Self-Organizing Maps for the Investigation of Tornadic Near-Storm Environments

ALEXANDRA K. ANDERSON-FREY AND YVETTE P. RICHARDSON

The Pennsylvania State University, University Park, Pennsylvania

ANDREW R. DEAN, RICHARD L. THOMPSON, AND BRYAN T. SMITH

Storm Prediction Center, Norman, Oklahoma

(Manuscript received 17 March 2017, in final form 20 May 2017)

ABSTRACT

In this work, self-organizing maps (SOMs) are used to investigate patterns of favorable near-storm environmental parameters in a 13-yr climatology of 14 814 tornado events and 44 961 tornado warnings across the continental United States. Establishing nine statistically distinct clusters of spatial distributions of the significant tornado parameter (STP) in the 480 km × 480 km region surrounding each tornado event or warning allows for the examination of each cluster in isolation. For tornado events, distinct patterns are associated more with particular times of day, geographical locations, and times of year. For example, the archetypal springtime dryline setup in the Great Plains emerges readily from the data. While high values of STP tend to correspond to relatively high probabilities of detection (PODs) and relatively low false alarm ratios (FARs), the majority of tornado events occur within a pattern of uniformly lower STP, with relatively high FAR and low POD. Overall, the two-dimensional plots produced by the SOM approach provide an intuitive way of creating nuanced climatologies of tornadic near-storm environments.

1. Introduction

Given a supercell, which requires moderate-to-strong convective available potential energy (CAPE) and strong 0–6-km vector shear magnitude (SHR6; Brooks et al. 2003), the combination of a low-to-moderate lifting condensation level height (LCL) and strong 0–1-km storm-relative helicity (SRH1) has been shown to help distinguish between nontornadic and significantly tornadic cases, likely through their role in preventing the damping effects of overly strong cold pools and enhancing the lifting and contracting of near-surface air via stronger vertical perturbation pressure gradients (Markowski et al. 2003; Markowski and Richardson 2014). These four parameters are combined in composite indices such as the significant tornado parameter (STP; Thompson et al. 2003), which in its fixed-layer form is defined as STP = (MLCAPE/1000 J kg^{-1}) × [(2000 − MLLCL)/1500 m] × (SRH1/100 m^2 s^{-2}) × (SHR6/20 m s^{-1}), where the ML prefix denotes mixed-layer values of the parameters. Values of STP (collected near tornadic events using point soundings) greater than

1 have been shown (e.g., Thompson et al. 2012) to discriminate fairly well between supercellular storms capable of producing significant hail or wind but no tornadoes, and supercellular storms capable of producing tornadoes with a rating of 2 or greater on the enhanced Fujita scale.

The classification and examination of *two-dimensional* patterns in weather data has a long history; principal component analysis (PCA) has been used to create prototypical synoptic patterns for over 50 years (Christensen and Bryson 1966). More recently, Mercer et al. (2009) and Shafer et al. (2010) have made use of PCA to deduce information about the spatial structure of the near-storm environments for tornadic and nontornadic severe weather outbreaks; Mercer et al. (2012) made use of a rotated principal component analysis (RPCA) to create synoptic composites of tornadic and nontornadic severe weather outbreaks.

Self-organizing maps (SOMs; Kohonen 1982) are tools that have yet to be fully adopted by the severe storms research community, despite being a prominent technique in many studies of climate (see review by Liu and Weisberg 2011). A SOM is an artificial neural network that can be used as a clustering method in order to group together events with similar spatial structure in their

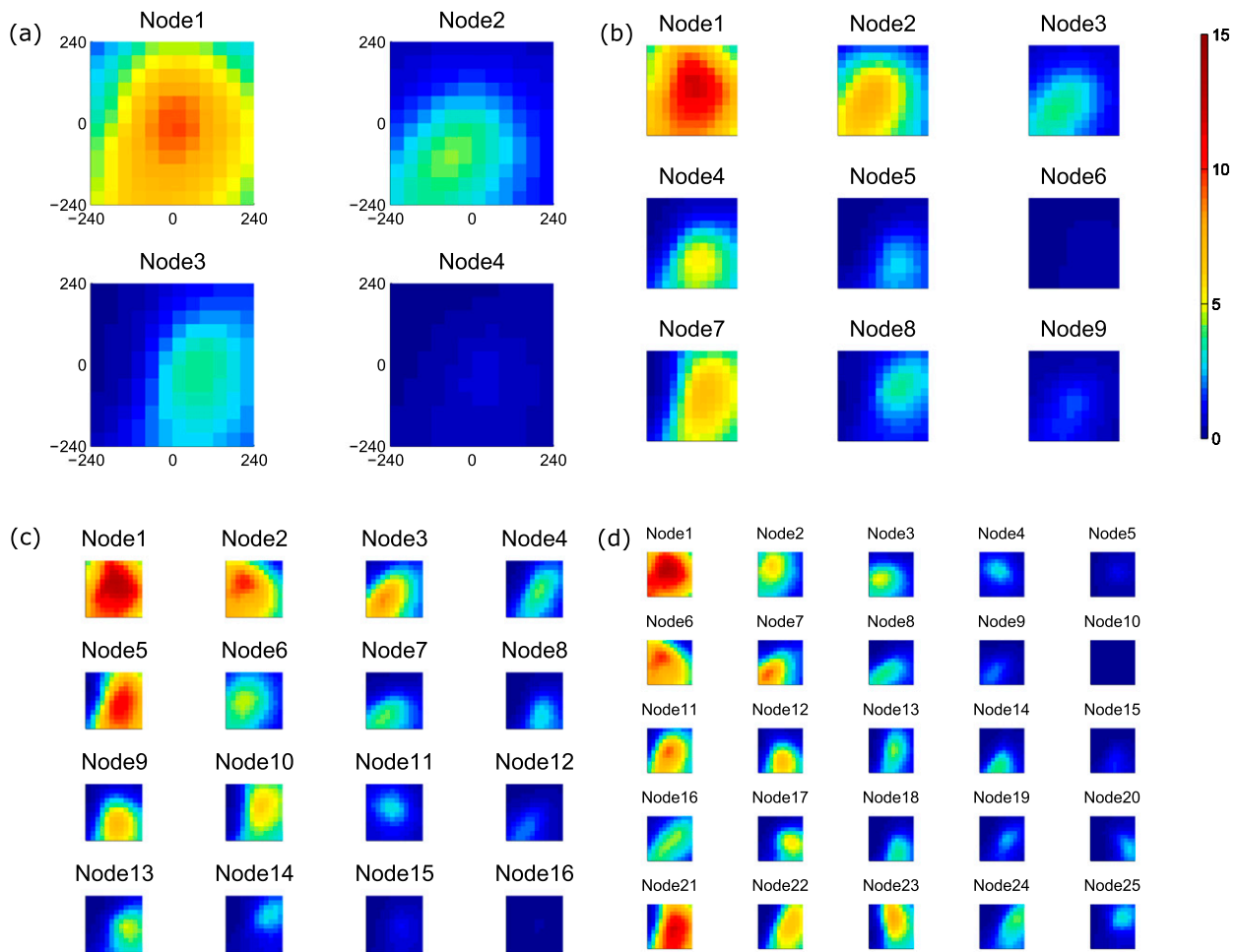*Corresponding author*: Alexandra K. Anderson-Frey, aka145@psu.edu

