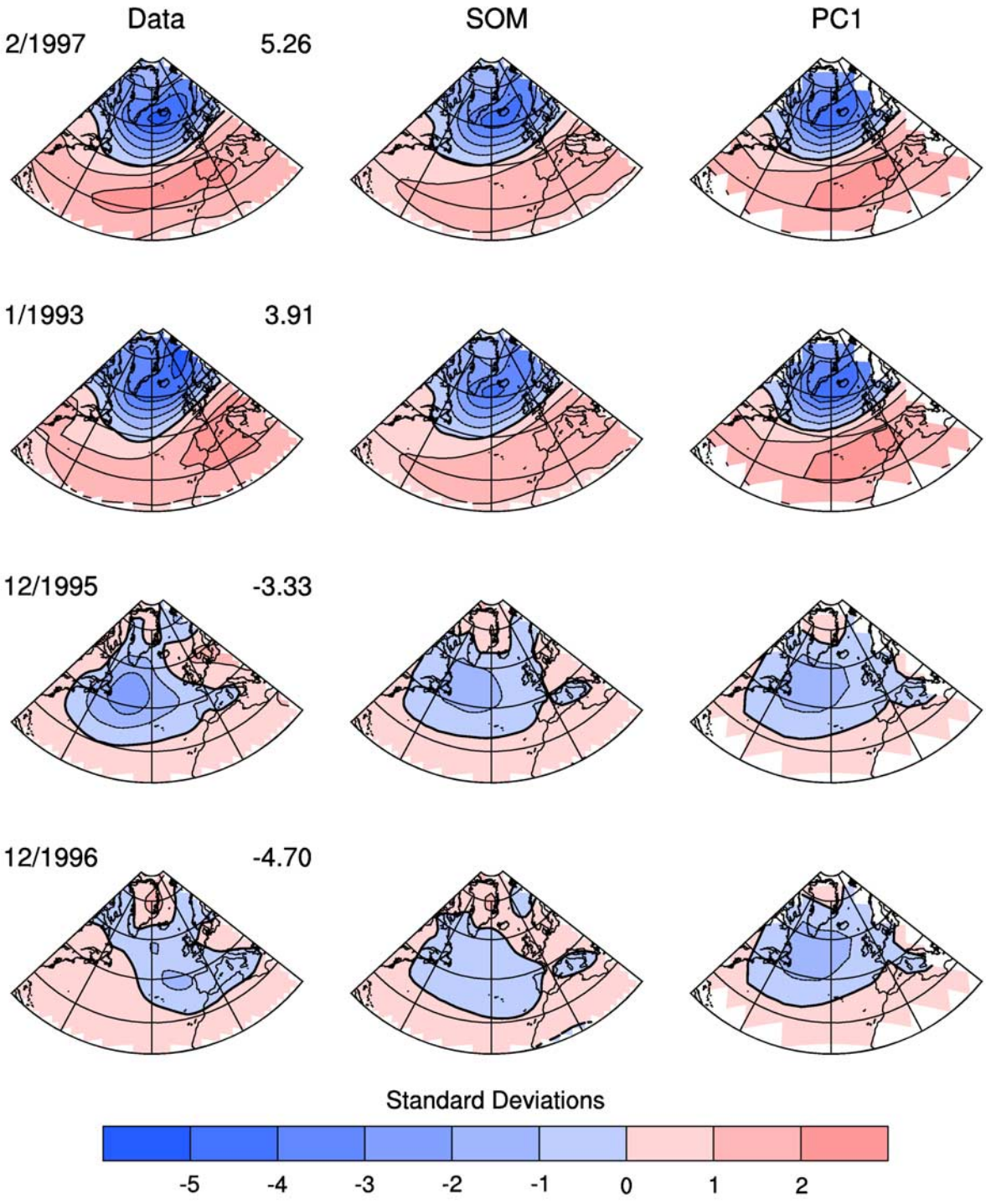

Figure 10. Selected MSLP patterns based on high/low values of the CRU monthly NAO index [*Jones et al.*, 1997]. Columns show the original data (with date and NAO index value) (first column), the corresponding SOM pattern (second column), and the corresponding PC1 pattern (third column). The mean and standard deviation are approximately 1015 mbar and 6 mbar, respectively.

