Progress in Physical Geography
35(1) 109–119

DOI: 10.1177/0309133310397582
ppg.sagepub.com
SAGE

The self-organizing map in synoptic climatological research

Scott C. Sheridan
Kent State University, USA

Cameron C. Lee
Kent State University, USA

Abstract

Self-organizing maps (SOMs) are a relative newcomer to synoptic climatology; the method itself has only been utilized in the field for around a decade. In this article, we review the major developments and climatological applications of SOMs in the literature. The SOM can be used in synoptic climatological analysis in a manner similar to most other clustering methods. However, as the results from a SOM are generally represented by a two-dimensional array of cluster types that 'self-organize', the synoptic categories in the array effectively represent a continuum of synoptic categorizations, compared with discrete realizations produced through most traditional methods. Thus, a larger number of patterns can be more readily understood, and patterns, as well as transitional nodes between patterns, can be discerned. Perhaps the most intriguing development with SOMs has been the new avenues of visualization; the resultant spatial patterns of any variable can be more readily understood when displayed in a SOM. This improved visualization has led to SOMs becoming an increasingly popular tool in various research with climatological applications from other disciplines as well.

Keywords

applied climatology, self-organizing map, SOM, synoptic climatology

I Introduction

Synoptic climatology was defined by Yarnal (1993) as directly linking the atmospheric circulation and an environmental response. To a synoptic climatologist, the basis for such linkages most often involves a successful classification of atmospheric conditions into one of a number of different states, often termed *patterns*, *weather types*, *air masses*, or *clusters*, among other terms, depending upon what is being classified and the preference of the researcher (Huth et al., 2008). Though not always the case, the most common synoptic classifications involve categorization of cases at the *daily* to *monthly* temporal level, with either a suite of variables at a *point* level or a single variable at a *regional* spatial level. To define classifications, one must first assume that the states of the atmosphere can validly be partitioned into such intervals, that each state will represent some relatively homogeneous set of cases, and that, fundamentally, these partitions represent a full complement of all relevant atmospheric states.

Corresponding author:
Scott C. Sheridan, Department of Geography, Kent State University, 443 McGilvrey Hall, Kent, OH 44242, USA
Email: ssherid1@kent.edu

Over the course of time the original *manual* map classifications, usually of surface pressure patterns (e.g. Muller, 1977), have evolved into a highly varied number of classification techniques of any of a number of atmospheric variables. Modern techniques are generally described as *objective*, to contrast against the clear subjectivity of manual methods, although Huth et al. (2008) among others note that the term *objective* can be misleading, as even with any mathematical classification, a number of subjective decisions still need to be made, and thus suggest *computer-assisted* may be a more appropriate term. Further, some methods, such as the Spatial Synoptic Classification (Sheridan, 2002) or the objectivized Grosswetterlagen (James, 2007) can be described as *hybrid* or *mixed* methods, involving some manual initial identification of different states.

Over the past 20 to 30 years, one of the most commonly utilized methods in synoptic climatology can be broadly termed as cluster analysis, and is usually preceded by a principal component analysis (PCA). The PCA step effectively serves as a data reduction technique (given the high levels of spatiotemporal collinearity in climate data sets), and the cluster analysis – which can be accomplished via a number of different methods (Barry and Perry, 2001; Yarnal, 1993) – groups the resulting principal components into classifications. This method, while utilized in myriad applications and easily replicable in statistical packages, involves a number of human inputs – from the method of clustering, to the number of principal components or clusters retained – all of which can significantly affect resulting classifications (e.g. Cuell and Bonsal, 2009; Kalkstein et al., 1987; Yarnal, 1993).

Aside from these considerations, one general shortcoming of many traditional synoptic classifications is that while they represent discrete realizations of an atmospheric system, they generally cannot be organized into a continuum. The self-organizing map (SOM), dating back in concept to the late 1980s and first presented in significant detail in Kohonen (1995), may solve this shortcoming. Though not the first to utilize the method (e.g. Cavazos, 1999, 2000), it is Hewitson and Crane (2002) which served as the seminal introduction to the concept of self-organizing maps in their utility to synoptic climatology, and in the years since, SOMs have become increasingly popular in the discipline. In this progress report, we first outline some of the fundamental aspects of SOMs, followed by a review of their applications to date across the field, how they compare with existing techniques, and their potential benefits to the discipline.

II Practical matters of the SOM

A wealth of information on the SOM can be found at the SOM page hosted by the Helsinki University of Technology's Laboratory of Computer and Information Science (http://www.cis.hut.fi/research/som-research), including public domain software for producing a SOM.