FIG. 1. Several possible choices for the number M of nodes: M = (a) 2×2, (b) 3×3, (c) 4×4, and (d) 5×5. Some axis labels and grids have been removed for the sake of clarity.

parameters; the statistics of each cluster [e.g., probability of detection (POD), the percentage of all tornadoes for which a warning was issued prior to the tornado's start time, and false alarm ratio (FAR), the percentage of all tornado warnings for which no tornado was reported within the warning boundaries over the duration of the warning] can then be calculated and assessed, effectively providing an example of, say, low-POD or high-FAR near-storm environments. For instance, Nowotarski and Jensen (2013) made use of 1185 SOM-clustered atmospheric soundings to characterize typical tornadic, nontornadic, and nonsupercellular soundings. Unlike PCAs, SOMs do not have any requirement of orthogonality and lend themselves well to the task of creating mesosynoptic composites of tornadic near-storm environments.

We begin with a detailed description of the dataset and methodology of this study in section 2. Section 3 provides a brief overview of the point values of near-storm

environmental parameters for tornado events and tornado warnings across the entire dataset, while section 4 makes use of SOM clustering to create and contrast the fully two-dimensional approach to the same problem for tornado events, and section 5 does the same for tornado warnings. Finally, section 6 concludes and summarizes our findings and provides guidance and motivation for future SOM-related studies by the severe storms community.

2. Methods and data

The tornado warning and event dataset is the same as was featured in Anderson-Frey et al. (2016): the study period of 2003–15 comprises 14 814 tornado events and 44 961 tornado warnings. The tornado event dataset was created by filtering county tornado segment data and keeping the highest (E)F-scale rating within an hour and a $40 \times 40 \, km^2$ area (Smith et al. 2012). Environmental data corresponding to each of these tornado

(a) Tornado Reports (Time of Day, 2003–2015)

(b) Tornado Reports (Season, 2003–2015)
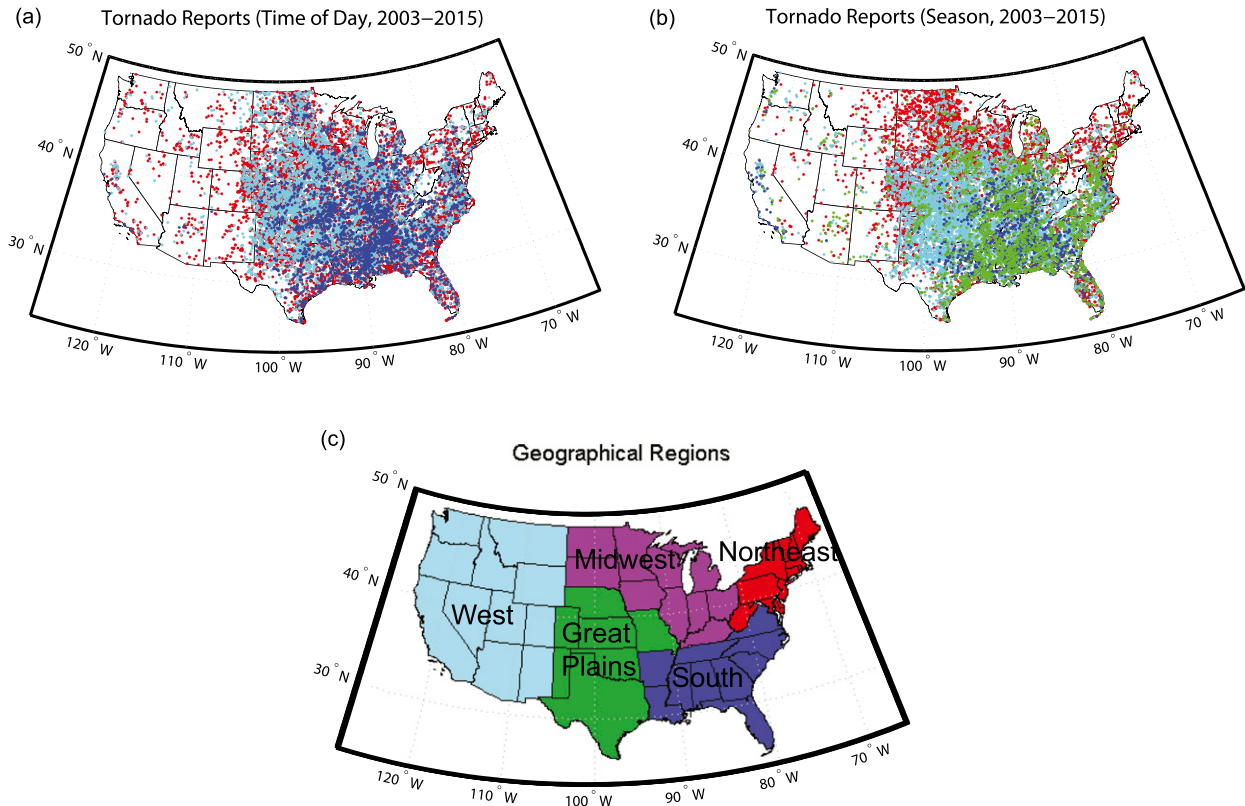
(c)