and 4 shows some similarities. For example, average PC1 values are high (low) in the lower left (upper right) corner. This reflects (and supports) usage of PC1 in the literature as a proxy NAO index. However, Figure 12 also has large negative values for all patterns in the rightmost column of the grid, not just the upper right corner. This suggests that the orthogonal characteristics of the PCA may be blending aspects of the NAO dipole with the variability in the Azores

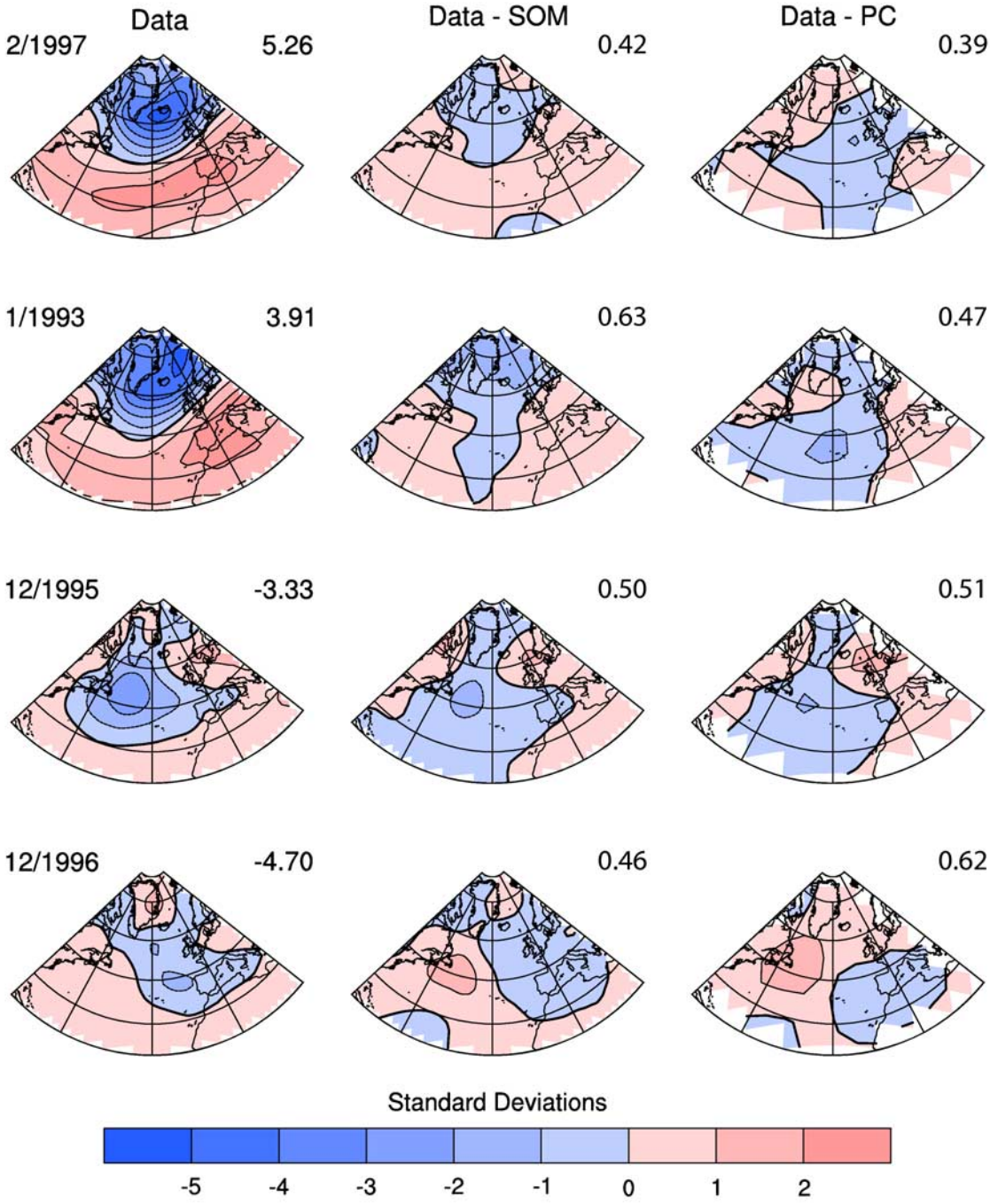

Figure 11. As in Figure 10 but showing residuals between data and analysis patterns. Columns are the original data (first column), the difference between the data and its best matching generalized SOM pattern (second column), and the difference between the data and its corresponding PC1 pattern (third column). In addition, the RMS (root mean square) of the residual (in standard deviations) is shown to the upper right of each pattern in each difference column.

High unrelated to the Icelandic Low (''non-NAO variability''), at least for the below-normal anomalies of the latter. In short, SOM results appear to more cleanly capture the decoupling of the Azores High from the joint behavior with the Icelandic Low (also seen in work with other tools such as *Cassou et al.* [2004]) when compared to the PCA results. It is beyond the scope of this paper to confirm this, but a

Table 3. RMS Statistics for Residuals Between Data and Corresponding SOM and PC1 Patterns

	Minimum	Maximum	Mean	Standard Deviation	Range
SOM	0.22	0.87	0.50	0.12	0.64
PC1	0.27	1.1	0.65	0.18	0.83

