ABSTRACT

CAMPBELL, TREVOR ALDRICH. Self-Organizing Maps for Classifying Variability in Southeastern U.S. High-Shear, Low-CAPE Environments. (Under the direction of Gary Lackmann).

Severe convection occurring in high-shear, low-CAPE (HSLC) environments is a common threat in the southeastern United States cool season. Previous studies of HSLC convection have demonstrated the increased operational challenge these environments present compared to their high-CAPE counterparts, corresponding to a higher false alarm rate and lower probability of detection. They are most frequent during the cool season and during the overnight hours in regions where societal awareness to threats associated with severe convection is relatively low. These environments often experience rapid destabilization in the hours prior to convection, sometimes caused by potential instability release via large-scale forcing for ascent. Self-organizing maps (SOMs) were implemented to objectively classify the environmental patterns in HSLC cool season severe events and determine if variations in the distribution of severe weather occurrence and HSLC composite parameters across these patterns exist. Individual meteorological patterns were found to exhibit slight variations within the HSLC convection subclass, but on average these events were defined similarly to previous literature: a strong surface cyclone accompanied by a deep, negatively-tilted mid-level trough and a northward extending region of instability. The presence of other convective ingredients such as lower-tropospheric wind shear, near-surface theta-E advection, and the release of potential instability varied in intensity more significantly across events. No one variable was able to consistently demonstrate differences in the distribution of severe weather occurrence across patterns; all patterns exhibited a higher frequency of lower-impact severe events. Assessing the meteorological patterns associated with the upper and lower quartiles of severe weather occurrence showed the release of potential instability to be most consistently associated with higher-impact events in comparison to other convective ingredients. On average, a HSLC composite parameter was able to accurately depict HSLC severe convection in high-impact events. However, instances of variability in the spatial location or maximum values of this parameter were found in events in which the parameter did not align with the distribution of severe weather. These differences were most heavily influenced by variables representing the release of potential instability.
© Copyright 2022 by Trevor Aldrich Campbell
All Rights Reserved
Self-Organizing Maps for Classifying Variability in Southeastern U.S. High-Shear, Low-CAPE Environments

by
Trevor Aldrich Campbell

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Marine, Earth, and Atmospheric Science

Raleigh, North Carolina
2022

APPROVED BY:

Matthew Parker

Sandra Yuter

Gary Lackmann
Chair of Advisory Committee
BIOGRAPHY

Trevor was born and raised in Coventry, Rhode Island where his passion for grew with curiosity in snowstorm forecasting. Each winter, trying to understand the reasoning behind whether school would be canceled due to snow developed this curiosity into a desired career path. In 2016, Trevor made the decision to attend Plymouth State University in Plymouth, New Hampshire to pursue a degree in Meteorology. During this time, he worked as a member of a research team studying the effects of weather on power outages in New Hampshire and participated in the PSU chapter of the American Meteorological Society. Trevor was also a member of the varsity cross-country and track and field teams during his undergraduate career. In 2020, he accepted the opportunity to move to Raleigh, North Carolina to pursue a graduate degree under the advisement of Dr. Gary Lackmann at North Carolina State University. Trevor still enjoys running and was active member in the Raleigh running community in his spare time.
ACKNOWLEDGMENTS

Thank you to Dr. Gary Lackmann for always giving valuable feedback, guidance, and mentorship during the completion of this project and to Drs. Parker and Yuter for their suggestions and contributions that helped improve this study. This project was funded by NOAA CSTAR grant NA17NWS4680002. Thank you to Dr. Maria Molina from NCAR for providing the self-organizing map code used in this study and for being willing to answer any and all questions about SOMs. Thank you to Keith Sherburn for providing material to help with the HSLC identification process and the implementation of MOSH and SHERBS3 into this study. Thank you to all the graduate students in the Marine, Earth, and Atmospheric science department who helped bring some normalcy to “these unprecedented times.” Finally, thank you to my friends and family for the continuous support throughout my studies. Whether it was from Rhode Island, New Hampshire, or Raleigh, it was all greatly influential in helping me get to where I am now.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2: Methods</td>
<td>9</td>
</tr>
<tr>
<td>2.1 HSLC Event Classification</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Self Organizing Maps</td>
<td>12</td>
</tr>
<tr>
<td>2.3 SOM Representativeness and Impact Metrics</td>
<td>15</td>
</tr>
<tr>
<td>Chapter 3: Results</td>
<td>24</td>
</tr>
<tr>
<td>3.1 SOM Representativeness</td>
<td>24</td>
</tr>
<tr>
<td>3.2 HSLC Environmental Patterns</td>
<td>25</td>
</tr>
<tr>
<td>3.2.1 Sea-Level Pressure Patterns</td>
<td>25</td>
</tr>
<tr>
<td>3.2.2 Potential Instability Patterns</td>
<td>28</td>
</tr>
<tr>
<td>3.2.3 MOSH Patterns</td>
<td>30</td>
</tr>
<tr>
<td>3.3 LSR Quartile SOMs</td>
<td>33</td>
</tr>
<tr>
<td>Chapter 4: Conclusions</td>
<td>69</td>
</tr>
<tr>
<td>4.1 Summary of Findings</td>
<td>69</td>
</tr>
<tr>
<td>4.2 Future Work</td>
<td>72</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>74</td>
</tr>
<tr>
<td>APPENDICIES</td>
<td>79</td>
</tr>
<tr>
<td>Appendix A SOM Training Variable Output</td>
<td>80</td>
</tr>
<tr>
<td>Appendix B Sea-Level Pressure Trained SOM Composite Variables</td>
<td>93</td>
</tr>
<tr>
<td>Appendix C 700-1000 hPa Theta-E Difference Trained SOM Composite Variables</td>
<td>107</td>
</tr>
<tr>
<td>Appendix D MOSH Trained SOM Composite Variables</td>
<td>121</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 2.1 Atmospheric variables used for SOMs training. Variables that were used as a part of a multivariate SOM are defined as such. .............................................................. 17

Table 2.2 Hyperparameters used in the SOM training process. Settings were kept constant throughout all experiments. ................................................................. 18
LIST OF FIGURES

Figure 1.1  SOM trained on sea level pressure (hPa) using Rapid Refresh analysis data for HSLC severe events from 2008-2021. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Nodes depict composite sea-level pressure fields of events that best match that node, contoured and shaded, as in legend, at 2 hPa intervals. Total number of events assigned to a node titled above each node. ...................................................... 6

