# P4.63 Application of a self-organizing map statistical technique to a RUC supercell proximity sounding database

ANDERS A. JENSEN,* CHRISTOPHER NOWOTARSKI

*Department of Meteorology, The Pennsylvania State University, University Park, PA*

## 1. Introduction

A wide breadth of work examining supercell thunderstorms has established links between the characteristics of such storms and their environments (see e.g., Maddox 1976; Thompson et al. 2003). Thermodynamic and wind shear data from proximity soundings have been used to create various forecast parameters for supercells and tornadoes including the supercell composite parameter (SCP) and the significant tornado parameter (STP) (Thompson et al. 2002). Most forecast parameters are limited to assessing wind shear or instability over pre-designated layers, often condensing a two-dimensional profile into a single quantity. Self-organizing maps (SOMs) allow for the classification of vertical profiles of relevant thermodynamic and wind variables, using information over the entire depth of the profile. This study applies the SOM technique to RUC proximity sounding data with the goal of determining which variables are most effective at discriminating between nonsupercell, supercell, and tornadic environments.

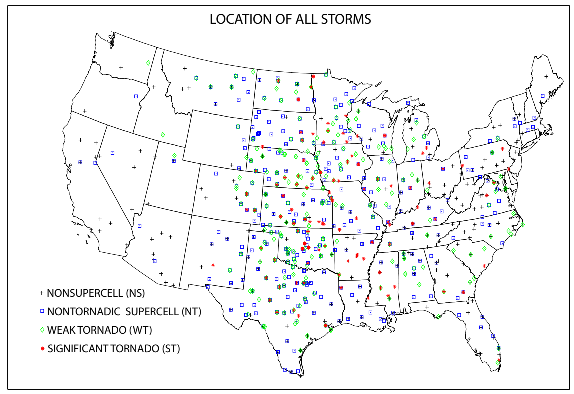


FIG. 1. The location and storm type of all 1185 RUC-2 soundings used for this analysis.

## 2. Data

The self-organizing maps in this study were created using the Rapid Update Cycle-2 (RUC-2) proximity sounding data set collected by Thompson et al. (2003) and augmented with additional data by Thompson et al. (2007). There were 1185 proximity soundings collected in the contiguous United States between April 1999 - June 2001 and January 2003 - March 2005 (fig. 1). Soundings with no surface-based convective available potential energy (CAPE) were removed in an effort to eliminate elevated supercells. Soundings were divided into four categories: Nonsupercell (NS), nontornadic (NT) supercell, where a cyclonic mesocyclone was detected by radar, weakly tornadic (WT) supercell, where an F0 or F1 tornado was reported, and significantly tornadic (ST) supercell, where an F2 tornado or greater was reported (fig. 2). Sounding data were interpolated to 100-m AGL intervals and vertical profiles of various thermodynamic (e.g., relative humidity, potential temperature, stability) and wind (e.g., ground- and storm-relative wind speed, direction, shear, helicity density, vorticity) variables were computed.

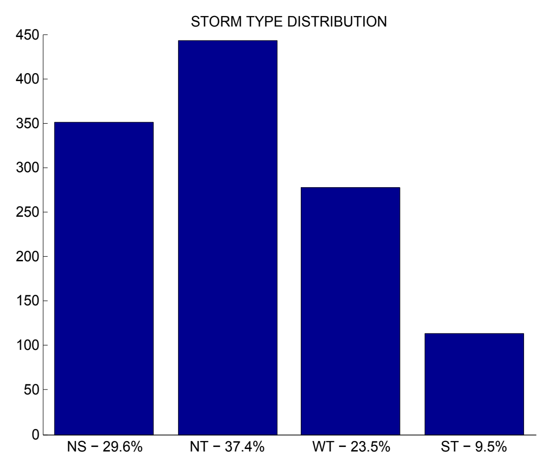


FIG. 2. The number of each storm type in the RUC-2 dataset.