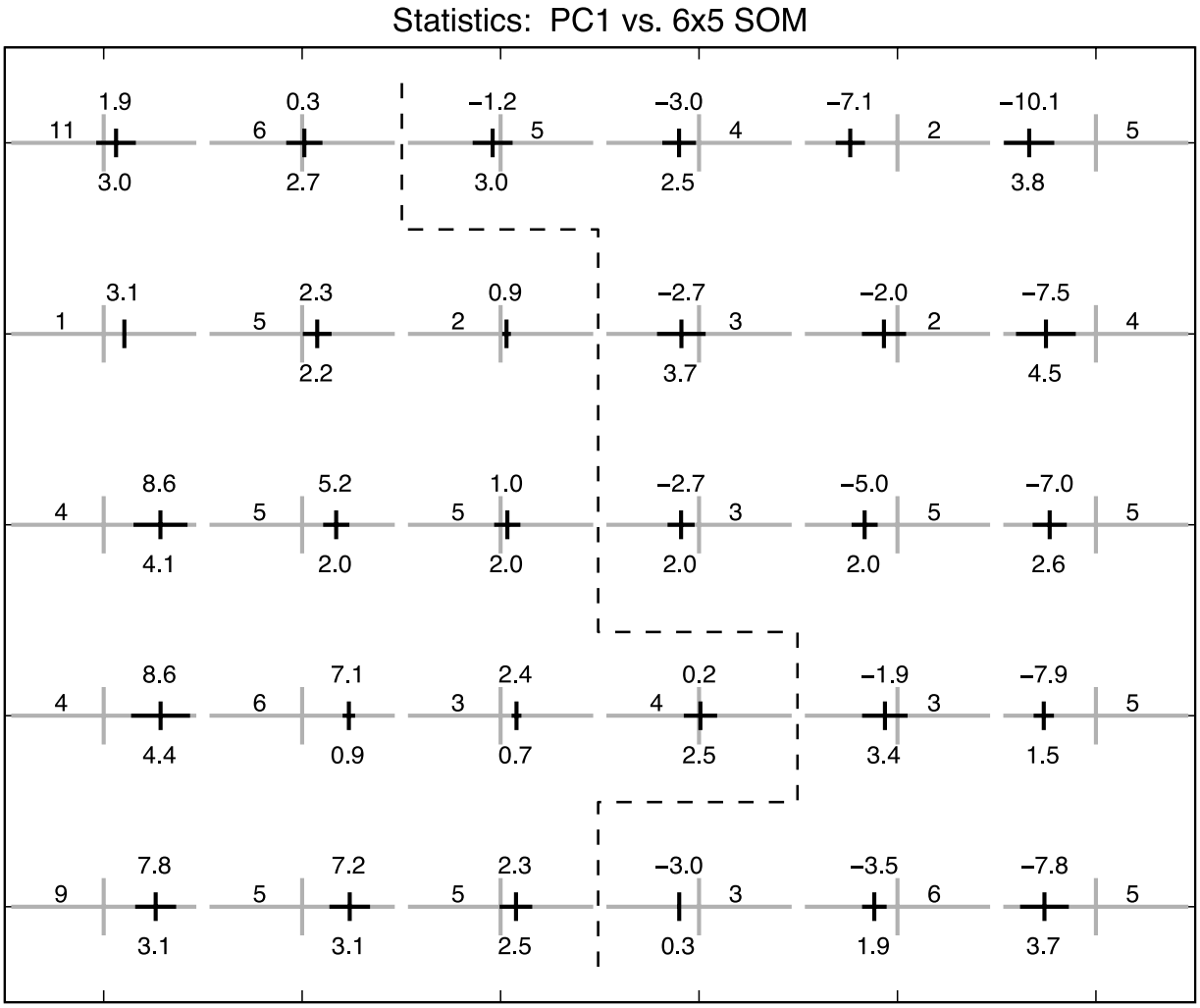

Figure 12. Statistics of PC1 for each SOM pattern. Description and values as in Figure 4 but for PC1.

comprehensive comparison of SOMs and PCA on this data set should include repeating the analyses with orthogonally and/or obliquely rotated PCA. (Note that comparisons of SOMs to PCA using synthetic data have recently been done by *Reusch et al.* [2005a], in the North Atlantic domain used here, and by *Liu et al.* [2006b], who also include observational oceanographic data.)

4.5. Other Nonlinear Studies

[42] It is not our intent to do a comprehensive comparison with all other nonlinear-tool-based studies of North Atlantic variability, but it is worth noting that, for example, the cluster analysis of *Cassou et al.* [2004] reaches many similar conclusions. For example, the RDG mode (a strong anticyclonic ridge [*Cassou et al.*, 2004, Figure 1c]) resembles the positive mode of the patterns we describe as "mid-North Atlantic-only" (pattern 1 in Figure 1). However, the below-normal "phase" of this pattern (pattern 30 in Figure 1) is not readily seen in the cluster analysis results. However, the cluster analysis does capture, at least in part, the changes in position and extent of the pressure systems between positive and negative NAO conditions. Thus, while cluster analysis is clearly useful for this data set, the SOM-based approach has a number of advantages (including reduced subjectivity) that make it more attractive from a usability point of view [e.g., *Hewitson and Crane*, 2002].

5. Conclusion

[43] The SOM-based approach to studying North Atlantic variability in general, and the NAO in particular, has great promise. Results from this boreal winter, monthly MSLP study readily identify both the canonical variability of the NAO and some of the nonlinear aspects of its behavior. Further work should lead to new calibrations of regional climate proxies (e.g., Greenland ice cores) to the North Atlantic climate record and a better understanding of the climate in this important region.

[44] In particular, four main points can be drawn from this work:

[45] 1. The SOM-based analysis of monthly MSLP decomposes variability into patterns representing NAO extrema, a "mid-North Atlantic-only" pattern reflecting variability in the Azores High independent of the Icelandic Low, and intermediate patterns reflecting smooth transitions between these climate states.

[46] 2. Frequency analyses of the data grouped by month show that the recent trend toward higher NAO index states

is coming primarily from trends during January (since 1972–1987) and February (since 1987–2002). The frequency maps also show notable differences in variability patterns between the months.

[47] 3. High (low) NAO indices are associated with a variety of high- (low-) state spatial patterns in the SOM. Standard two-points-in-space indices are not expected to capture such spatial variability but the SOM makes identifying it much easier.