Figure 1.2  Sample SOM trained on sea level pressure (hPa) using Rapid Refresh analysis data for HSLC severe events from 2008-2021. Sea-level pressure is contoured at 2 hPa intervals, and node-average CAPE (J/kg) is shaded, as in legend. Total number of events assigned to a node are listed above each node....................... 7

Figure 1.3  Sample SOM trained on sea level pressure (hPa) using Rapid Refresh analysis data for HSLC severe events from 2008-2021. Sea-level pressure is contoured at 2 hPa intervals, and node-average 0-6 km vertical wind shear (kts) is shaded, as in legend. Total number of events assigned to a node are listed above each node. 8

Figure 2.1  Domain used to identify HSLC severe events, bounded by 40°N, 29°N, 75°W, and 94°W. LSRs must have occurred within the domain for an event to be considered........................................................................................................... 19

Figure 2.2  LSR sub-domains used to determine if local environments meet HSLC criteria for the 02/13/2019 0000Z HSLC severe event. Boxes represent a 200 km x 200 km domain centered around each LSR from the event. CAPE is shaded, as in legend, to demonstrate the “low-CAPE” local environments. ......................... 20

Figure 2.3  Sample event-relative 2000 km x 2000 km centroids for the 03/01/2017 1800Z and 01/19/2019 1200Z HSLC severe events. Centroids are centered on the average latitude and longitude of all LSRs within each event. ......................... 21

Figure 2.4  Quantization (red) and topographic (blue) errors for SLP trained SOMs from sizes [2, 2] through [7, 7] for A. σ = x - 0.5 where x is the maximum SOM dimension and B. σ=0.5. ...................................................................................................................... 22

Figure 2.5  Quantization (red) and topographic (blue) errors for the SLP trained SOMs of the upper-quartile of number of LSR events from sizes [2, 2] through [5, 5] for A. σ = x - 0.5 where x is the maximum SOM dimension and B. σ=0.5. ......................... 23

Figure 3.1  Histogram plots depicting the correlation coefficients for all cases within a node with respect to the node composite for A. a SLP trained SOM, B. a 700 - 1000 hPa theta-E difference trained SOM, and C. a MOSH trained SOM. The mean correlation coefficient in a node is indicated in red and the median correlation coefficient in a node is indicated in blue, as in legend. ........................................... 37
Figure 3.2 SOM trained on sea level pressure (hPa) from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Pressure is contoured in dashed black every 2 hPa, node-composite MUCAPE is shaded every 500 J/kg, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.3 Histogram plots depicting the amount of LSRs for all cases within each SLP trained SOM node for A. all LSRs and B. tornado-only LSRs. Outliers are excluded to better represent node-to-node LSR spread. The mean number of LSRs in a node is indicated in red and the median number of LSRs in a node is indicated in blue, as in legend.

Figure 3.4 SOM trained on sea level pressure (hPa) from HSLC null events. Data are event-relative 2000 km x 2000 km grids centered on 34.5 °N and -84.5 °W. Sea-level pressure is contoured in black every 2 hPa and node-composite MUCAPE is shaded, as in legend. Node number and total number of events listed above each node.

Figure 3.5 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is shaded every 6 K, as in legend, node composite sea-level pressure is contoured in black every 2 hPa, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Node number and total number of events listed above each node.

Figure 3.6 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is contoured in black and shaded in K, as in legend, and node composite 500-hPa vertical velocity is contoured in dashed black every 3 -µb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.7 Histogram plots depicting the amount of LSRs for all cases within each 700-1000 hPa theta-E difference trained SOM node for A. all LSRs and B. tornado-only LSRs. Outliers are excluded to better represent node-to-node LSR spread. The mean number of LSRs in a node is indicated in red and the median number of LSRs in a node is indicated in blue, as in legend.
Figure 3.8 SOM trained on 700-1000 hPa theta-E difference from HSLC null events. Data are event-relative 2000 km x 2000 km grids centered on 34.5 °N and -84.5 °W. 700-1000 hPa theta-E difference is contoured in black and shaded, as in legend, and node composite 500-hPa vertical velocity is contoured in dashed black every 2 -µb/s. Node number and total number of events listed above each node. 

Figure 3.9 SOM trained on MOSH from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red every 0.25, node-composite SLP is contoured every 2 hPa, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Node number, total number of events, and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.10 SOM trained on MOSH from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red every 0.25 and node-composite 0-1 km vertical wind shear is shaded, as in legend. Node number, total number of events, and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.11 SOM trained on MOSH from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red 0.25, node composite theta-E difference is shaded, as in legend, and node composite 500-hPa vertical velocity is contoured in dashed black every 3 -µb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.12 Histogram plots depicting the amount of LSRs for all cases within each MOSH trained SOM node for A. all LSRs and B. tornado-only LSRs. Outliers are excluded to better represent node-to-node LSR spread. The mean number of LSRs in a node is indicated in red and the median number of LSRs in a node is indicated in blue, as in legend.

Figure 3.13 SOM trained on MOSH from HSLC null events. Data are event-relative 2000 km x 2000 km grids centered on 34.5 °N and -84.5 °W. MOSH is contoured in red every 0.25, node composite 500-hPa vertical velocity is contoured in dashed black every 3 -µb/s and node-composite 700-1000 hPa theta-E difference is shaded, as in legend. Node number and total number of events listed above each node.
Figure 3.14 SOM trained on sea level pressure (hPa) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Pressure is contoured in dashed black every 2 hPa, MUCAPE is shaded, as in legend, and MOSH is contoured in red every 0.25. Total number of events and the average number of wind and tornado LSRs for each node listed above each node. 

Figure 3.15 SOM trained on sea level pressure (hPa) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Pressure is contoured in dashed black every 2 hPa, MUCAPE is shaded, as in legend, and MOSH is contoured in red every 0.25. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.16 SOM trained on 0-1 km vertical wind shear (kts) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Vertical wind shear is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.17 SOM trained on 0-1 km vertical wind shear (kts) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Vertical wind shear is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.18 SOM trained on near-surface theta-E advection (K/day) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E advection is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.19 SOM trained on near-surface theta-E advection (K/day) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E advection is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.20 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is contoured and shaded, as in legend, and node-averaged 500-hPa vertical velocity is contoured in black every 3 μb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.21 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is contoured and shaded, as in legend, and node-averaged 500-hPa vertical velocity is contoured in black every 3 μb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.22 SOM trained on MOSH from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red 0.25, node composite theta-E difference is contoured in dashed black and shaded every 6 K, node composite 500-hPa vertical velocity is contoured in dashed black every 3 μb/s, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