## 3. Self-Organizing Maps

A self-organizing map (SOM) works by using an initialized map of nodes that learn from the input data. The number of nodes is user-defined and each SOM node is represented just like the input data. Due to the SOM learning algorithm (see Kohonen 1995), each node becomes more like the input data

**Corresponding author address*: Anders Jensen, Department of Meteorology, The Pennsylvania State University, 413 Walker Building, University Park, PA 16802; *e-mail*: anders@psu.edu.

to which it is closest. The node that is closest to a given input vector is know as the best-matching unit (BMU). Over many iterations the nodes become representative of the input data because of a neighborhood function that weights the amount of learning depending on how close a given node is to each input vector. SOMs are created using various thermodynamic and wind variables. The SOMs have no information about storm type prior to learning; the only information given to the SOM is height level and a specified variable. Therefore, the SOMs learn objectively and can classify storms without any information on storm type. After the SOM is created, the profiles are binned according to their BMU. By comparing the percentage of storm type in a given node to the overall percentage of that storm type we can assess how well a SOM of a given variable discriminates between storm environments.

## 4. Results

### *4.1* $3 \times 3$ *0-6 km GR wind speed SOM*

Of all the variables, ground-relative (GR) wind speed performed best in differentiating between all storm types. The $3 \times 3$ 0-6 km GR wind speed SOM is shown in figure 3. The topology of this SOM results in increasing wind speed (and shear) from lower left to upper right. This coincides with a transition from the NS to the ST regime. Figure 4 shows the location and storm type of each profile corresponding to each node. Node 7 shows a lot of NS storms which mainly occur in the southern and eastern United States. Node 3 which includes more ST storms due to increased 0-1 km shear, shows these storms to occur more often in the Ohio River Valley. Figure 5 displays the time of day time of year of each profile. This figure confirms that the seasonality of ST storms is mainly Spring and Fall (node 3).

In order to measure how well a SOM node performs against the entire dataset, the percentage difference between each storm type in each node and the entire dataset is computed and plotted in figure 6. Large positive values indicate a given storm type in a certain node is more likely to occur as compared to the entire dataset. The average hodographs for each node are displayed in figure 7. Comparing nodes 2 and 3 from figure 3 show both nodes to have similar bulk 0-6 km wind shear. Both the larger 0-1 km wind shear and curvature that is evident in the average hodograph for node 3 as compared to node 2 is what leads to the increased ST storms in node 3.

### *4.2* $3 \times 3$ *0-1 km streamwise vorticity SOM*

The $3 \times 3$ 0-1 km streamwise vorticity SOM (nodes 7 and 9 shown in fig. 8) performed best in discriminating significant tornado environments, though it was less adept at differentiating other storm types. The profiles matching node 9 are relatively rare (approximately 6% of all profiles), but such profiles are characterized by large streamwise vorticity in the lowest 500 m AGL which leads to the increases percentage of WT and ST storms. The profiles composing node 7 have larger surface-based CAPE (SBCAPE) on average, but the lack of low-level streamwise vorticity and the increased crosswise vorticity as seen by the average hodograph lead to node 7 having a lower percentage of ST storms than node 9.

### *4.3* $3 \times 3$ *0-6 km* $d\theta/dz$ *SOM*

Nodes 5 and 9 from the $3 \times 3$ 0-6 km $d\theta/dz$ SOM are shown in figure 9. Both of these nodes have some degree of low-level stability below 3 km. However, the profiles in node 9 tend to be more stably stratified. Despite a similar, if not better, average hodograph than node 5 and similar location there are very few ST profiles in node 9. Strong surface-based convective inhibition (SBCIN) and the prevalence of overnight storms in node 9 suggest surface-based stability may be impeding tornadoes in this regime. Enhanced 0-1 km shear in node 9 suggests some of these profiles may be in a nocturnal low-level jet scenario.