The SOM methodology utilizes a neural-network algorithm to determine and display the distribution function of a multidimensional data set. It accomplishes this by creating an *array* or *lattice* (the SOM or master SOM) that is generally a two-dimensional matrix of *nodes*, which for the purposes of synoptic climatology can be thought of as analogous to clusters. Because these nodes span the entire data space of the input data, the SOM methodology involves no a priori assumptions about the distribution of the data (Hewitson and Crane, 2002). A map with a certain user-selected number of nodes is initialized with random values (e.g. Cassano et al., 2006a; Hewitson and Crane, 2002), or eigenvectors of the data set (e.g. Gutiérrez et al., 2005). Each input vector (case) is added to the SOM, and a 'winning' node is determined by virtue of identifying the node whose Euclidean distance from the input vector is the lowest (Hewitson and Crane, 2002). This node is then modified through a learning-rate parameter, effectively nudging it in the direction of the new

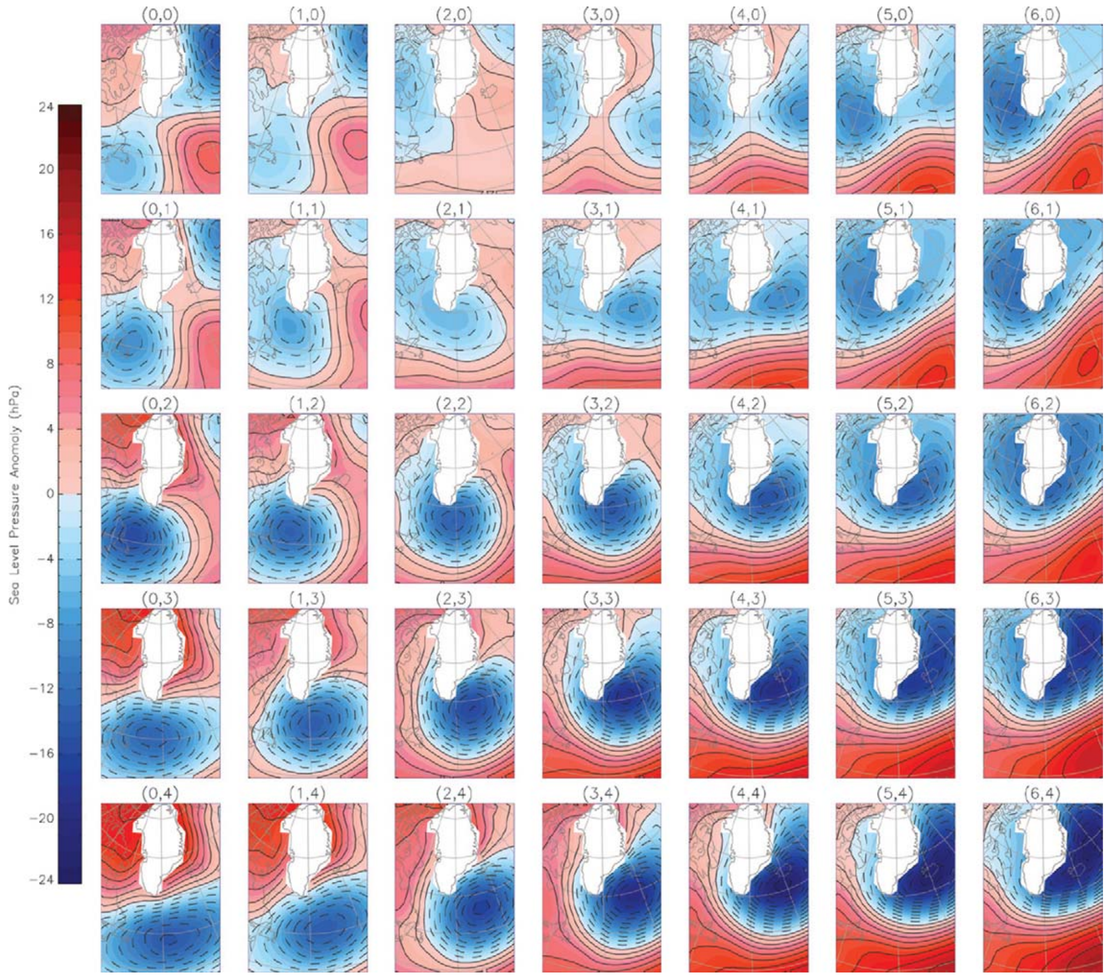

Figure 1. Example SOM of SLP anomalies based on ERA-40 SLP data from 1961 to 1999.
Source: Schuenemann et al. (2009)

case. The learning rate can be altered, as higher values yield a faster but potentially unstable learning process, increasing the potential for an unrealistic SOM, while lower values yield a smoother process (Gutiérrez et al., 2005). What is unique about SOMs is that, in addition to modifying the winning node, the winning node's neighbors on the array are also modified, through a distance decay function (Hewitson and Crane, 2002). Kohonen (2001) points out that the SOM methodology is effectively the same as the k-means method, if the neighbor modification were to be removed.