Figure 3.23 SOM trained on MOSH from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red 0.25, node composite theta-E difference is contoured in dashed black and shaded every 6 K, node composite 500-hPa vertical velocity is contoured in dashed black every 3 μb/s, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
CHAPTER 1

1. Introduction

There are many environmental patterns associated with severe convection across the United States, most of which produce societal threats, including damaging wind, hail, and tornadoes. These patterns are often characterized by the ingredients required for convection, such as instability (often quantified using convective available potential energy, CAPE) and vertical wind shear. Two of the main environmental “subclasses” of severe convection in the United States contain strong deep layer wind shear (0-6 km vertical wind shear ≥ 18 m/s). However, one subclass consists of large mixed-layer CAPE (MLCAPE) (MLCAPE ≥ 1000 J/kg) and predominantly occurs in the central United States, while the other consists of small MLCAPE (MLCAPE ≤ 1000 J/kg) and primarily occurs in the southeastern United States (Schneider et al. 2006). Henceforth, these subclasses are referred to as high-shear, high-CAPE (HSHC) and high-shear, low-CAPE (HSLC) respectively. While the probabilities of thunderstorms and tornadoes are high in HSHC environments, there is a higher frequency of total environmental hours characterized by HSLC and thus a large percentage of total thunderstorm and tornado events occur in HSLC environments despite lower probabilities of occurrence (Schneider and Dean 2008). This results in a relatively high false-alarm ratio for severe watches and warnings that pose a challenge to operational forecasters, especially considering the relatively low probabilities of detection of severe weather threats in HSLC environments compared to HSHC environments (Schneider and Dean 2008; Dean and Schneider 2008; Dean and Schneider 2012). The spatial distribution of HSLC severe events amplifies their impact as they typically occur in the southeastern United States, where the population density is greater relative to that in regions where HSHC convection occurs (Dean and Schneider 2012). Impact amplification is also due to the temporal distribution of HSLC severe events: They are most frequent during the cool season (Guyer et al. 2006; Schneider et al. 2006; Sherburn and Parker 2014; Sherburn et al. 2016), during the overnight hours (Kis and Straka 2010; Sherburn and Parker 2014; Sherburn et al. 2016). Further, HSLC convection occurs in regions where societal awareness to threats associated with severe convection is relatively low (Ashley 2007; Ashley et al. 2008). These factors likely contribute to the higher percentage of tornadoes that are deadly in the southeastern United States (3.8%) in comparison to the entire United States (2.0%; Anderson-Frey et al.
Given the operational difficulty and significant impact associated with HSLC severe events, a greater understanding of the variability in environmental patterns and their link to event impact is needed.

Within the HSLC subclass, multiple combinations of other environmental variables exist, ranging from low lifted-condensation levels (LCLs) and large low-level moisture to well-mixed boundary layers that are deep and dry (Sherburn and Parker 2014). However, low LCLs and moist boundary layers within the warm sector or along the cold front of a strong surface cyclone are the dominant pattern of HSLC severe events in the eastern United States, while in the western United States dry boundary layers near a surface triple-point or in an upslope regime are the typical pattern (Sherburn and Parker 2014). HSLC severe events in the southeastern United States are also characterized by strong synoptic-scale forcing for ascent at all levels and a region of enhanced low-level instability extending towards the event center (Sherburn et al. 2016). Case studies of observed tornado-producing storms in HSLC environments revealed that mid-level dry intrusions seem to coincide with onset of severe convection, suggesting the release of potential instability (Lane and Moore 2006; Clark 2009). Case study simulations reinforced the importance of potential instability as a factor in the rapid environmental destabilization in the hours prior to HSLC severe convection in some events, along with surface moisture advection, surface warm-air advection, and strong 0-1 km wind shear; destabilization can occur in time periods ≤ 3 hours prior to convection onset time (King et al. 2017). Additionally, there is difficulty in confirming tornadic vortices using radar-derived azimuthal shear in HSLC environments at distances ≥60 km from radar sites and high false alarm rates associated with azimuthal shear and radar reflectivity signatures close to the radar (Davis and Parker 2014). These difficulties of detecting radar signatures in HSLC environments along with the short time scales of rapid destabilization contribute to increased operational challenges, despite knowledge of the general environments and ingredients conducive to HSLC severe convection.

Understanding the convective dynamics of HSLC environments is also critical to properly assess and forecast HSLC severe events. While most previous literature in the storm-scale dynamics of supercell thunderstorms is focused on high-CAPE supercells, simulations by McCaul and Weisman (2001) found that robust convection can occur in low-CAPE environments when there is a concentration of instability at low levels. This low-level concentration of instability produced stronger and more intense convection than other
configurations of vertical CAPE distribution. The concentration of low-level instability (represented by large low-level lapse rates) produces strong lower-tropospheric updrafts caused by vertical perturbation pressure gradient accelerations (Sherburn and Parker 2019), which is an essential vertical forcing for environments with strong shear and marginal instability. It should be noted that over 23% of all southeastern HSLC tornadoes occur within the quasi-linear convective system (QLCS) convective mode (Anderson-Frey et al. 2019). QLCSs are narrow lines or arcs of convective storms with contiguous precipitation and often form along cold fronts and in environments containing strong vertical wind shear (Markowski and Richardson 2010). Differences in convective mode and dynamics are evident between southeastern HSLC convection and central United States HSHC convection. HSLC convection is generally characterized by strong synoptic-scale forcing and low-level destabilization that creates rapidly evolving convective features (sometimes embedded within QLCSs), whereas HSHC convection is generally characterized by weaker synoptic forcing and ample instability that is more commonly associated with cellular convective features (Smith et al. 2012).

To try and bridge the gap between HSLC climatologies, the dynamical processes associated with HSLC convection, and operational forecasts, several studies have examined the skill of environmental variables in discriminating between HSLC severe and HSLC non-severe events. Sherburn and Parker (2014) identified the 0-3 km and 500-700 hPa lapse rates as the most skillful discriminants. Combining these variables with fixed-level wind and fixed-layer shear variables (SHERB and SHERBE; Sherburn and Parker 2014) or also including variables representing the release of potential instability (MOSH and MOSHE; Sherburn et al. 2016) provided environmental metrics that show promising skill in locating HSLC threat areas. These environmental parameters, along with the increase of high-resolution convection-allowing models (CAMs) in recent years, have provided a baseline for improving HSLC severe convection forecasts. Analysis of 0-3 km updraft helicity (UH) forecasts, used as a proxy for identifying HSLC severe convection, was examined for the High-Resolution Ensemble Forecast (HREF) system, an ensemble of convection-allowing models (Graham 2021). This study found that lower UH thresholds were more skillful in predicting HSLC severe events than higher thresholds, however the false alarm rate in lower thresholds is exceptionally high. While Graham 2021 focuses on UH forecasts, which is consistent with previous severe convection verification literature (Sobash and Kain 2017; Roberts et al. 2020), many of the aforementioned HSLC
environment and ingredient studies suggest that variables beyond UH are essential to identifying HSLC severe convection. Additionally, the temporal extent of a high-resolution verification study in HSLC environments is limited by the relatively recent deployment of operational convection-allowing model ensembles, as the second version of the HREF was implemented in 2017 as NCEP’s first operational CAM ensemble (Roberts et al. 2019). Without a reforecast, there is not a sufficient period of record to conduct a meaningful analysis of HSLC forecast skill for the HREF system.