## 5. Discussion and Conclusions

This study showcases the potential ability of the SOM technique in predicting supercells and tornadoes through objective classification of storm environments based on the shape of vertical profiles of relevant variables. The best performing variables in terms of their ability to discriminate between all storm types were ground-relative wind speed, storm-relative wind speed, streamwise vorticity, ground-relative u and v wind components, and ground-relative helicity density. Though wind variables are generally better discriminators than thermodynamic variables, stability variables show some skill. The worst performing variables were wind direction, relative humidity, and crosswise vorticity. In general, 0-6 km SOMs were better than 0-1 km SOMs in differentiating amongst all storm types, though some 0-1 km SOMs were better than their 0-6 km counterparts in predicting tornadoes. In addition to predicting storm type, the SOMs also show some ability in discerning certain weather patterns, location, and seasonal regimes. We plan on testing the SOM method with other height levels (0-3 km) and different numbers of nodes. We also plan to extend the SOM technique to three dimensions, using hodographs and profiles of dew point and temperature. Eventually SOMs might be a useful forecasting tool wherein real-time data can be compared with SOMs and a conditional probability of storm type may be issued.

*Acknowledgments. We wish to thank Roger Edwards and Rich Thompson for providing the RUC proximity sounding data set. We would also like to thank Harold Brooks, Paul Markowski and Jim Marquis for helpful discussions related to this work. Funding for this work was provided by NSF grant AGS-0644533 and US Department of Energy grant DE-FG02-05ER64058.*

REFERENCES

Maddox, R. A., 1976: An evaluation of tornado proximity wind and stability data. *Mon. Wea. Rev.*, **104**, 133142.

Kohonen, T., 1995: Self-Organizing Maps. Springer, 392 pp.

Thompson, R. L., R. Edwards, and J. A. Hart, 2002: Evaluation and interpretation of the supercell composite and significant tornado parameters at the Storm Prediction Center. Preprints. *21st Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., J11-J14.