Ultimately, through this spatialized pattern of adjustment, the array of nodes 'self-organizes' into a cohesive pattern, with more similar nodes in closer proximity, and more dissimilar nodes farther away. The four corners of the SOM can thus be thought of as the most extreme nodes in terms of climate variability, with a smooth continuum in between. For instance (Figure 1), in a typical SOM with sea-level pressure (SLP) values over a study area, high-pressure dominant patterns will align on one side, while low-pressure dominant patterns will align on the other (e.g. Brown et al., 2010; Cassano and Cassano, 2010;

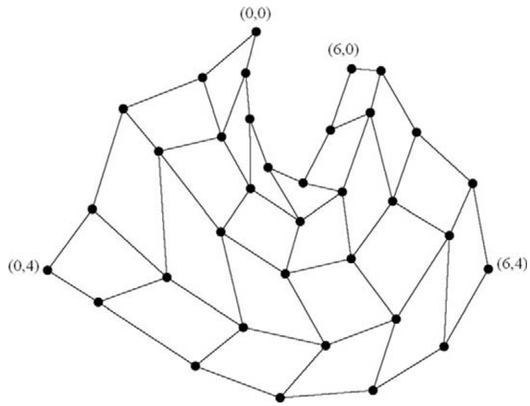

Figure 2. Sammon map for the SOM in Figure 1. Corner nodes are labeled.
Source: Schuenemann et al. (2009)

Cassano et al., 2006b; Schuenemann et al., 2009). Similarly, an oscillation such as the El Niño-Southern Oscillation (ENSO) will effectively be partitioned into opposite corners of the SOM (e.g. Leloup et al., 2007), with transitional phases in between. Further, by virtue of this learning process, the probability density is preserved, so that more nodes are placed in areas of high data density (i.e. more common weather situations), and outlier cases are less likely to be merged into an unrepresentative cluster (Nicholls et al., 2010). Additionally, the SOM methodology is more robust in handling missing values, as they can be interpolated within a SOM (Hewitson and Crane, 2002; Richardson et al., 2003). The ultimate result of the SOM process can be described as a 'flexible lattice (of nodes) folding on to the cloud formed by the data' (Gutiérrez et al., 2005). A Sammon map (Sammon, 1969), in which multidimensional vectors can be rendered in two-dimensional (2-D) space, with distances effectively relating the level of dissimilarity between different nodes, helps visualize this lattice. In Figure 2, for example, it can be seen that the nodes along the top row are more similar than the nodes across the bottom. An unstable learning process could also be discerned from a Sammon map that folds over on itself. Plots arranged in the lattice format, either maps of the mean fields or of diagnostic values such as frequency or occurrence of another event (e.g. precipitation), can also improve interpretability over traditional synoptic techniques.

Implementing the SOM requires several separate decisions, such as the learning rate and the radius in which adjacent nodes are modified. The initial selection of the values of these parameters, where discussed, is based on trial and error in many cases (e.g. Schuenemann et al., 2009). As a SOM iterates, and final nodes are determined, both the learning rate and radius are typically linearly decreased (Tozuka et al., 2008). Generally, the entire process is repeated twice, first with relatively large values to generalize the SOM lattice, and then with a smaller radius to refine the nodes in areas of high data density (Crane and Hewitson, 2003). Once a stable solution is reached, a final classification is performed, in which each case can be assigned to the node it most closely resembles. In some studies, it is suggested that the differences in the final SOM as one varies parameters such as the learning rate and radius are relatively minimal, and thus the final solution may be quite robust regardless of these subjective decisions (e.g. Cassano et al., 2006b; Johnson et al., 2008), although in other cases this does not appear to be so (Schuenemann et al., 2009). In most articles, unfortunately, little or no discussion is given to the exact choice of these parameters.

Many of the other user decisions with SOMs are similar to all synoptic-typing methods, such as which variable to map and its domain, whether to standardize or normalize the data, and how to address seasonality. Predictably, a decision which has received ample attention across SOM literature involves the number of nodes to create. Although some studies have determined the number of nodes to retain via a metric such as an elbow criterion, in which the number of nodes to retain is based on the point after which additional nodes 'fail to add a significant amount of information' (Schuenemann et al., 2009), most studies have generally utilized a more subjective methodology. In many cases, several different quantities of nodes are produced (e.g. 2×3, 4×3, and 7×5 arrays in Leloup et al., 2007), and then the resultant nodes are evaluated in each case, to determine which quantity most sufficiently accounts for the climatic variability across the study area, or which helps stratify the analysis variable best. For instance, Tymvios et al. (2010) tested several different array sizes, and ultimately chose a 36-node array of 500-mb geopotential heights for analysis after showing that this size array best delineated among the precipitation events that were the focus of their research. Reusch et al. (2007) has suggested the array not be symmetric for the stability of the calculations, although symmetric arrays have been presented in some articles (e.g. Tozuka et al., 2008). Generally, because of the spatial cohesiveness of the lattice structure of nodes, and the resulting ease in interpretability, most studies that utilize SOMs retain a greater number of nodes than is typically seen with traditional synoptic methods. Once the ideal number of nodes is chosen, a number of different SOMs of same size, but with varying initial parameters (i.e. learning rate, update radius, and number of iterations), are then compared; the ultimate goal generally is to select the SOM that minimizes the average Euclidean distance between the SOM and the input data, or the 'quantization error' (Schuenemann et al., 2009).