While previous studies have assessed the patterns and ingredients of HSLC severe convection and verified forecasts of UH in these environments, little work has been done to objectively classify the different synoptic and meso-beta scale patterns that accompany HSLC severe convection. This prompts the question: Are there distinct HSLC patterns that are associated with significant variation in storm impacts? To the author’s knowledge, the only relevant objective classification involving HSLC environments was performed by Anderson-Frey et al. 2019, who used self-organizing maps (SOMs) to classify tornadic near-storm environments in tornado outbreaks, in which the dataset included southeastern HSLC events. They found that the dominant geographic distribution of local significant tornado parameter (STP) maxima was located more westward for southeastern United States outbreaks and that within a single outbreak, the variance of STP patterns between different convective modes can be used to determine typical and atypical STP patterns to a given region. SOMs are an unsupervised machine learning technique that trains a neural network using competitive learning and maps a user-specified number of data nodes across a dataset (Kohonen 2001). SOMs have been used in meteorological studies primarily for classification of synoptic and climate patterns (Hewitson and Crane 2002; Schuenemann and Cassano 2009; Mechem et al. 2018) in addition to the previously mentioned study. The purpose and unique contribution of this study is to objectively classify the patterns that comprise southeastern U.S. HSLC severe convection using SOMs and determine the variations, if any, in severe weather distribution and HSLC composite parameters (SHERBS3 and MOSH) across these patterns.

To illustrate our methods, Figures 1.1-1.3 show output from a SOM trained on event-relative sea level pressure (SLP) patterns drawn from a 13 year HSLC severe event dataset. The SLP node composites (Figure 1.1) exhibit a variety of patterns: A strong surface cyclone with a north-to-south pressure trough (Node 1) and a weak surface cyclone and northeast-to-southwest
pressure trough (Node 9). While the difference in these patterns are prominent, it should be noted that these node composites can be subject smoothing due to spatial variability also, and composite patterns such as the weak surface cyclone could be due to the smoothing of dissimilar cases. Measures to minimize sources of error via inter-node variability of cases are discussed further in Section 2. The SOM nodes in each figure are organized by SLP patterns, but additional meteorological variables can be examined for each SOM configuration, including CAPE (Figure 1.2) and vertical wind shear (Figure 1.3), even when those variables are not used to train the SOM. The differences in CAPE and shear patterns are not as distinct as SLP, but there are subtleties to each. The amount of instability present and the extent to which it extends northward (Figure 1.2) and the relative magnitude and orientation of the 0-6 km wind shear distribution (Figure 1.3) differs based on the associated sea level pressure pattern. This motivates further investigation into the types of environmental patterns that comprise the HSLC subclass and the relative storm impacts associated with each, which would provide valuable context for operational forecasters who are relied on to provide impact-based decision support services. The analyses presented in this study are designed to address the following questions:

1. What are the objectively defined environmental patterns that comprise southeastern United States HSLC severe convection?
2. Are there significant variations in the distribution of severe weather occurrence in association with these environmental patterns?
3. Are there particular environmental patterns that are more prevalent in higher-impact events, where event impact is defined by the occurrence of Local Storm Reports (LSRs)?
4. Are the environmental patterns identified in SOMs consistent with previous environmental HSLC composite parameters such as the SHERBS3 and MOSH?

Section 2 will discuss the methods used for this study, including the data and processes used in creating the HSLC severe event dataset, the choices made in implementing SOMs, and defining analysis metrics for SOM output. The results of this study will be presented in Section 3, emphasizing the variations in HSLC environmental patterns demonstrated by SOM output and the distribution of severe weather occurrence across these patterns. The key findings of Section 3
will be summarized in Section 4, as well as discussing future work and improvements that could be made to this study.

Figure 1.1 SOM trained on sea level pressure (hPa) using Rapid Refresh analysis data for HSLC severe events from 2008-2021. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Nodes depict composite sea-level pressure fields of events that best match that node, contoured and shaded, as in legend, at 2 hPa intervals. Total number of events assigned to a node titled above each node.
Figure 1.2 Sample SOM trained on sea level pressure (hPa) using Rapid Refresh analysis data for HSLC severe events from 2008-2021. Sea-level pressure is contoured at 2 hPa intervals, and node-average CAPE (J/kg) is shaded, as in legend. Total number of events assigned to a node are listed above each node.
Figure 1.3 Sample SOM trained on sea level pressure (hPa) using Rapid Refresh analysis data for HSLC severe events from 2008-2021. Sea-level pressure is contoured at 2 hPa intervals, and node-average 0-6 km vertical wind shear (kts) is shaded, as in legend. Total number of events assigned to a node are listed above each node.
CHAPTER 2