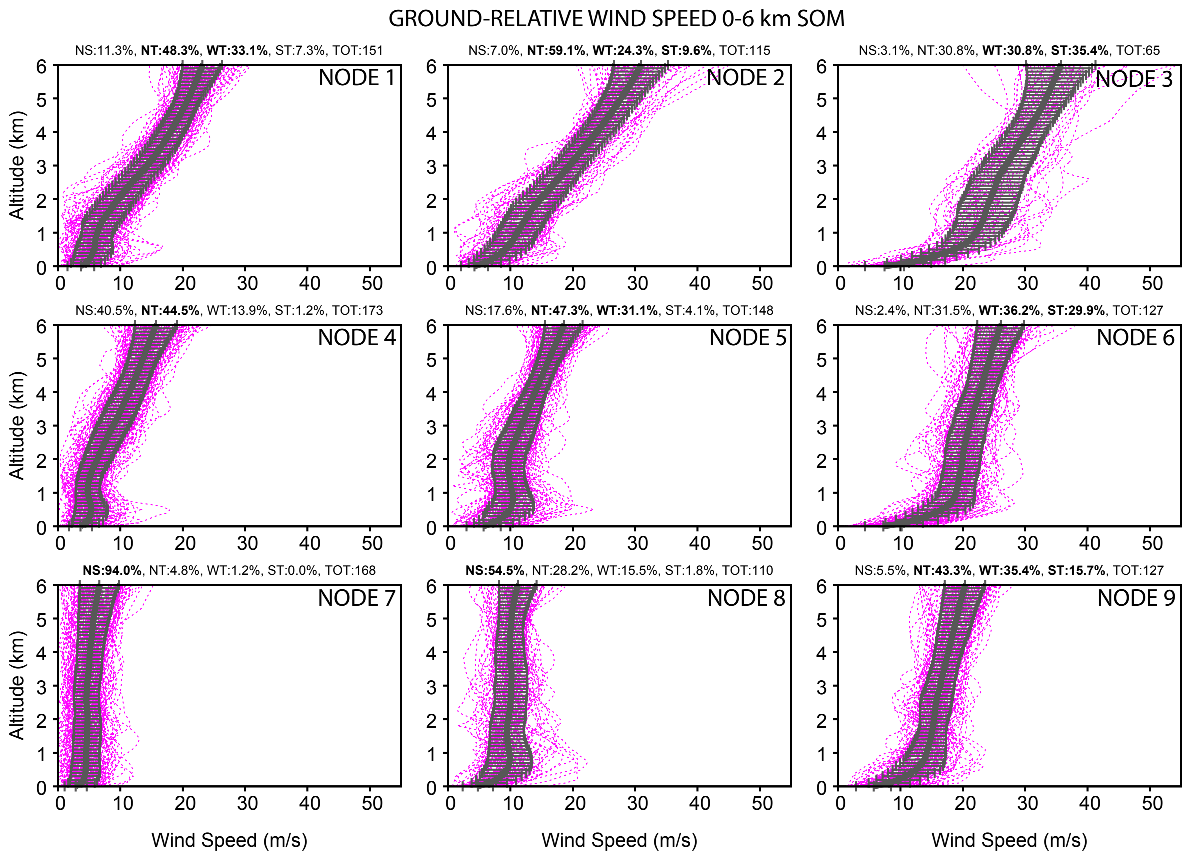


FIG. 3. The $3 \times 3$ 0-6 km ground-relative wind speed SOM where the profiles are binned according to each node. The profiles are plotted in magenta and the average and one standard deviation is plotted in gray.

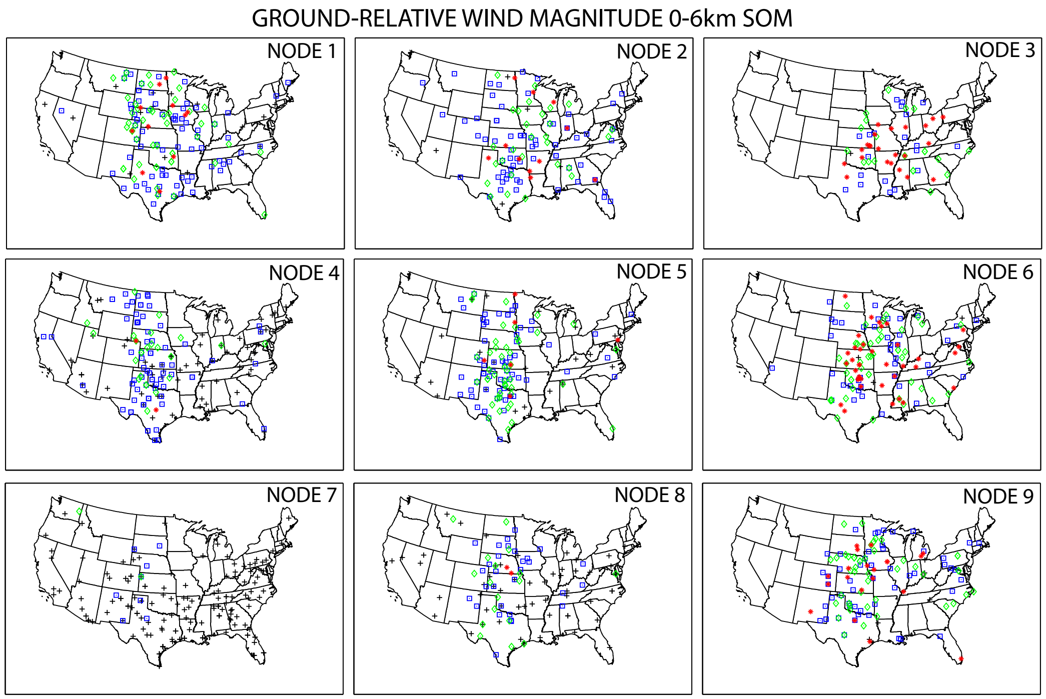


FIG. 4: The location of the profiles corresponding to each node labelled by storm type (see fig. 1 for key).