In other research, a two-part clustering is performed. First, a number of nodes are created that well exceeds typical synoptic climatological work, such as the 273-node (21×13) SOM utilized in Guèye et al. (2010). These nodes themselves may then be clustered, using a hierarchical agglomerative cluster (HAC) analysis such as Ward's (e.g. Guèye et al., 2010) to determine the optimal number of node clusters. In other cases, a more subjective approach is used, such as Finnis et al. (2009a) in which 48 nodes are clustered into seven 'subregions' of the SOM based on subjective grouping around the dominant pressure centers. By virtue of the properties of the SOM, these clusters are composed of spatially contiguous nodes across the array, further enhancing interpretability. Interestingly, when a data distribution is highly non-linear and/or bimodal, nodes may end up being placed in regions of significant data paucity; in some cases 'empty nodes' may actually result where no cases are ultimately assigned to a node. Reusch et al. (2005b) created a SOM based on mean seasonal anomalies of 700-mb temperature, geopotential height, and humidity over Antarctica for 15 winters, and analyzed the winters using a 5×3 array; only seven of the 15 nodes had at least one winter mapped to it.

III Literature review of uses of the SOM

SOMs have appeared throughout the breadth of synoptic climatological literature. As a relative newcomer to the discipline, a number of articles focus on providing a demonstration of the SOM methodology and output, with the application to a particular research question as a secondary goal (e.g. Khedairia and Khadir, 2008; Michaelides et al., 2007; Richardson et al., 2003). Good overviews of the methodology can be found in Cassano et al. (2006b), Gutiérrez et al. (2005), Hewitson and Crane (2002), Johnson et al. (2008), Sang et al. (2008), and Schuenemann et al. (2009).

One synoptic climatological application that has seen increased attention over recent years has been the validation of the general circulation model (GCM; Sheridan and Lee, 2010), and the SOM method has appeared in much of this research (Brown et al., 2010; Cassano et al., 2006b; Finnis et al., 2009a, 2009b; Gutowski et al., 2004; Higgins and Cassano, 2010; Lynch et al., 2006; Skific et al., 2009a, 2009b; Tennant, 2003; Tozuka et al., 2008), with a majority of the studies focused on the Arctic and Antarctic. Some GCM validation studies using synoptic methods have focused on the assessment of model bias in terms of weather patterns created.

For instance, Cassano et al. (2006b) and Lynch et al. (2006) for the Arctic and Antarctic, respectively, compare the performance of a suite of GCMs at rendering the range of sea-level pressure (SLP) patterns shown in reanalysis data sets, with the ultimate goal of projecting future synoptic frequencies. GCMs varied considerably in their ability to reproduce historical synoptic types, as well as in terms of which features each GCM would over- or underemphasize. Ensembles of the GCMs were produced, which performed better, although still not without significant divergence from reanalysis values in certain cases.

A number of studies that apply a SOM to GCM output examine how models generate precipitation, in order to diagnose model difficulties in this area. In a pair of articles, Finnis et al. (2009a, 2009b) utilize SOMs to evaluate the performance of 14 GCMs in simulating synoptic conditions and precipitation generation across four Arctic river basins. Their work also identifies considerable variability among GCMs in their ability to adequately render synoptic types, although they generate a subset of models which performs better. Their methodology involves generating 48-node solutions, which were then subjectively grouped into seven (Mackenzie River Basin; Finnis et al., 2009a) and nine (Eurasian watersheds; Finnis et al., 2009b) clusters. In both papers, use of the lattice structure to display (and contour) frequencies, mean precipitation amounts, and model biases, by node, greatly enhances interpretability of model bias. Their results suggest that despite model deficiencies in simulating node frequencies, model precipitation bias is largely a function of simulated precipitation processes within the model, rather than discrepancies in synoptic type, an observation also noted by Hewitson and Crane (2006) and Brown et al. (2010), in their SOM analyses of South African and Australian precipitation, respectively.

SOM-based studies have also invoked GCMs with the goal of projecting future climate scenarios. Skific et al. (2009a, 2009b) evaluate moisture transport and convergence in the Arctic by creating a SOM based on SLP values which is based on an aggregate of past and future modeled values, to assess the general migration of pattern frequency over time (and then to assess the relative change in transport that is circulation-related). Hope (2006), in investigating projected future precipitation trends in Australia, and Steynor et al. (2009), in assessing future runoff for the Breede River in South Africa, both base their work on the SOM node-precipitation relationship. In both of these studies, however, the SOM is trained only on past observations; as a result, future synoptic patterns that do not resemble historical ones may not be properly represented as they will be classified into the corner node they most closely resemble, making results difficult to interpret.