Geographical Regions

FIG. 2. Geographical plots of 2003–15 tornado events. (a) Tornado events for daytime (red; events occurring between local sunrise and 2 h before local sunset), EET (cyan; events occurring between 2 h before and 2 h after local sunset), and nocturnal (blue; events occurring between 2 h after local sunset and local sunrise) storms. Note the general west-to-east progression from daytime to EET to nocturnal tornadic activity, consistent with the upscale development and progression of storms throughout the day. (b) Tornado events for spring (cyan), summer (red), fall (green), and winter (blue) storms. Note the shift in seasonality, with the Great Plains and much of the Midwest exhibiting the traditional spring storms, the northern states showing a tendency toward summer storms, and southern states showing a great deal of tornadic activity either associated with hurricane season (fall) or even in the winter. (c) Map of geographical regions of the United States used in this study. [Adapted from Anderson-Frey et al. (2016).]

events was obtained by taking archived mesoanalysis gridded data from the grid box closest to the reported tornado (Smith et al. 2012); the soundings were obtained from the Rapid Update Cycle model (RUC; Benjamin et al. 2004) for January 2003–April 2012, or the Rapid Refresh model (RAP; Benjamin et al. 2016) for later dates. Since tornado warnings may encompass multiple model grid boxes, the proximity sounding chosen is that which is nearest to the location of maximum STP contained within the warning. Note that this could lead to slightly different values for a given event that appears in both the event and warning databases. See Thompson et al. (2012) and Anderson-Frey et al. (2016) for discussion of the limitations of the dataset.

Our SOMs receive as input a series of 480 km × 480 km gridded RUC/RAP mesoanalysis datasets centered on either the grid point nearest the tornado (for the tornado event dataset) or the grid point nearest the location of maximum STP contained within the tornado

warning region (for the tornado warning dataset). The algorithm used to create the SOM clusters is loosely summarized as follows (Vesanto et al. 2000):

1) Create a user-specified number M of "nodes", that is, 480 km × 480 km maps of randomly generated parameter values. The number of nodes is also the final number of clusters.

2) Select the first input map randomly from the list of input maps. Compare this map with each of the M nodes via point-by-point analysis of Euclidean distance.