[48] 4. The PCA and SOM results both capture "North Atlantic variability" but do so in different ways, leading to different perspectives on the data. Comparison of PCA to SOM results suggests that the orthogonality constraint of standard PCA may be leading to a mixing of patterns that the SOM is able to extract more readily. Further, the SOM-based approach has a distinct advantage in its more readily accessible visualization aspects (e.g., a SOM grid plot summarizes all of a data set's variability in a page).

[49] The work reported here has only involved MSLP but preliminary work with other single-variable analyses has shown similarly encouraging results. Further, we expect additional insights when multiple variables are analyzed together using the SOM-based methods. For example, our analysis of the MSLP mean and standard deviation showed how these aspects of the pressure field varied jointly. Similarly, joint surface and upper air analyses will help to show the broader behavior of the full atmosphere. This will help with, for example, understanding variability in storm tracks as seen in pressure, moisture and wind fields throughout the atmosphere.

[50] Longer term, integration of our work on the atmosphere with high-resolution ice core climate proxy records (e.g., major ions, trace metals, accumulation) from Greenland is an essential next step. Results will help improve our understanding of the proxies (e.g., how they record climate) and yield an extended record of climate in the North Atlantic through artificial neural network-based calibration [*Reusch et al.*, 2005b] of the atmosphere to the ice core records of past climate. Given the extensive body of work on the NAO and North Atlantic variability, it is likely that not all these results will be entirely novel. Even so, it is still of interest that an independent analysis path was able to confirm prior knowledge. That which is different will add to our understanding of this system.

[51] **Acknowledgments.** This work was supported by NSF grant ATM 04-25592 to D. B. Reusch and R. B. Alley. We are also grateful to NCAR's Visualization and Enabling Technologies Section, Scientific Computing Division, for their tireless support of NCL (the NCAR Command Language).

References

Alley, R. B., and K. M. Cuffey (2001), Oxygen- and hydrogen-isotopic ratios of water in precipitation: Beyond paleothermometry, in *Stable Isotope Geochemistry, Reviews in Mineralogy and Geochemistry*, vol. 43, edited by J. W. Valley and D. Cole, pp. 527–553, Mineral. Soc. Am., Chantilly, Va.

Appenzeller, C., T. F. Stocker, and M. Anklin (1998), North Atlantic oscillation dynamics recorded in Greenland ice cores, *Science*, *282*(5388), 446–449.

Armstrong, R. L., and M. J. Brodzik (1995), An earth-gridded SSM/I data set for cryospheric studies and global change monitoring, *Adv. Space Res.*, *10*, 155–163.

Barnston, A. G., and R. E. Livezey (1987), Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Mon. Weather Rev.*, *115*(6), 1083–1126.

Barry, R. G., and A. M. Carleton (2001), *Synoptic and Dynamic Climatology*, 620 pp., Routledge, Boca Raton, Fla.

Cassano, E. N., A. H. Lynch, J. J. Cassano, and M. R. Koslow (2006a), Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA, *Clim. Res.*, *30*(2), 83–97.

Cassano, J. J., O. Uotila, and A. Lynch (2006b), Changes in synoptic weather patterns in the polar regions in the 20th and 21st centuries, part 1: Arctic, *Int. J. Climatol.*, *26*(8), 1027–1049, doi:10.1002/joc.1306.

Cassou, C., L. Terray, J. W. Hurrell, and C. Deser (2004), North Atlantic winter climate regimes: Spatial asymmetry, stationarity with time, and oceanic forcing, *J. Clim.*, *17*(5), 1055–1068.

Cattell, R. B. (1952), *Factor Analysis: An Introduction and Manual for the Psychologist and Social Scientist*, 462 pp., HarperCollins, New York.

Cavazos, T. (1999), Large-scale circulation anomalies conducive to extreme events and simulation of daily rainfall in northeastern Mexico and southeastern Texas, *J. Clim.*, *12*(5), 1506–1523.

Cavazos, T. (2000), Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans, *J. Clim.*, *13*(10), 1718–1732.

Cavazos, T., A. C. Comrie, and D. M. Liverman (2002), Intraseasonal variability associated with wet monsoons in southeast Arizona, *J. Clim.*, *15*(17), 2477–2490.