Even when excluding GCM-based studies, the relationship between precipitation and atmospheric circulation is the most common focus of SOM-based articles. In two of the earliest climate applications, Cavazos (1999, 2000) utilize SOMs to feed a neural network model to predict localized precipitation amounts over the Texas-Mexico region and the eastern Balkans, respectively. Other papers have focused on quantifying the precipitation-synoptic type relationship, such as Verdon-Kidd and Kiem (2008), who use SOMs on monthly SLP values over Australia, and Schuenemann et al. (2009) and Cassano and Cassano (2010), who use SOMs on daily SLP values over Greenland and the Canadian Arctic, respectively. Tadross et al. (2005) apply a SOM to precipitation data to assess the variability and trends in the onset of the maize-growing season across southern Africa. Tymvios et al. (2010) utilize a SOM derived from 500-mb height fields to assess the likelihood of extreme precipitation events over Cyprus. Alexander et al. (2010) and Hope et al. (2006) both use long-term records of winter SLP data to dissect the relationship between the secular precipitation decline in Australia and SOM pattern-related trends. Gutiérrez et al. (2005) break down the SOM-precipitation relationship in Peru among El Niño phases, and utilize this information in testing methods for seasonal precipitation forecasting. Tennant and Hewitson (2002) and Tennant and Reason (2005) create SOMs based on barotropic kinetic energy values over the southern polar and mid-latitude regions to relate circulation to precipitation patterns across Australia and South Africa.

Three articles take the SOM-precipitation relationship further, to understand the evolution of a seasonal monsoon. Chattopadhyay et al. (2008) study the Indian intra-seasonal monsoon with 3×3 and 9×9 SOMs that are created from a field of six different indices comprised of pressure, moisture, wind, and geopotential height values. The lower-order SOM is used to isolate primary states, while the higher-order SOM is able to identify subsets of these states, and could serve as a forecast tool for a medium-range prediction of monsoonal rainfall. Guèye et al. (2010) looks at phase transitions within the monsoon in Senegal using a large SOM array that is further subdivided into nine classes, while Cavazos et al. (2002) utilize a 3×5 SOM to assess variability in forcing mechanisms relative to the strength and pattern of the Arizona monsoon.

Several authors have also utilized SOMs to analyze teleconnections to better understand phase transitions. Tozuka et al. (2008) create a 7×7 SOM from monthly sea-surface temperature values to assess temporal and spatial trends in the Indian Ocean Dipole, as well as models' ability to reproduce it. Leloup et al. (2007) base a 10×10 SOM on multiple metrics of sea-surface temperature – clustered using a HAC into 12 subgroups – to assess the role of phase transitions of ENSO events. Reusch et al. (2007) study the North Atlantic Oscillation (NAO), using wintertime monthly SLP anomalies, to identify the causes of the trends in the teleconnection, and visualize the intermediate patterns between the opposite phases. Johnson et al. (2008) utilize the continuum provided by SOM nodes in analyzing the spatial and temporal trends in the NAO to assess the differences in circulation regimes between the pre-1978 and post-1978 periods. Later work by Johnson and Feldstein (2010) incorporates SOMs in studying North Pacific SLP patterns, with the goal of explaining the variability and relationship among several regional teleconnection indices with pattern frequencies. Maslanik et al. (2007) evaluate modes of the Arctic Oscillation (AO) using SOMs, and suggest that the decoupling of sea ice-AO trends over time may be due to a shift to nodes that resemble the AO, but would not be picked up by traditional calculation methods.

The diversity of the self-organizing map is exhibited in a number of other climatological studies with wider-ranging applications. Two studies have created SOMs derived from back trajectories of atmospheric flow in order to stratify ambient pollution levels: Kassomenos et al. (2007) for Athens, Greece, and Karaca and Camci (2010) for Istanbul, Turkey. Other studies have included SOMs in analyses of fire weather across the southwestern US (Crimmins, 2006), seasonal to decadal temperature differences in Melbourne, Australia (Nicholls et al., 2010), the distance-decay of precipitation correlation across Europe (Hofstra and New, 2009), Antarctic sea ice (Reusch and Alley, 2007; Reusch et al., 2005b), low-level jet patterns in Antarctica (Seefeldt and Cassano, 2008), and extreme temperature and wind events at Barrow, Alaska (Cassano et al., 2006a). MacKellar et al. (2010) assess the climatic response due to vegetation changes in southern Africa, by running a regional circulation model (RCM) and evaluating the results separately by SOM node.

Similar to an R-mode principal component analysis (as in Malmgren and Winter, 1999), several studies have clustered weather stations using SOMs to create climate regions. Crane and Hewitson (2003) insert the variance-covariance matrix of precipitation values among 104 stations, and latitude and longitude, to create a regionalization of precipitation climate zones across the eastern USA. Similar studies have been performed by Malmgren and Winter (1999) for Puerto Rico, though using principal component scores as determinants, Penlap et al. (2004) for Cameroon, using monthly anomalous precipitation values, and Thomas et al. (2007) for South Africa, using several precipitation metrics important for agriculture.