3) "Nudge" each node: nodes that are more similar to the input map (i.e., those with smaller Euclidean distances) are more strongly nudged toward the input map values. Nodes that are less similar to the input map are not as strongly nudged. Refer to Nowotarski and Jensen (2013, their Fig. 3) for a diagram depicting this step.
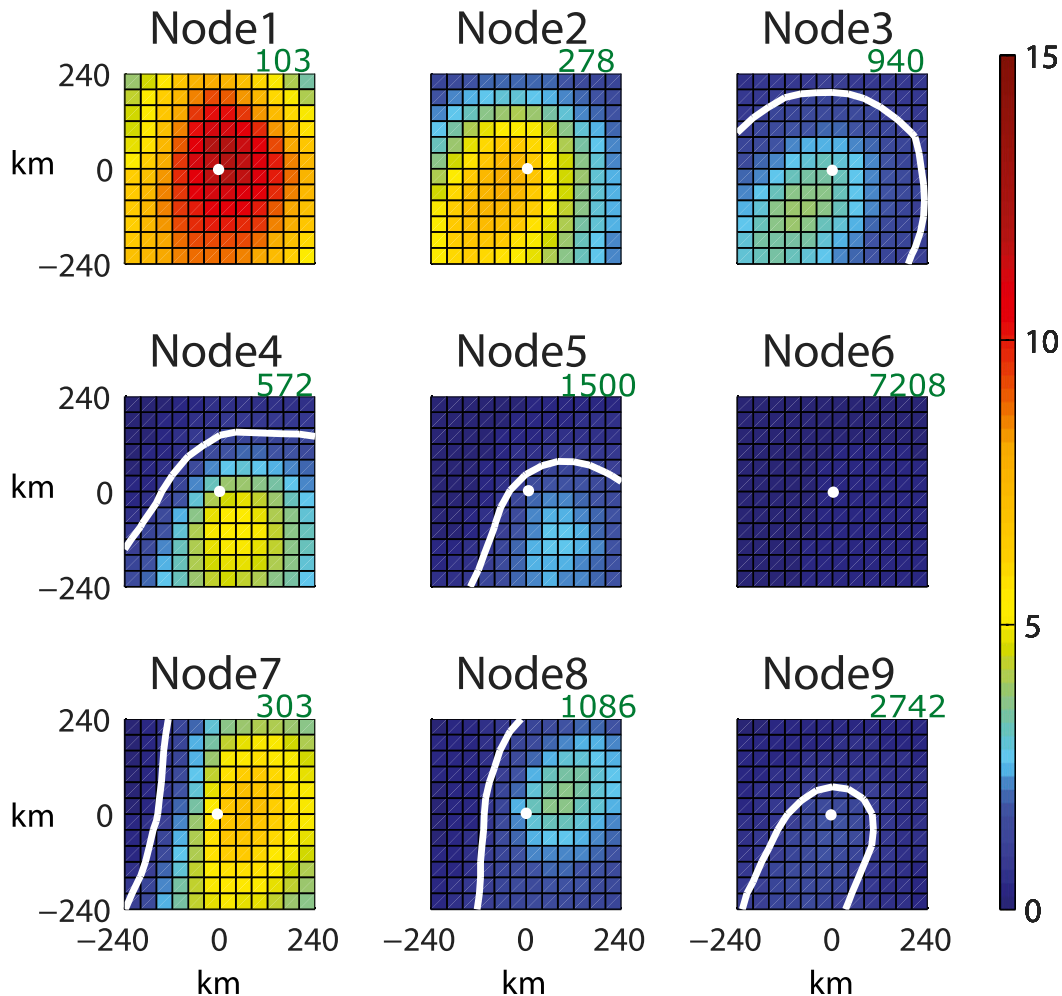
FIG. 3. SOM results for 3 × 3 nodes of 2003–15 tornado event values of STP (unitless) on a 480 km × 480 km grid with the location of the tornado at the center (white dot) of each node. The numbers of events are listed at the top-right corner of each node in green. White contours delineate STP = 1.

4) Select a second input map randomly from the list of input maps. Compare this map with each of the M new nodes point by point, resulting in the same nudging process.

5) Repeat this process across all input maps, and then iterate several times until the nodes stabilize into M statistically distinct maps.

6) Compare each input map with the nodes for the final iteration. Each input map is then assigned to the cluster corresponding to the node it most closely resembles. Each cluster can then be analyzed separately.

The input into the SOM algorithm is thus a large number of 480 km × 480 km maps of any environmental variable (in our case, STP), and the output is a much smaller number M of clusters of 480 km × 480 km maps of that variable, each of which can be summarized as a statistically distinct characteristic node. The only user-specified quantity in the SOM process is the number M of nodes, and this flexibility is both a blessing and a curse. On the one hand, it can be extremely valuable to specify the complexity of the result (i.e., the final number of clusters). On the other hand, the specification process is often fairly arbitrary. A SOM that uses too few nodes means that the maps contained in each cluster will have very similar statistics to the dataset as a whole, since each cluster samples a large portion of the dataset. A SOM that uses too many nodes, on the other hand, will result in redundant nodes with functionally identical appearance (i.e., two nodes will show essentially the same map). Through sensitivity testing, for much of the following work $M = 9$ nodes seemed to strike an optimal balance between capturing important features in the data and reducing redundancy in the final nodes; see Fig. 1 for several alternative examples of numbers of
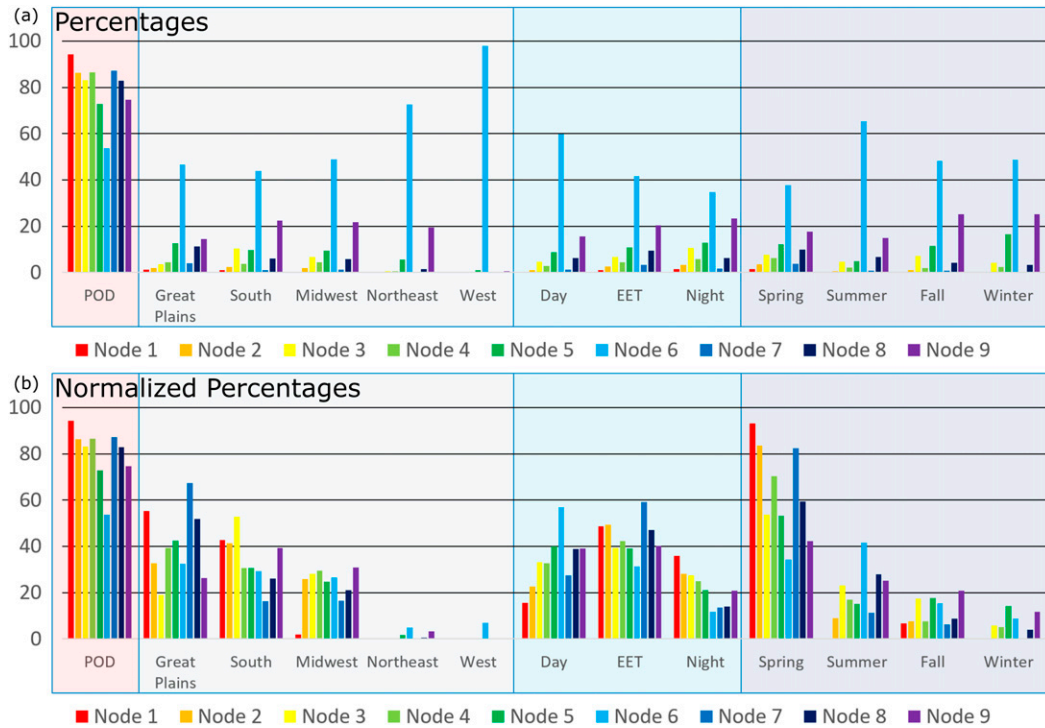
FIG. 4. (a) Bar plot of percentages corresponding to each tornado event STP SOM node (see Fig. 3) for each category (POD, geographical location, time of day, and season). These plots are not normalized by the number of events sorted into each node; for instance, the percentages within the Great Plains category will sum to 100%. (b) Bar plot of normalized percentages, all but POD divided through by the number of events sorted into each node. For instance, the percentages for the node 1 bars within the five geographical categories will sum to 100%.