2. Methods

2.1 HSLC Event Classification

A 13-season dataset of HSLC severe events in the Southeastern U.S. was created to provide a sufficient sample size for analysis. Individual HSLC events were defined as six-hour windows occurring during southeastern United States cool-seasons that met specific criteria. These windows are defined as a six-hour period beginning at either 00 UTC, 06 UTC, 12 UTC, or 18 UTC. This interval for event classification was selected to replicate the process of Graham 2021, in which six-hour intervals were determined to resemble the frequency of operational forecast products. The desire for operational relevance influences this study: using six-hour intervals allows for storm development and progression into new forecast domains and keeping a constant interval removes the subjective nature of self-defined event start and end times. Consecutive six-hour windows would not traditionally be considered as independent meteorological events, but assessing these windows separately differentiates the environmental patterns at different stages of storm development where the associated storm impacts could vary. Assessing the variations in severe weather occurrence with respect to environmental patterns is one of the primary goals of this study, thus even consecutive windows are treated as separate events. As previously stated, HSLC environments occur primarily in the southeastern United States (Schneider et al. 2006), so a domain bounded by 40°N, 29°N, 75°W, and 94°W was selected as our analysis region (Figure 2.1), consistent with Graham 2021. In this study, cool-seasons were defined as October 1st through April 30th and the entire dataset spanned the cool seasons of 2008-2009 through 2020-2021, deviating from Graham 2021 to include the entire months of October and April. The process of identifying HSLC severe events was automated using a two-step Python script that checks the Storm Prediction Center’s (SPC) Local Storm Reports (LSRs) archive (NOAA/SPC, 1999) and the National Center for Environmental Prediction (NCEP) 20-km Rapid Refresh/Rapid Update Cycle (RAP/RUC) analysis data (NOAA/NCEP, 2005) provided by the National Centers for Environmental Information (NCEI).
Identifying the presence of severe weather was the first step in determining if a HSLC severe event occurred. Each six-hour window through all cool seasons was examined to identify if SPC LSRs were reported within the southeastern US domain. The limitations of LSRs have been documented as being prone to population density biases and inconsistent reporting practices (Brooks et al. 2003; Doswell et al. 2005; Trapp et al. 2006) and recommendations for the use of radar-derived rotation track data over LSRs for verification purposes have been proposed (Dawson et al. 2017). Despite these disadvantages, LSRs were used in lieu of methods involving radar signatures because of the limitations of distinguishing between tornadic and non-tornadic signatures at distances greater than 60 km from a radar site and the lack of density in the current radar network for HSLC convection (Davis and Parker 2014). Hail reports were excluded from this analysis, because significant hail (> ¾ in.) was not expected to be the primary hazard in these environments and is more commonly associated with high-instability environments (Schneider and Dean 2008). Therefore, wind reports and tornado reports were the only LSR type used in the identification process. If one or more tornado report or wind report occurred within a six-hour window it was defined as a severe event.

To verify if a given severe event met HSLC criteria, the environments around individual LSRs were examined to verify the thresholds of most unstable CAPE (MUCAPE) ≤ 1000 J/kg and 0-6 km wind shear ≥ 18 m/s used in previous studies (Sherburn and Parker 2014; Sherburn et al. 2016; King et al. 2017; Graham 2021) were met. These variables were collected and calculated using 20-km RAP/RUC hourly analysis data as provided by NCEI. The RAP model was commissioned in May 2012, so subsequent events use RAP analysis data while those preceding use RUC analysis data. The focus of the study is on the broader environmental patterns rather than finer meteorological features in individual cases; the RUC/RAP resolution (20-km grid length) and hourly output motivate its use, in addition to serving as the boundary conditions for one of the more widely-used convection-allowing models (High Resolution Rapid Refresh, or HRRR model). The analysis data were taken from the event onset time to measure the environment surrounding the severe reports prior to the onset of severe convection, as would be seen by a forecaster. All LSRs within each event were adjusted from a latitude and longitude coordinate to the closest RUC/RAP analysis data grid cell by calculating the nearest neighbor. Then, a 200 km x 200 km sub-grid around the adjusted LSR was created to analyze the local
environment. A sample of these sub-grids with respect to the CAPE environment of a single event is shown in Figure 2.2.

Next, the events were separated into three categories: events with all LSRs occurring in a HSLC environment (sub-grid averaged MUCAPE ≤ 1000 J/kg and 0-6 km wind shear ≥ 18 m/s), events with no LSRs occurring in a HSLC environment, and events with LSRs in both HSLC and non-HSLC environments. The two former categories were simply defined as averaged MUCAPE ≤ 1000 J/kg and 0-6 km wind shear ≥ 18 m/s HSLC severe events and non-HSLC severe events respectively, but the latter category required additional definition. To accommodate events where strong gradients of CAPE or shear were present or where significant geographic spread in LSRs resulted in notably different local environments, all LSR CAPE and shear sub-grids for a given event were averaged. If the HSLC criteria were met on this averaged grid (MUCAPE ≤ 1000 J/kg and 0-6 km wind shear ≥ 18 m/s), the event was defined as a HSLC severe event. Finally, for events that still did not meet HSLC severe event criteria, if the ratio of LSR sub-grids in HSLC environments to LSR sub-grids not in HSLC environments was ≥ 5, the event was defined as a HSLC severe event. This final criterion was added to accommodate events where averaging all LSR grids would not be representative: for example, most LSRs occur in HSLC environments, but few occur in environments categorized by significant CAPE or minimal shear. These categorizations were implemented with the intention of transferring human judgment of classifying less-distinct HSLC severe events to an automated process. All remaining events were defined as non-HSLC severe events and excluded from future analysis. Out of 11,027 possible six-hour windows over 13 cool seasons, 587 HSLC severe events were defined as having at least one tornado or severe wind report in the southeastern US domain and meeting the HSLC environment criteria.

For comparison between HSLC severe events and HSLC non-severe events, a HSLC “null” event dataset was created. The methods used in identifying null events deviated from the processes used in Sherburn and Parker (2014) and Sherburn et al. (2016). Null events in previous studies were defined as either warning nulls, where a severe thunderstorm or tornado warning was issued in a HSLC environment where no LSRs were reported in that convective day (Sherburn and Parker 2014; Sherburn et al. 2016), or radar-based nulls, where the WSR-88D’s storm cell identification and tracking algorithm and 45 dBZ minimum reflectivity threshold were used to identify non-severe convection (Sherburn et al. 2016). In this study only environmental
thresholds were required, and no convective prerequisites were implemented. Since the local LSR environments cannot be examined in a non-severe event, null events were defined as six-hour windows where no LSRs were reported and a 5,000 km\(^2\) minimum area was observed meeting the HSCL environment requirements (MUCAPE \(\leq\) 1000 J/kg and 0-6 km wind shear \(\geq\) 18 m/s). To identify environments defined by no instability, a minimum MUCAPE threshold of 10 J/kg was also required. These HSCL requirements were assessed for a 1,000 km x 1,000 km region centered on the latitude and longitude center of the HSCL event identification domain (Figure 2.1) to avoid the excessive influence of instability in coastal regions. This resulted in 780 HSCL null events out of 11,027 possible six-hour windows over 13 cool seasons.