IV Summary

The SOM has significant promise for the field of synoptic climatology. Beyond the traditional benefits of all synoptic methods, the self-organization that is an inherent part of the process offers further advantages. A larger number of patterns can be more readily understood, since the nodes exist in a continuum, rather than as discrete realizations of patterns. This continuum enables understanding of phases as well as transitional nodes between phases, which can be helpful in understanding climate variability, and may be useful for forecasting mode transitions in features such as monsoon circulations. Perhaps the most novel development with SOMs has been the new avenues of visualization; patterns can be much more readily understood when displayed in the SOM array thus reducing the time needed to familiarize oneself with many distinct cluster types, a shortcoming of much traditional synoptic literature. Use of the lattice format to render analytical variables such as pattern frequency, node transition, or occurrence of some other event also improves interpretability. Several creative means of exploring climate variability via the SOM are also emerging, such as the dipole of ENSO in Leloup et al. (2007), the 423-node test SOM arranged as a spherical SOM in Sugimoto and Tachibana (2008), or the Growing Hierarchical SOM (GHSOM), capable of creating hierarchical node structures, in Liu et al. (2006).

To date, few studies have directly compared the efficacy of SOMs to other methods. Among those few studies, however, Reusch et al. (2005a, 2007) directly compare PCA patterns to SOMs, and discover that both methods have unique advantages, though the SOMs show improved visualization of data, are less likely to mix patterns than PCA, and thus more clearly depict 'blended' patterns when a SOM is expanded to include additional nodes. Michaelides et al. (2001) show that SOMs characterize precipitation variability better than traditional clustering methods. However, Kassomenos et al. (2007) discover differences in the classifications produced by SOMs, k-means, and a HAC in assessing atmospheric trajectories, though none of the three were significantly better at stratifying pollution levels.

While it is clear that SOMs are an excellent complement to current research methods, offering some distinct advantages over classical synoptic climatological techniques, many of the same disadvantages of these classifications are still present in the SOM methodology as well (e.g. the number of clusters to retain, seasonality, variables). Further, nodes created within a SOM are unique in synoptic climatology in that they are theoretical patterns based on the distribution of the data and not technically patterns based on a single case or a composite of actual cases (hence, the possibility of an 'empty' node). While this does not appear to have any significant impact on analysis, little has been done to assess specifically the impact this distinction may have, and more work comparing the efficacy of SOMs vis-a-vis other synoptic methodologies should continue to be undertaken.

Acknowledgements

We express our gratitude to the editor, Tim Warner, and one anonymous reviewer, for their helpful comments that improved this manuscript.

References

Alexander LV, Uotila P, Nicholls N, and Lynch A (2010) A new daily dataset for Australia and its application to the assessment of changes in synoptic patterns during the last century. *Journal of Climate* 23: 1111–1126.

Barry RG and Perry AH (2001) Synoptic climatology and its applications. In: Barry RG and Carleton AM (eds) *Synoptic and Dynamic Climatology*. London: Routledge: 547–603.

Brown JR, Jakob C, and Haynes JM (2010) an evaluation of rainfall frequency and intensity over the Australian region in a global climate model. *Journal of Climate*. doi: 10.1175/2010JCLI3571.1.

Cassano EN and Cassano JJ (2010) Synoptic forcing of precipitation in the Mackenzie and Yukon River basins. *International Journal of Climatology* 30: 658–674.

Cassano EN, Lynch AH, Cassano JJ, and Koslow MR (2006a) Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA. *Climate Research* 30: 83–97.

Cassano JJ, Uotila P, and Lynch A (2006b) Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part I: Arctic. *International Journal of Climatology* 26: 1027–1049.

Cavazos T (1999) Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *Journal of Climate* 12: 1506–1523.

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

Cavazos T, Comrie AC, and Liverman DM (2002) Intraseasonal variability associated with wet monsoons in southeast Arizona. *Journal of Climate* 15: 2477–2490.

Chattopadhyay R, Sahai AK, and Goswami BN (2008) Objective identification of nonlinear convectively coupled phases of monsoon intraseasonal oscillation: Implications for prediction. *Journal of the Atmospheric Sciences* 65: 1549–1569.

Crane RG and Hewitson BC (2003) Clustering and upscaling of station precipitation records to regional patterns using self-organizing maps (SOMs). *Climate Research* 25: 95–107.

Crimmins MA (2006) Synoptic climatology of extreme fireweather conditions across the southwest United States. *International Journal of Climatology* 26: 1001–1016.

Cuell C and Bonsall B (2009) An assessment of climatological synoptic typing by principal component analysis and kmeans clustering. *Theoretical and Applied Climatology* 98: 361–373.

Finnis J, Cassano J, Holland M, Serreze M, and Uotila P (2009a) Synoptically forced hydroclimatology of major Arctic watersheds in general circulation models; Part 1: The Mackenzie River Basin. *International Journal of Climatology* 29: 1226–1243.