nodes. The SOM process is iterated 200 times (sensitivity tests did not show appreciable differences with higher numbers of iterations) before grouping each map into its component cluster.

3. Environmental overview

It is useful to first examine some parameter-space diagrams in order to properly characterize the environments in which tornadic activity occurs. Anderson-Frey et al. (2016) found that tornado events and tornado warnings both occur over a wide variety of environments, with the highest density centered on about $1250 \, \mathrm{J \, kg^{-1}}$ of MLCAPE, $25 \, \mathrm{m \, s^{-1}}$ of 0–6-km shear, 800-m MLLCL heights, and $200 \, \mathrm{m^2 \, s^{-2}}$ of 0–1-km storm-relative helicity (their Fig. 2). Generally speaking, forecast skill tends to increase (i.e., POD is higher and FAR is lower) as MLCAPE and SHR6 increase; that is, as we move into environments more traditionally regarded as favorable for supercellular storms (Anderson-Frey et al. 2016, their Fig. 4). For the MLLCL–SRH1 parameter space, POD exhibits a fairly clear threshold at around $100 \, \mathrm{m^2 \, s^{-2}}$ of SRH1, below which the skill is considerably lower. In contrast, the FAR exhibits little variability with either

SRH1 or MLLCL heights (Anderson-Frey et al. 2016, their Fig. 4).

Figure 2 depicts the geographical distribution of the tornado events used in the following analysis. The points are colored according to time of day (Fig. 2a) and time of year (Fig. 2b). Figure 2c shows the five geographical regions to be used in the discussion of the results.

4. Self-organizing maps: Tornado events

Figure 3 shows the output for a SOM created for the tornado events based on patterns of STP values. Strictly speaking, the nine nodes depicted are in fact the mean environments of each of the nine clusters, but in the following discussion the term "node" will be used to discuss each cluster. Note that, for instance, node 1 consists of extremely high values of STP centered on the location of the tornado, whereas node 7 shows a dramatic east–west gradient in STP values and node 6 is characterized by low values of STP.

Figure 4 depicts graphical summaries of relevant categories for each node, including the geographical regions and time of day and year making up the node. In

(a) Node1 SRH1   (b) Node1 MLCAPE

(c) Node2 MLLCL   (d) Node2 MLCAPE
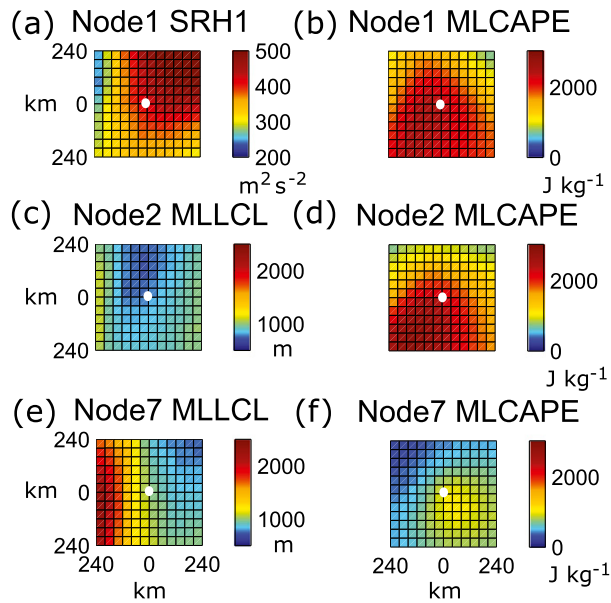

(e) Node7 MLLCL   (f) Node7 MLCAPE

FIG. 5. Tornado event SOM plots of interest. These are the mean values of the given parameter averaged across each node as defined in Fig. 3: (a) node 1 SRH1 ($m^2 s^{-2}$), (b) node 1 MLCAPE ($J kg^{-1}$), (c) node 2 MLLCL (m), (d) node 2 MLCAPE ($J kg^{-1}$), (e) node 7 MLLCL (m), and (f) node 7 MLCAPE ($J kg^{-1}$).

Fig. 4a, for each category, the number of events in each node is divided by the total number of events in that category so that one can easily see the mix of nodes making up the category. In Fig. 4b, for each node, the number of events in a particular category is divided by the total number in that node, such that one can see the categories that are most influential to a particular node. Note that only POD, and not FAR, can be calculated here, given that false alarm warnings are not included in the event database.