### 2.2 Self-Organizing Maps

To objectively classify the environmental patterns associated with the developed HSCL severe event dataset, self-organizing maps (SOMs) were implemented. SOMs are an unsupervised machine learning technique that trains a neural network using competitive learning and maps a user-specified number of data nodes across a dataset (for this study, these data nodes are environmental patterns; Kohonen 2001). A simplified breakdown of the SOM training process (Vesanto et al. 2000) begins with selecting a user-defined number of nodes and the input data, followed by initializing the node weights randomly. Prior to training, all data are normalized using a z-score (Eq. 2.1). In (2.1) \(z\) is the two-dimensional field of a variable for a given event, \(\bar{z}\) is the mean of the variable field across all events, and \(\sigma_z\) is the standard deviation of the variable field across all events.

\[
Z = \frac{z - \bar{z}}{\sigma_z} \tag{2.1}
\]

Once defined, the SOM training starts by selecting a data point and computing the winning node using the node weights. All nodes are then updated using a Euclidean distance calculation and this process is repeated for all data points over a user-defined number of iterations. As this process is repeated, the nodes take the shape of the dataset with similar cases appearing in the same node. SOMs have previously been used in the classification of synoptic and climate patterns, including northeastern U.S. sea-level pressure patterns (Hewitson and Crane 2002), Greenland precipitation and storm track patterns (Schuenemann and Cassano 2009), northeastern
Atlantic 500-hPa geopotential height patterns and associated cloud variability (Mechem et al. 2018), and the impacts of North Atlantic jet regimes by the North Atlantic warming hole (Gervais et al. 2020). In this study, version 2.2.9 of the minisom Python module (Vettigli 2018) was used to create SOMs trained on a set of atmospheric variables (Table 2.1) to define the patterns that comprise southeastern US HSLC severe convection.

These atmospheric variables were selected with previous literature in mind, where strong synoptic forcing was shown to be a key ingredient in producing HSLC severe convection (Sherburn et al. 2016), along with surface warming and moistening, low-level forcing for ascent and the release of potential instability (King et al. 2017). In addition to typical meteorological diagnostics, two multivariate environmental predictors for HSLC severe convection introduced in prior studies are analyzed: SHERBS3 (Sherburn and Parker 2014) and MOSH (Sherburn et al. 2016). The analysis of skill in an ensemble of convection-allowing models performed by Graham (2021) is confined to the period in which the model was in operation, as the ensemble used was implemented in 2017; changes in HREF configuration further limit the sample size available for predictability analysis. However, we still wish to assess predictive tools in HSLC environments in keeping with our original project goals. Therefore, we incorporate environmental predictors along with the basic variables associated patterns, and SHERBS3 and MOSH are evaluated for the various SOM patterns. The SHERBS3 (Eq. 2.2) is a combined parameter designed to separate significant HSLC severe convection from non-severe HSLC storms using a threshold of 1 (significant severe weather is more likely for SHERBS3 > 1). In (2.2), S3MG is the 0-3 km shear magnitude, LLLR is the 0-3 km lapse rate, and LR75 is the 700-500 hPa lapse rate (Sherburn and Parker 2014). The MOSH (Eq. 2.3) is a combined parameter designed to improve upon the SHERB by accounting for low-level forcing that approximates the release of potential instability, with LLLR defined the same as in SHERBS3, S15MG as the 0-1.5 km shear magnitude, and MAXTEVV as the maximum dθe/dz * ω product calculated from the 0-2 km layer through the 0-6 km layer at 0.5 km intervals (Sherburn et al. 2016). It should be noted that these composite variables are calculated at fixed-layer intervals. Sherburn and Parker (2014) and Sherburn et al. (2016) also provide composite variables calculated using effective bulk shear. However, this study chose to prioritize the fixed layer MOSH and SHERBS3 as it used by forecasters due to the inability to calculate effective bulk shear in forecaster products at the time of this study (author correspondence to NOAA personnel - Keith Sherburn).
Two methods of formatting the SOM input data were used for different environmental classification processes, both using the previously mentioned RAP/RUC analysis data at the event onset time. The first method was creating an event-relative 2000 km x 2000 km centroid centered on the average LSR latitude and longitude coordinate for each event and converting that coordinate to the closest grid point. This was designed to exhibit LSR frequency and distribution in relation to meteorological features. Examples of the location and extent of these event-relative centroids are shown in Figure 2.3. Despite the grid’s relatively large size and location in the southeastern United States, there were no instances in which the event-relative centroid encountered the edge of RAP/RUC data domain. The second method was using the southeastern U.S. event identification domain as a fixed domain, designed to relate the LSR frequency and distribution to geographic features. For both domains, all SOMs were trained on atmospheric variables outlined in Table 2.1. For the purpose of this study, only the event-relative data trained SOMs are presented in Section 3. Fixed-domain SOMs produced output that was organized by the geographic location of environmental patterns, while event-relative SOMs produced output organized by location of environmental patterns relative to severe weather occurrence. Assessing the distribution of severe weather occurrence across the environmental patterns is a primary goal of this study, thus, fixed-domain SOMs were not included for further analysis.

Determining the desired number of nodes for SOM analysis required balancing the excessive smoothing of distinct environmental patterns from using too few nodes and the overpopulation of similar environmental patterns within the SOM from using too many nodes. To optimize the number of nodes, a combination of quantitative analysis of error metrics and qualitative assessment of performed sensitivity tests was used. The error metrics used in this study are quantization error and topographic error. Quantization error defines “how close each
synoptic setting is to its best-matching node” (Mechem et al. 2018), and decreases with increasing number of nodes, while topographic error defines “the ratio of the number of synoptic settings whose second best-matching node is nonadjacent to its best-matching node” (Mechem et al. 2018), and increases with increasing nodes. An optimal SOM node configuration will minimize both errors to the greatest extent possible. Calculating these error metrics for all variables showed a level of error variability that is highly dependent on the “smoothness” of the training variable. For example, a 4 x 4 dimension SOM trained on a smoother variable such as 500-hPa geopotential height trained SOM has a lower quantization error (32) than the nosier near-surface temperature advection trained SOM (93). Additionally for this study, the SOM hyperparameter σ was defined as the largest SOM dimension – 0.5 (ex. 3.5 for a 4 x 4 dimension SOM). This results in a lower topographic error and a smaller range of quantization errors, albeit slightly larger values, across SOMs of all sizes than if sigma were held constant at 0.5 (Figure 2.4). Given reduced clarity in the exact SOM dimensions that produce the lowest combined error by assessing topographic error against quantization error, the sensitivity of varying SOM dimensions was tested by examining the variability of events within individual nodes for SOMs of different training variables (not shown). The 4 x 4 dimension SOM was most consistent in balancing the overpopulation of similar environmental patterns with the underrepresentation of distinct environmental patterns and was thus selected for use in analysis. The remaining hyperparameters used in the SOM training process were kept constant throughout all experiments presented in this study and are presented in Table 2.2.