Cohen, S. J. (1983), Classification of 500 mb height anomalies using obliquely rotated principal components, *J. Appl. Meteorol.*, *22*(12), 1975–1988.

Crane, R. G., and B. C. Hewitson (1998), Doubled CO_2 precipitation changes for the Susquehanna basin: Down-scaling from the GENESIS general circulation model, *Int. J. Climatol.*, *18*, 65–76.

Cullen, H. M., R. D. D'Arrigo, E. R. Cook, and M. E. Mann (2001), Multiproxy reconstructions of the North Atlantic Oscillation, *Paleoceanography*, *16*(1), 27–39.

Deser, C. (2000), On the teleconnectivity of the "Arctic Oscillation," *Geophys. Res. Lett.*, *27*(6), 779–782.

Gibson, J. K., M. Fiorino, A. Hernandez, P. Kållberg, X. Li, K. Onogo, S. Saarinen, and S. Uppala (1999), The ECMWF 40 year re-analysis (ERA-40) project—Plans and current status, in *10th Global Change Studies*, pp. 369–372, Am. Meteorol. Soc., Boston, Mass.

Hewitson, B. C., and R. G. Crane (2002), Self-organizing maps: applications to synoptic climatology, *Clim. Res.*, *22*(1), 13–26.

Hewitson, B. C., and R. G. Crane (2005), Gridded area average precipitation via conditional interpolation, *J. Clim.*, *18*(1), 41–57, doi:10.1175/JCLI3246.1.

Hurrell, J. W. (1995), Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation, *Science*, *269*, 676–679.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck (Eds.) (2003), *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr. Ser.*, vol. 134, edited by J. W. Hurrell et al., 279 pp., AGU, Washington, D. C.

Jolliffe, I. T. (2002), *Principal Component Analysis*, 487 pp., Springer, New York.

Jones, P. D., T. Jonsson, and D. Wheeler (1997), Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland, *Int. J. Climatol.*, *17*, 1433–1450.

Kohonen, T. (2001), *Self-Organizing Maps*, 501 pp., Springer, New York.

Kutzbach, J. E. (1967), Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America, *J. Appl. Meteorol.*, *6*(5), 791–802.

Lamb, P. J., and R. A. Peppler (1987), North Atlantic Oscillation: Concept and an application, *Bull. Am. Meteorol. Soc.*, *68*(10), 1218–1225.

Liu, Y., and R. H. Weisberg (2005), Patterns of ocean current variability on the West Florida Shelf using the self-organizing map, *J. Geophys. Res.*, *110*, C06003, doi:10.1029/2004JC002786.

Liu, Y., R. H. Weisberg, and R. He (2006a), Sea surface temperature patterns on the West Florida Shelf using growing hierarchical self-organizing maps, *J. Atmos. Oceanic Technol.*, *23*(2), 325–338.

Liu, Y., R. H. Weisberg, and C. N. K. Mooers (2006b), Performance evaluation of the self-organizing map for feature extraction, *J. Geophys. Res.*, *111*, C05018, doi:10.1029/2005JC003117.

Lorenz, E. N. (1956), Empirical orthogonal functions and statistical weather prediction, in *Statistical Forecasting Project*, *Sci. Rep. 1*, 49 pp., Dep. of Meteorol., Mass. Inst. of Technol., Cambridge, Mass.

Lynch, A., O. Uotila, and J. J. Cassano (2006), Changes in synoptic weather patterns in the polar regions in the 20th and 21st centuries, part 2: Antarctic, *Int. J. Climatol.*, *26*(9), 1181–1199, doi:10.1002/joc.1305.

Marshall, J., et al. (2001), North Atlantic climate variability: Phenomena, impacts and mechanisms, *Int. J. Climatol.*, *21*, 1863–1898, doi:10.1002/joc.693.

Michelangeli, P.-A., R. Vautard, and B. Legras (1995), Weather regimes: Recurrence and quasi stationarity, *J. Atmos. Sci.*, *52*(8), 1237–1256.