Figure 3 shows that the environments clustered into node 1 feature extreme values of STP across a wide region, with the highest values centered on the location of the tornado. When we plot some of the STP-related parameters within this cluster (Fig. 5), we find that these

high STP values are generally due to extreme values of SRH1 (Fig. 5a); in order to compare quantitatively, we define, for instance, $\overline{MLCAPE}_N$ as the mean MLCAPE across node N and $\langle \overline{MLCAPE} \rangle$ as the mean MLCAPE value computed from the nine nodes' $\overline{MLCAPE}$ values (see Table 1 for the list of values). MLCAPE at node 1 (Fig. 5b), for instance, is fairly high compared to other nodes ($\overline{MLCAPE}_1$ = 2237 J kg^{-1} is more than one standard deviation higher than $\langle \overline{MLCAPE} \rangle$ = 1251 J kg^{-1}), but the SRH1 values show a more dramatic difference ($\overline{SRH1}_1$ = 429 m^2 s^{-2} is more than two standard deviations higher than $\langle \overline{SRH1} \rangle$ = 222 m^2 s^{-2}). These environments with particularly high values of STP correspond to the highest values of POD in the dataset (Fig. 4a; 94.2%, as compared with the 66.4% average for the entire dataset); this is what we might expect based on the proximity sounding studies in Anderson-Frey et al. (2016), but this plot confirms that having a widespread geographical region of high STP values seems particularly helpful for issuing accurate warnings. Node 1 environments are overwhelmingly likely to occur in either the Great Plains (55.3%) or the South (42.7%), are more likely to occur during the early evening transition (EET; 48.5%) or the night (35.9%) than during the day (15.5%), and are overwhelmingly springtime events (93.2%) (Fig. 4b).

Nodes 2 and 7 show the value of a two-dimensional approach over a point-based proximity sounding approach. The nodes feature similar values of STP at the location of the tornado (6.5 and 6.6, respectively), but the spatial distribution of STP values surrounding the tornado are quite different (Fig. 3). As Fig. 5 shows, node 7 has a strong east–west gradient in MLLCL heights (Fig. 5e) and MLCAPE (Fig. 5f), which is typical of, for example, a dryline setup; in contrast, node 2 shows smaller gradients in both parameters (Figs. 5c,d), but stronger MLCAPE to the southwest dominates the STP signal. As depicted in Fig. 4b, nodes 2 and 7 have similar values of POD, but node 2 is slightly dominated by the South (41.4%), with many Great Plains (32.7%)

TABLE 1. Mean values of STP component parameters (MLCAPE, SHR6, MLLCL, and SRH1) for the tornado event nodes depicted in Fig. 3, along with probability of detection values for each node.

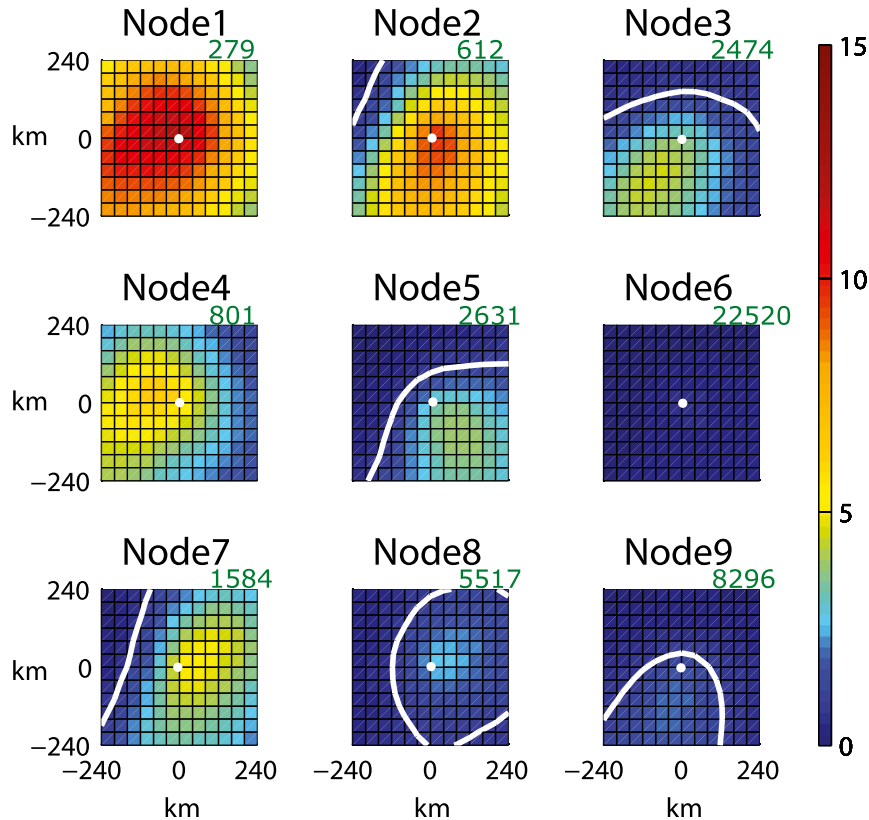|	MLCAPE (J kg^{-1})	SHR6 (m s^{-1})	MLLCL (m)	SRH1 (m^2 s^{-2})	POD (%)
Node 1	2237	31	738	429	94.2
Node 2	2114	30	807	406	86.3
Node 3	1517	28	836	340	83.0
Node 4	1739	30	806	373	86.5
Node 5	1244	27	832	289	72.8
Node 6	956	21	1065	134	53.8
Node 7	2499	30	972	323	87.1
Node 8	2168	27	1058	278	82.8
Node 9	1209	26	846	284	74.7
All events	1251	24	966	222	66.4