2.3 SOM Representativeness and Impact Metrics

While the quantization error defines how representative SOM nodes are, the values are only useful in understanding the relative error between SOMs trained on different variables. To better define SOM representativeness, a correlation coefficient was calculated between each event within a SOM node and the node composite for all variables. The data for both single events and node composites were archived in the SOM output as arrays of data for each grid point. Thus, a correlation coefficient value between these two arrays was calculated and used to assess the similarity between an event and its node composite. Similar to number of LSRs, the mean and median of correlation coefficients are calculated to understand which nodes are the
most representative in an individual SOM and reinforce quantization error by defining which training variables provide more representative SOM nodes on average.

In addition to qualitatively analyzing HSLC environmental patterns, impact metrics were calculated to determine the variation in severe weather frequency between SOM nodes. Given that LSRs were used in this study to characterize severe weather, the impact metrics were defined by the spread of total number of LSRs across all events within a node. Comparing the variability in mean and median values of total LSRs across nodes was designed to assess if there are meteorological patterns that are more conducive to producing more severe weather. This was calculated for both total LSRs and tornado-only LSRs to assess whether particular patterns are more or less conducive to tornadic activity. Comparing the spread of tornado-only LSRs to total LSRs is a means of overcoming the limitations of LSRs by verifying that the cause of severe weather is due to the HSLC environment and not as a product of, for example, a synoptically-forced severe wind environment with strong background wind speeds. If the distribution of total LSR frequency generally matches the distribution of tornado-only LSR frequency, there is further justification that LSRs are a product of an environment in which severe wind reports are heavily favored. However, the frequency of tornado LSRs is much less than wind LSRs, so significant variations in the distribution of average tornado reports across nodes can be less prevalent than for total LSRs.

This spread of severe weather occurrence across nodes is also visualized as a LSR density field. LSR density is calculated at each grid point as the fraction of events where an LSR is reported to total events within a node, where a nearest neighbor calculation is used for the conversion of an LSR coordinate to assigned grid point. Given that the data are event relative, the maximum values of LSR density were generally in the center of the domain. However, assessing the spatial distribution of LSR density compared to the orientation of a given meteorological variable provided reinforcement of the variable’s impact.

To create a further distinction between the environmental patterns associated with varying degrees of impact, as defined by number of LSRs, a final set of SOMs was trained on the upper- and lower-quartile LSR HSLC severe events. The upper quartile was defined as events with ≥36 LSRs and the lower quartile was defined as events with ≤4 LSRs. Given the reduced number of events in the respective datasets, additional error metric sensitivity testing, as performed in section 2.2, was implemented to arrive at a desired SOM dimension of 3 x 3
(Figure 2.5). The set of training variables used is consistent with previous SOMs (Table 2.1) and the only change in SOM hyperparameters (Table 2.2) was in input length to reflect the number of events in each quartile: 148 for the upper quartile and 160 for the lower quartile respectively.

**Table 2.1** Atmospheric variables used for SOMs training. Variables that were used as a part of a multivariate SOM are defined as such.

<table>
<thead>
<tr>
<th>Training Variables</th>
<th>Used in Multivariate SOM?</th>
</tr>
</thead>
<tbody>
<tr>
<td>500-hPa Geopotential Height</td>
<td>Yes (with Sea-Level Pressure)</td>
</tr>
<tr>
<td>Sea-Level Pressure</td>
<td>Yes (with 500-hPa Geopotential Height)</td>
</tr>
<tr>
<td>700-1000 hPa Theta-E Difference</td>
<td>Yes (with 500-hPa Vertical Velocity)</td>
</tr>
<tr>
<td>MUCAPE</td>
<td>No</td>
</tr>
<tr>
<td>0-6 km Vertical Wind Shear</td>
<td>No</td>
</tr>
<tr>
<td>0-1 km Vertical Wind Shear</td>
<td>No</td>
</tr>
<tr>
<td>0-1 km Storm Relative Helicity</td>
<td>No</td>
</tr>
<tr>
<td>Near-Surface Temperature Advection</td>
<td>No</td>
</tr>
<tr>
<td>Near-Surface Dewpoint Temperature Advection</td>
<td>No</td>
</tr>
<tr>
<td>Near-Surface Theta-E Advection</td>
<td>No</td>
</tr>
<tr>
<td>500-hPa Vertical Velocity</td>
<td>Yes (with 700-1000 hPa Theta-E difference)</td>
</tr>
<tr>
<td>SHERB</td>
<td>No</td>
</tr>
<tr>
<td>MOSH</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 2.2 Hyperparameters used in the SOM training process. Settings were kept constant throughout all experiments.

<table>
<thead>
<tr>
<th>Hyperparameter Name</th>
<th>Hyperparameter Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Length</td>
<td>587</td>
</tr>
<tr>
<td>Learning Rate</td>
<td>0.5</td>
</tr>
<tr>
<td>Decay Function</td>
<td>Asymptotic Decay</td>
</tr>
<tr>
<td>Neighborhood Function</td>
<td>Gaussian</td>
</tr>
<tr>
<td>Topology</td>
<td>Rectangular</td>
</tr>
<tr>
<td>Activation Distance</td>
<td>Euclidean</td>
</tr>
<tr>
<td>Random Seed</td>
<td>1</td>
</tr>
<tr>
<td>Number of Iterations</td>
<td>100000</td>
</tr>
<tr>
<td>Random Order</td>
<td>True</td>
</tr>
<tr>
<td>Verbose</td>
<td>True</td>
</tr>
</tbody>
</table>
Figure 2.1 Domain used to identify HSCL severe events, bounded by 40°N, 29°N, 75°W, and 94°W. LSRs must have occurred within the domain for an event to be considered.
Figure 2.2 LSR sub-domains used to determine if local environments meet HSLC criteria for the 02/13/2019 0000Z HSLC severe event. Boxes represent a 200 km x 200 km domain centered around each LSR from the event. CAPE is shaded, as in legend, to demonstrate the “low-CAPE” local environments.
Figure 2.3 Sample event-relative 2000 km x 2000 km centroids for the 03/01/2017 1800Z and 01/19/2019 1200Z HSLC severe events. Centroids are centered on the average latitude and longitude of all LSRs within each event.
Figure 2.4 Quantization (red) and topographic (blue) errors for SLP trained SOMs from sizes [2, 2] through [7, 7] for A. $\sigma = x - 0.5$ where $x$ is the maximum SOM dimension and B. $\sigma=0.5$. 
Figure 2.5 Quantization (red) and topographic (blue) errors for the SLP trained SOMs of the upper-quartile of number of LSR events from sizes [2, 2] through [5, 5] for A. $\sigma = x - 0.5$ where $x$ is the maximum SOM dimension and B. $\sigma=0.5$. 