Finnis J, Cassano, J, Holland M, Serreze M, and Uotila P (2009b) Synoptically forced hydroclimatology of major Arctic watersheds in general circulation models; Part 2: Eurasian watersheds. *International Journal of Climatology* 29: 1244–1261.

Guèye AK, Janicot S, Niang A, Sawadogo S, Sultan B, Diongue-Niang A, and Thiria S (2010) Weather regimes over Senegal during the summer monsoon season using self-organizing maps and hierarchical ascendant classification. Part I: Synoptic time scale. *Climate Dynamics*. doi: 10.1007/s00382-010-0782-6.

Gutiérrez JM, Cano R, Cofiño AS, and Sordo C (2005) Analysis and downscaling multi-model seasonal forecasts in Peru using self-organizing maps. *Tellus* 57A: 435–447.

Gutowski WJ, Otieno FO, Arritt RW, Takle ES, and Pan Z (2004) Diagnosis and attribution of a seasonal precipitation deficit in a USA regional climate simulation. *Journal of Hydrometeorology* 5: 230–243.

Hewitson BC and Crane RG (2002) Self-organizing maps: Applications to synoptic climatology. *Climate Research* 22: 13–26.

Hewitson BC and Crane RG (2006) Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa. *International Journal of Climatology* 26: 1315–1337.

Higgins ME and Cassano JJ (2010) Response of Arctic 1000 hPa circulation to changes in horizontal resolution and sea ice forcing in the Community Atmospheric Model. *Journal of Geophysical Research* 115: D17114.

Hofstra N and New M (2009) Spatial variability in correlation decay distance and influence on angular-distance weighting interpolation of daily precipitation over Europe. *International Journal of Climatology* 29: 1872–1880.

Hope PK (2006) Projected future changes in synoptic systems influencing southwest Western Australia. *Climate Dynamics* 26: 765–780.

Hope PK, Drosdowsky W, and Nicholls N (2006) Shifts in the synoptic systems influencing southwest Western Australia. *Climate Dynamics* 26: 751–764.

Huth R, Beck C, Philipp A, Demuzere M, Ustrnul Z, Cahynová M, et al. (2008) Classifications of atmospheric circulation patterns: Recent advances and applications. *Annals of the New York Academy of Sciences* 1146: 105–152.

James PM (2007) An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe. *Theoretical and Applied Climatology* 88: 17–42.

Johnson NC and Feldstein SB (2010) The continuum of North Pacific sea level pressure patterns: Intraseasonal, interannual, and interdecadal variability. *Journal of Climate* 23: 851–867.

Johnson NC, Feldstein SB, and Tremblay B (2008) The continuum of Northern Hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organizing maps. *Journal of Climate* 21: 6354–6371.

Kalkstein LS, Tan G, and Skindlov JA (1987) An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Applied Meteorology* 26: 717–730.

Karaca F and Camci F (2010) Distant source contributions to PM10 profile evaluated by SOM based cluster analysis of air mass trajectory sets. *Atmospheric Environment* 44: 892–899.

Kassomenos P, Vardoulakis S, Borge R, Lumbreras J, Papaloukas C, and Karakitsios S (2007) Comparison of statistical clustering techniques for the classification of modelled atmospheric trajectories. *Theoretical and Applied Climatology* 102: 1–2.

Khedairia S and Khadir MT (2008) Self-organizing map and k-means for meteorological day type identification for the region of Annaba, Algeria. *Proceedings, 7th Computer Information Systems and Industrial Management Applications*: 91–96.

Kohonen T (1995) *Self-organizing Maps*. Vol. 30, Springer Series in Information Sciences. Berlin: Springer.

Kohonen T (2001) *Self-Organizing Maps*, third edition. Berlin: Springer.

Leloup JA, Lachkar Z, Boulanger J-P, and Thiria S (2007) Detecting decadal changes in ENSO using neural networks. *Climate Dynamics* 28: 147–162.

Liu Y, Weisberg RH, and He R (2006) Sea surface temperature patterns on the West Florida Shelf using growing hierarchical self-organizing maps. *Journal of Atmospheric and Oceanic Technology* 23: 325–338.

Lynch A, Uotila P, and Cassano JJ (2006) Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part II: Antarctic. *International Journal of Climatology* 26: 1181–1199.

MacKellar N, Tadross M, and Hewitson B (2010) Synoptic-based evaluation of climate response to vegetation change over southern Africa. *International Journal of Climatology* 30: 774–789.

Malmgren BA and Winter A (1999) Climate zonation in Puerto Rico based on principal components analysis and an artificial neural network. *Journal of Climate* 12: 977–985.

Maslanik J, Drobot S, Fowler C, Emery W, and Barry R (2007) On the Arctic climate paradox and the continuing role of atmospheric circulation in affecting sea ice conditions. *Geophysical Research Letters* 34: L03711.