FIG. 6. As in Fig. 3, but for the 2003–15 tornado warning dataset rather than the tornado event dataset. Keep in mind that while the STP SOM nodes produced for the tornado warning dataset may resemble those in Fig. 3 corresponding to the tornado events, the datasets are different (14 814 tornado events vs 44 961 tornado warnings) and hence the node 1 for tornado events is not the same node 1 as is shown in tornado warnings. Qualitative comparisons can, however, still be made between similar-looking environments for events and warnings, and the warning node order has been changed to facilitate comparison with event nodes.

and Midwest (25.9%) events as well, whereas node 7 is strongly dominated by the Great Plains (67.3%), with relatively few South (16.2%) or Midwest (16.5%) events. The isolation of node 7's prototypical Great Plains springtime likely dryline setup would not be possible with a single proximity sounding value of STP; with the SOM approach, the distinction is clear. The SOM method described herein can objectively and efficiently differentiate between these two distinct scenarios. As a result, more precise environmental climatologies can be developed for a particular meteorological scenario (e.g., Great Plains springtime dryline setup) that is climatologically more frequent to a geographic region. Furthermore, this method can indirectly incorporate other associated variables (e.g., dryline, surface low placement relative to the warm sector, upper-level flow pattern) common to these setups and can be inferred by the spatial distribution of STP.

Finally, node 6 contains the most marginal values of STP across the board (Fig. 3), together with the lowest

value of POD (53.8%). This extremely marginal node also has a higher percentage of daytime events than EET events and is characterized by more summer (41.7%) than spring (34.4%) events and makes up nearly 50% of the dataset.

5. Self-organizing maps: Tornado warnings

Similar to node 1 in the tornado events database (Fig. 3), node 1 in the tornado warnings database (Fig. 6) consists largely of high values of STP. Using a similar approach to that taken for the analysis of the event SOMs in Fig. 4, Fig. 7b shows that the extreme environments in node 1 feature relatively low FAR, with a value of 67.7% (cf. with the dataset's overall FAR of 76.3%). Note that only FAR, and not POD, can be calculated here, given that events for which no warning was issued are not included in the warning database. Node 1 environments are also dominated by the South (65.2%), with some Great Plains warnings (29.4%) and very few warnings in any
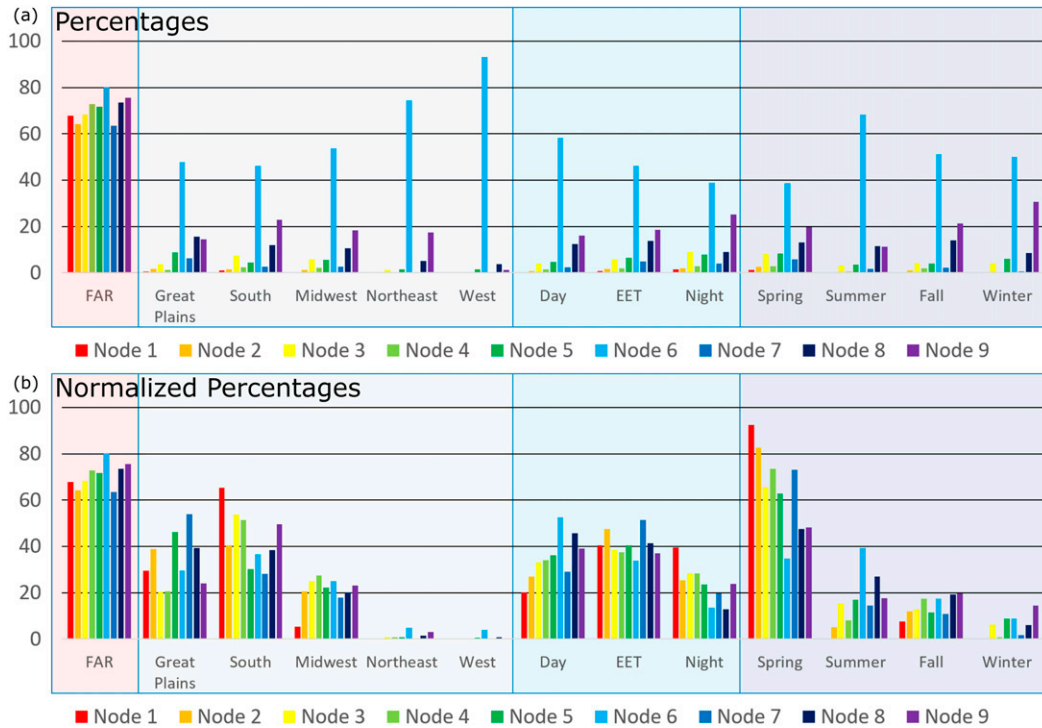
FIG. 7. As in Fig. 4, but for the 2003–15 tornado warning dataset rather than the tornado event dataset. Here, POD has been replaced by FAR.

other region. In terms of time of day, node 1 features an even split between EET and nocturnal events (40.5% and 39.4%, respectively), with many warnings also occurring during the day (20.1%). Much like the high-STP node in the events database (node 1 in Fig. 3), node 1 in the warnings database is heavily dominated by springtime warnings (92.5%). Thus, if a tornado warning is issued within a broad region of extreme values of STP, chances are that the warning in question occurred in the South or the Great Plains, during the evening or night, and during the springtime.

Nodes 3 and 5 (Fig. 6) again show the advantage of considering the fully two-dimensional distribution of STP values rather than relying on a single point parameter value: the two nodes feature similar values of STP (3.3 and 2.9, respectively) at the location of the tornado warning (white dots in Fig. 6), but node 5 has more favorable environments to the southeast, and node 3 has more favorable environments to the southwest. Figure 7 shows that nodes 3 and 5 have similar values of FAR, but node 5 is dominated by Great Plains warnings (46.2%), while node 3 contains a higher proportion of South warnings (53.8%).