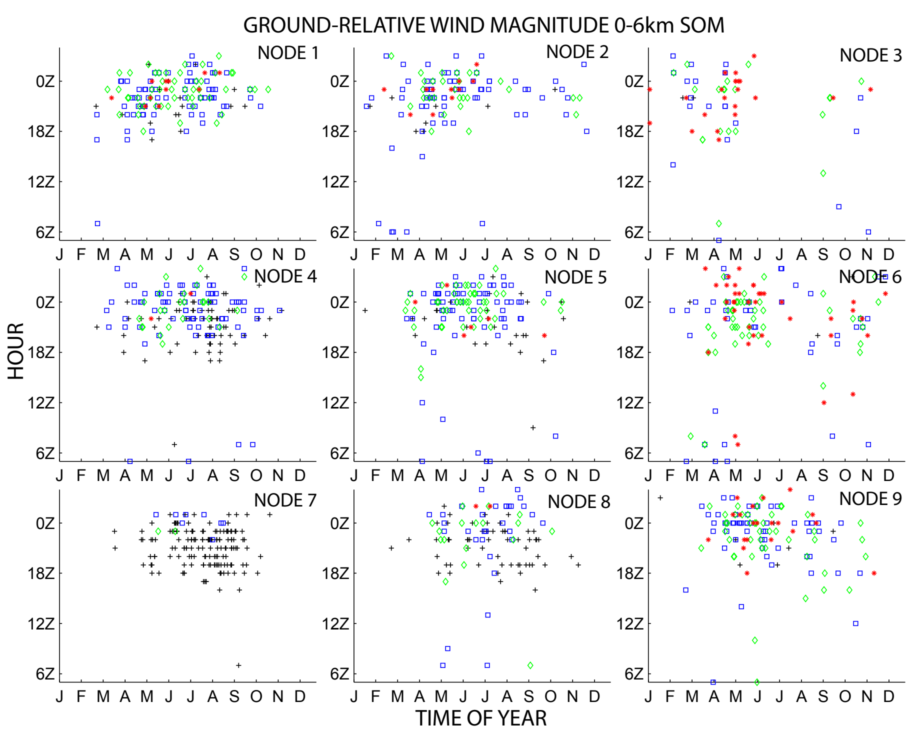


FIG. 5. The time of day and time of year of the profiles corresponding to each node labelled by storm type (see fig. 1 for key).

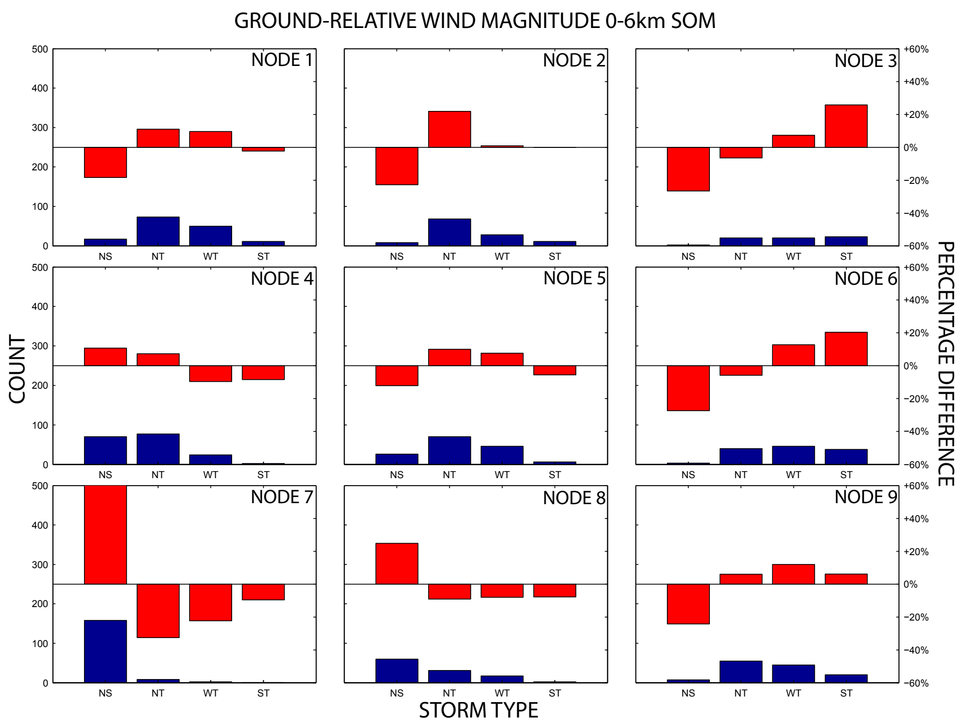


FIG. 6. The number of each storm type binned according to node (blue bars) and the percentage difference of the number of each storm type in each node relative to the total number of each storm type (red bars).