Ogi, M., K. Yamazaki, and Y. Tachibana (2003), Solar cycle modulation of the seasonal linkage of the North Atlantic Oscillation (NAO), *Geophys. Res. Lett.*, *30*(22), 2170, doi:10.1029/2003GL018545.

Reusch, D. B., R. B. Alley, and B. C. Hewitson (2005a), Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data, *Polar Geogr.*, *29*(3), 227–251.

Reusch, D. B., B. C. Hewitson, and R. B. Alley (2005b), Towards ice core-based synoptic reconstructions of West Antarctic climate with artificial neural networks, *Int. J. Climatol.*, *25*(5), 581–610, doi:10.1002/joc.1143.

Richman, M. B. (1981), Obliquely rotated principal components: An improved meteorological map typing technique?, *J. Appl. Meteorol.*, *20*(10), 1145–1159.

Richman, M. B. (1986), Rotation of principal components, *J. Climatol.*, *6*, 293–335.

Rogers, J. C. (1984), The association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere, *Mon. Weather Rev.*, *112*(10), 1999–2015.

Sammon, J. W. (1969), A nonlinear mapping for data structure analysis, *IEEE Trans. Comput.*, *C-18*(5), 401–409.

Smith, T. M., R. W. Reynolds, R. E. Livezey, and D. C. Stokes (1996), Reconstruction of historical sea surface temperatures using empirical orthogonal functions, *J. Clim.*, *9*(6), 1403–1420.

Stephenson, D. B., H. Wanner, S. Bronniman, and J. Luterbacher (2003), The history of scientific research on the North Atlantic Oscillation, in *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr. Ser.*, vol. 134, edited by J. W. Hurrell et al., pp. 37–50, AGU, Washington, D. C.

Stocker, T. F. (1999), Abrupt climate changes: From the past to the future—A review, *Int. J. Earth Sci.*, *88*, 365–374.

Tennant, W. (2004), Considerations when using pre-1979 NCEP/NCAR reanalyses in the southern hemisphere, *Geophys. Res. Lett.*, *31*, L11112, doi:10.1029/2004GL019751.

Thompson, D. W. J., and J. M. Wallace (2000), Annular modes in the extratropical circulation. Part I: Month-to-month variability, *J. Clim.*, *13*(5), 1000–1016.

Thompson, D. W. J., and J. M. Wallace (2001), Regional climate impacts of the Northern Hemisphere annular mode, *Science*, *293*(5527), 85–89.

Thompson, D. W. J., S. Lee, and M. P. Baldwin (2003), Atmospheric processes governing the Northern Hemisphere annular mode/North Atlantic Oscillation, in *The North Atlantic oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr. Ser.*, vol. 134, edited by J. W. Hurrell et al., pp. 81–112, AGU, Washington, D. C.

Wallace, J. M. (2000), North Atlantic Oscillation/annular mode: Two paradigms–one phenomenon, *Q. J. R. Meteorol. Soc.*, *126*(564), 791–805.

Wallace, J. M., and D. S. Gutzler (1981), Teleconnections in the geopotential height field during the Northern Hemisphere winter, *Mon. Weather Rev.*, *109*, 784–812.

Walsh, J. E. (1978), Temporal and spatial scales of the Arctic circulation, *Mon. Weather Rev.*, *106*(11), 1532–1544.

Wanner, H., S. Bronnimann, C. Casty, D. Gyalistras, J. Luterbacher, C. Schmutz, D. B. Stephenson, and E. Xoplaki (2001), North Atlantic Oscillation—Concepts and studies, *Surv. Geophys.*, *22*(4), 321–382.

R. B. Alley, Department of Geosciences and Earth and Environmental Systems Institute, College of Earth and Mineral Sciences, Pennsylvania State University, University Park, PA 16802, USA.

B. C. Hewitson, Department of Environmental and Geographical Sciences, University of Cape Town, Private Bag, Rondebosch 7701, South Africa.

D. B. Reusch, Earth and Environmental Systems Institute, College of Earth and Mineral Sciences, Pennsylvania State University, University Park, PA 16802, USA. (dbr@geosc.psu.edu)