Finally, node 6 for the tornado warnings (Fig. 6), like node 6 for the tornado events (Fig. 3), is an environment in which extremely marginal or even zero values of STP

extend over a considerable distance. Unfortunately, the majority of warnings (just over 50%) are occurring in these low-skill environments.

6. Conclusions and summary

For tornado reports and warnings, relatively favorable environments (i.e., high values of STP) at the location of the tornado correspond to relatively high POD and relatively low FAR. The importance of the spatial distribution and heterogeneity of favorable environments, however, becomes clear when we consider examples such as nodes 2 and 7 in the tornado event SOM (Fig. 3): these two environments have similar values of STP at the location of the tornado, but the spatial distribution characterizing node 7 is likely a Great Plains springtime dryline event, whereas node 2 represents a favorable environment that is more evenly split geographically between the Great Plains and the South. Likewise, in the tornado warning SOM (Fig. 6), nodes 3 and 5 have similar values of STP where the warning was issued, but node 5 has more favorable environments to the southeast, while node 3 has more favorable environments to the southwest. Node 5 is dominated by Great Plains warnings, while node 3 is dominated by South warnings. Uniformly low-STP environments dominate the tornado event and

warning databases as a whole; the majority of tornado events and warnings occur in these environments with relatively high FAR and relatively low POD, which is a pattern that persists even when (E)F0 tornadoes are removed from the dataset (not shown). It is therefore reasonable to conclude that tornado warning verification skill metrics will be associated with the dominant warning nodes during a specific time period; consequently, dominant warning nodes may disproportionately influence warning skill metrics for a particular period compared with longer-term climatology.

The two-dimensional plots produced by self-organizing maps are intuitive and immediately highlight similarities and differences between environments at a more precise level than the traditional proximity sounding approaches. The methodology outlined and demonstrated in this work is intended to provide a framework that can be used for any studies that would benefit from enhanced and more nuanced climatologies of the near-storm environments of severe weather events. In upcoming work, we will use the SOM approach to probe the near-storm environments and statistics of tornado outbreaks, as well as the environments and statistics of the warnings issued during these events; we also will explore the use of the SOM approach for different combinations of convective ingredients.

REFERENCES

Anderson-Frey, A., Y. Richardson, A. Dean, R. Thompson, and B. Smith, 2016: Investigation of near-storm environments for tornado events and warnings. *Wea. Forecasting*, **31**, 1771–1790, doi:10.1175/WAF-D-16-0046.1.

Benjamin, S., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, doi:10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.

——, and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, doi:10.1175/MWR-D-15-0242.1.

Brooks, H., J. Lee, and J. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94, doi:10.1016/S0169-8095(03)00045-0.

Christensen, W., Jr., and R. Bryson, 1966: An investigation of the potential of component analysis for weather classification. *Mon. Wea. Rev.*, **94**, 697–709, doi:10.1175/1520-0493(1966)094<0697:AIOTPO>2.3.CO;2.

Kohonen, T., 1982: Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, **43**, 59–69, doi:10.1007/BF00337288.

Liu, Y., and R. Weisberg, 2011: A review of self-organizing map applications in meteorology and oceanography. *Self-Organizing Maps—Applications and Novel Algorithm Design*, J. Mwasiagi, Ed., InTech, 253–272.

Markowski, P., and Y. Richardson, 2014: The influence of environmental low-level shear and cold pools on tornadogenesis: Insights from idealized simulations. *J. Atmos. Sci.*, **71**, 243–275, doi:10.1175/JAS-D-13-0159.1.

——, J. Straka, and E. Rasmussen, 2003: Tornadogenesis resulting from the transport of circulation by a downdraft: Idealized numerical simulations. *J. Atmos. Sci.*, **60**, 795–823, doi:10.1175/1520-0469(2003)060<0795:TRFTTO>2.0.CO;2.

Mercer, A., C. Shafer, C. Doswell III, L. Leslie, and M. Richman, 2009: Objective classification of tornadic and nontornadic severe weather outbreaks. *Mon. Wea. Rev.*, **137**, 4355–4368, doi:10.1175/2009MWR2897.1.

——, ——, ——, ——, and ——, 2012: Synoptic composites of tornadic and nontornadic outbreaks. *Mon. Wea. Rev.*, **140**, 2590–2608, doi:10.1175/MWR-D-12-00029.1.

Nowotarski, C., and A. Jensen, 2013: Classifying proximity soundings with self-organizing maps toward improving supercell and tornado forecasting. *Wea. Forecasting*, **28**, 783–801, doi:10.1175/WAF-D-12-00125.1.

Shafer, C., A. Mercer, L. Leslie, M. Richman, and C. Doswell III, 2010: Evaluation of WRF Model simulations of tornadic and nontornadic outbreaks occurring in the spring and fall. *Mon. Wea. Rev.*, **138**, 4098–4119, doi:10.1175/2010MWR3269.1.

Smith, B., R. Thompson, J. Grams, C. Broyles, and H. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, doi:10.1175/WAF-D-11-00115.1.

Thompson, R., R. Edwards, J. Hart, K. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, doi:10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2.

——, B. Smith, J. Grams, A. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, doi:10.1175/WAF-D-11-00116.1.

Vesanto, J., J. Himberg, E. Alhoniemi, and J. Parkhankangas, 2000: SOM toolbox for MATLAB 5. Helsinki University of Technology Tech. Rep. A57, 8 pp. [Available online at http://www.cis.hut.fi/somtoolbox/package/papers/techrep.pdf.]