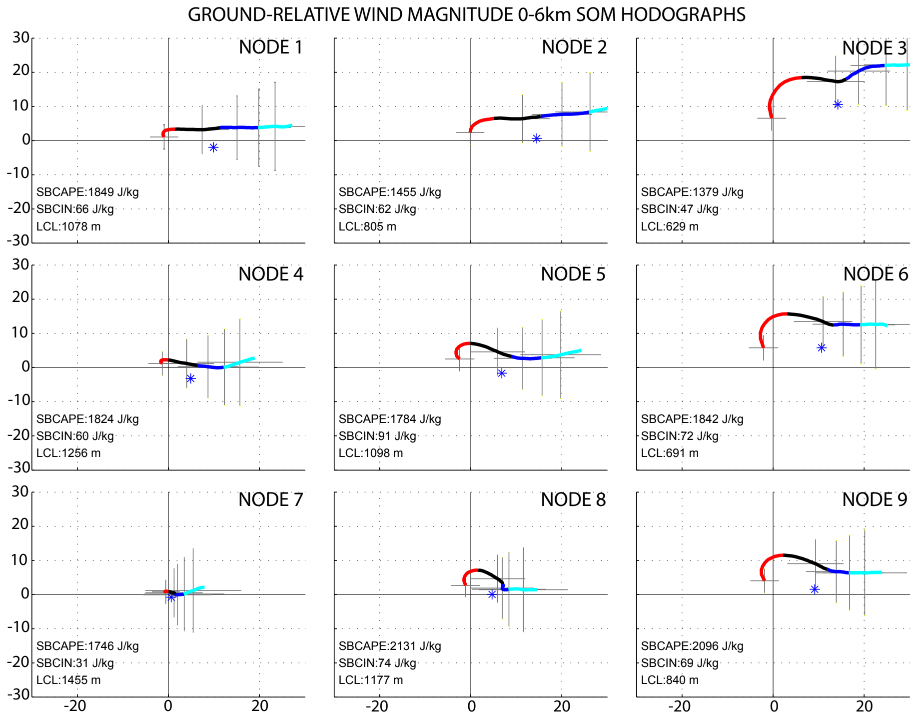


FIG. 7. The average hodographs and one standard deviation of all the profiles binned according to each node. The hodographs are plotted from 0-1 km (red) 1-3 km (black), 3-6 km (blue), and 6-10 km (cyan). Average storm motion (blue asterisk), SBCAPE, SBCIN, and lifting condensation level (LCL) are also shown.