Michaelides SC, Liassidou F, and Schizas CN (2007) Synoptic classification and establishment of analogues with artificial neural networks. *Pure and Applied Geophysics* 164: 1347–1364.

Michaelides SC, Pattichis CS, and Kleovoulou G (2001) Classification of rainfall variability by using artificial neural networks. *International Journal of Climatology* 21: 1401–1414.

Muller RA (1977) A synoptic climatology for environmental baseline analysis: New Orleans. *Journal of Applied Meteorology* 16: 20–33.

Nicholls N, Uotila P, and Alexander L (2010) Synoptic influences on seasonal, interannual and decadal temperature variations in Melbourne, Australia. *International Journal of Climatology* 30: 1372–1381.

Penlap EK, Matulla C, von Storch H, and Mkankam FK (2004) Empirical downscaling in the tropics – possible changes of the little rainy season (March–June) in Cameroon. *Climate Research* 26: 85–96.

Reusch DB and Alley RB (2007) Antarctic sea ice: A self-organizing map-based perspective. *Annals of Glaciology* 46: 391–396.

Reusch DB, Alley RB and Hewitson BC (2005a) Relative performance of self-organizing maps and principle component analysis in pattern extraction from synthetic climatological data. *Polar Geography* 29: 188–212.

Reusch DB, Alley RB, and Hewitson BC (2007) North Atlantic climate variability from a self-organizing map perspective. *Journal of Geophysical Research* 112: D02104.

Reusch DB, Hewitson BC, and Alley RB (2005b) Towards ice core-based synoptic reconstructions of West Antarctic climate with artificial neural networks. *International Journal of Climatology* 25: 581–610.

Richardson AJ, Risien C, and Shillington FA (2003) Using self-organizing maps to identify patterns in satellite imagery. *Progress in Oceanography* 59: 223–239.

Sammon J (1969) A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers* 18: 401–409.

Sang H, Gelfand AE, Lennard C, Hegerl G, and Hewitson B (2008) Interpreting self-organizing maps through space-time data models. *The Annals of Applied Statistics* 2: 1194–1216.

Schuenemann KC, Cassano JJ and Finnis J (2009) Synoptic forcing of precipitation over Greenland: Climatology for 1961–99. *Journal of Hydrometeorology* 10: 60–78.

Seefeldt MW and Cassano JJ (2008) An analysis of low-level jets in the Greater Ross Ice Shelf Region based on numerical simulations. *Monthly Weather Review* 136: 4188–4205.

Sheridan SC (2002) The redevelopment of a weather type classification scheme for North America. *International Journal of Climatology* 22: 51–68.

Sheridan SC and Lee CC (2010) Synoptic climatology and the general circulation model. *Progress in Physical Geography* 34: 101–109.

Skific N, Francis JA, and Cassano JJ (2009a) Attribution of projected changes in atmospheric moisture transport in the Arctic: A self-organizing map perspective. *Journal of Climate* 22: 4135–4153.

Skific N, Francis JA, and Cassano JJ (2009b) Attribution of seasonal and regional changes in Arctic moisture convergence. *Journal of Climate* 22: 5115–5134.

Steynor AC, Hewitson BC, and Tadross MA (2009) Projected future runoff of the Breede River under climate change. *Water South Africa* 35: 433–440.

Sugimoto N and Tachibana K (2008) A first attempt to apply high speed spherical self-organizing map to huge climate datasets. *Scientific Online Letters on the Atmosphere* 4: 41–44.

Tadross MA, Hewitson BC, and Usman MT (2005) The interannual variability of the onset of the maize growing season over South Africa and Zimbabwe. *Journal of Climate* 18: 3356–3372.

Tennant WJ (2003) An assessment of intra-seasonal variability from 13-year GCM simulations. *Monthly Weather Review* 131: 1975–1991.

Tennant WJ and Hewitson BC (2002) Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *International Journal of Climatology* 22: 1033–1048.

Tennant WJ and Reason CJ (2005) Associations between the global energy cycle and regional rainfall in South Africa and Southwest Australia. *Journal of Climate* 18: 3032–3047.

Thomas C, Twyman C, Osbahr H, and Hewitson B (2007) Adaptation to climate change and variability: Farmer responses to intra-seasonal precipitation trends in South Africa. *Climatic Change* 83: 301–322.

Tozuka T, Luo J-J, Masson S, and Yamagata T (2008) Tropical Indian Ocean variability revealed by self-organizing maps. *Climate Dynamics* 31: 333–343.

Tymvios F, Savvidou K, and Michaelides SC (2010) Association of geopotential height patterns with heavy rainfall events in Cyprus. *Advances in Geosciences* 23: 73–78.

Verdon-Kidd D and Kiem AS (2008) On the relationship between large-scale climate modes and regional synoptic patterns that drive Victorian rainfall. *Hydrology and Earth System Sciences* Discussions 5: 2791–2815.

Yarnal B (1993) *Synoptic Climatology in Environmental Analysis*. London: Belhaven Press.