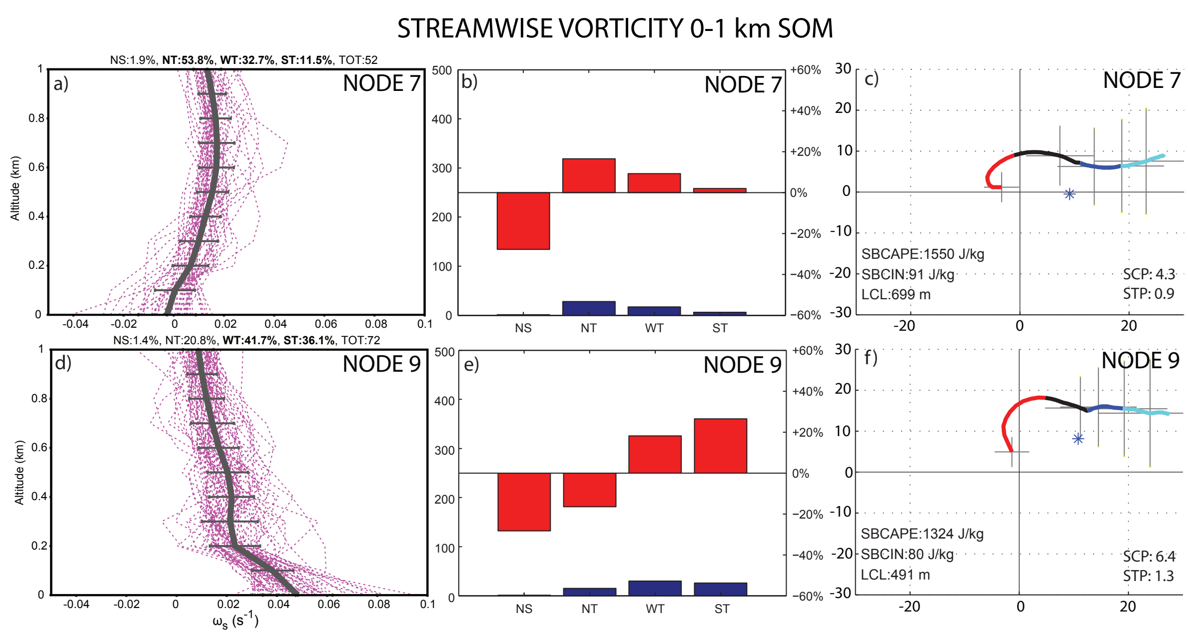


FIG. 8. The profiles (magenta) and their average and standard deviation (gray) binned with nodes 7 and 9 of the 0-1 km streamwise vorticity SOM. Similar to Figures 6 and 7, percentage difference and average hodograph are also plotted for profiles binned in nodes 7 and 9.

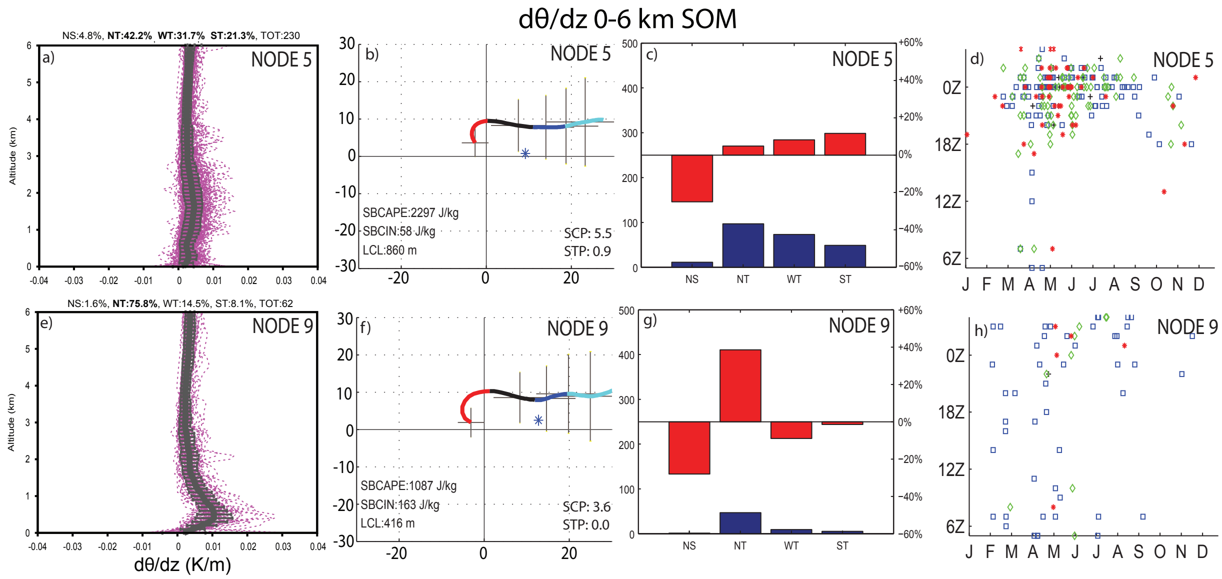


FIG. 9. The profiles (magenta) and their average and standard deviation (gray) binned with nodes 5 and 9 of the 0-6 km $d\theta/dz$ SOM. Similar to Figures 7, 6, and 5, average hodograph, storm type percentage difference, and storm time are also plotted for profiles binned in nodes 5 and 9.

Thompson, R. L., K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243-1261.

Thompson, R. L., C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102-115.