CHAPTER 3

3. Results

Although all SOMs trained on the variables outlined in Table 2.1 provide distinction in the patterns of HSLC environments, the goals of this study are to determine if the objectively defined environmental patterns of HSLC severe convection are consistent with previous literature and to assess if there are variables associated with significant variations in the distribution of severe weather. Identification of a complete set of HSLC synoptic patterns could serve to increase forecaster situational awareness, in the event that less common, undocumented patterns exist. To best address these goals while keeping the discussion concise, SOMs trained on sea-level pressure (SLP), 700-1000 hPa theta-E difference (TE710), and MOSH are selected for further discussion in this section. These SOMs allow for a comprehensive analysis by encapsulating environmental pattern variability within the scope of synoptic patterns, convective ingredients, and predictive parameters. SOMs trained on the remaining variables and composites of other parameters for the aforementioned SOMs are presented in Appendices A-D.

3.1 SOM Representativeness

It is important to remain aware when examining SOM output is that the node output plots are composites of the cases within the node, and case-to-case variability exists within the composite sample. A well-behaved SOM will have nodes that consist of cases that closely resemble the node’s composite field. This is measured by the quantization error of the SOMs, which generally decreases as the SOM matrix dimension increases. As discussed in section 2, the level of variability within SOM nodes is highly dependent on the smoothness or degree of spatial variability of the training variable. SLP is a slightly smoother field than TE710, and both are smoother and continuous as compared to noisier, discontinuous MOSH fields. Values of quantization error, which is the average Euclidean distance between each input field and its composite SOM node, are 55 for the SLP trained SOM, 60 for the TE710 trained SOM, and 78 for the MOSH trained SOM respectively. To assess the meaning of these values, a correlation coefficient was calculated for all cases within a node with respect to the node composite for each
SOM (Figure 3.1a-c). The results were node mean correlation coefficients between 0.65 – 0.85 for the SLP SOM (Figure 3.1a), 0.75 - 0.85 for the TE710 SOM (Figure 3.1b), and 0.2 - 0.6 for MOSH SOM (Figure 3.1c). The variability in correlation coefficients can be explained by examining the individual cases within a SOM node. For example, node 1 of the SLP trained SOM (Figure 3.2) is defined by the node composite as a weak surface cyclone with northeast-to-southwest pressure trough. The individual events within this node are characterized by both surface cyclones in the western region of the domain with eastward extending pressure troughs (and associated implied warm fronts), and the southern portion of a southward extending pressure trough (with associated implied cold fronts), where the surface cyclone is located close to the northern extent of the domain (not shown). Despite differences in surface cyclone location and pressure trough orientation, the values of sea-level pressure in the center of the domain were similar. While a SOM wasn’t excluded in this study for a larger quantization error or lower correlation coefficients as compared to other variables, it is a valuable reminder that the results of composting contain these variations, and that even in nodes made up of similar cases, as defined by the SOM, there is enough variability to distort some details of the orientation and gradients of the respective meteorological patterns associated with individual cases.

3.2 HSLS Environmental Patterns

3.2.1 Sea-Level Pressure Patterns

Training a HSLS severe event SOM on SLP defines the variability in synoptic environments by distinguishing the sea-level pressure and frontal patterns associated with HSLS convection (Figure 3.2). The SLP trained SOM is organized by surface cyclone location and pressure trough structure (orientation) on the x-axis and central surface cyclone pressure (intensity) on the y-axis. The left SOM matrix column is defined by surface cyclones located farther eastward with northeast-to-southwest oriented pressure troughs, while the right column is characterized by surface cyclones located further west with north-to-south oriented pressure troughs. For a given column, weaker surface cyclones are presented in the top row, while stronger surface cyclones are located in the bottom row. This SOM implies that HSLS severe convection can occur at any given stage of a surface cyclone life cycle (given HSLS conditions):
the patterns in the top row depict a weaker surface cyclone and pressure trough, and each row beneath shows a progression into a stronger surface cyclone with a strong pressure trough and associated implied frontal system. This progression occurs for surface cyclones of increasing intensity from the left column to the right column. All nodes are defined by some amount of northward extending instability, as defined by MUCAPE. The amount of instability and the zonal location of northward extent varies with SLP patterns: instability is greatest in stronger surface cyclones and extends northward in regions corresponding to the associated pressure trough. Comparison of SLP patterns in relation to all other composite variables (Table 2.2) is presented in Appendix B.

Assessing the distribution of severe weather across these patterns is accomplished by examining the plotted LSR density in relation to meteorological patterns and the spread in the number of LSRs per case within each node. In the SLP SOM (Figure 3.2), the centered LSR frequency maximum is oriented parallel to a trough of lower pressure and an implied frontal boundary in all nodes regardless of the differences in trough orientation, although some have a clearer alignment than others (node 5 vs node 8). This finding suggests that thermally direct frontal circulations may frequently play a role in these severe convection events; enhanced horizontal potential temperature gradients associated with fronts are also consistent with strong vertical wind shear via the thermal wind relation. Nodes with stronger surface cyclones and north-to-south oriented pressure troughs generally feature more LSRs per case (a measure of “impact”) than those with weaker surface cyclones and northeast-to-southwest oriented pressure troughs, where impact is defined by the node average number of LSRs.

This general trend is supported by Figure 3.3a, where most nodes with a greater number of average LSRs also have greater median number of LSRs per event. However, the node depicted by the strongest composite surface cyclone (node 16) has fewer average LSRs (37) than some of the adjacent nodes (nodes 11 and 12, with 53 and 46 average LSRs respectively), and the median number of LSRs for node 16 also depicts this trend. Nodes 1 and 16 can be defined as the two most distinct in terms of synoptic pattern, and the difference between average LSRs (17) is nearly double the standard deviation of average LSRs across all nodes (10). If the same comparison was made with node 11 (the highest impact node, which depicts a similar synoptic pattern) instead of node 16, the difference in average LSRs would be upwards of 3x the standard deviation of average LSRs across all nodes. Both of which suggest a meaningful separation of
severe weather occurrence. It is clear from Figure 3.3a that all nodes are skewed towards events defined by fewer total LSRs, however, the difference in median number of LSRs across nodes generally follows the difference in mean number of LSRs across nodes. Nodes with higher mean values of tornado LSR frequency (Figure 3.3b) are located along the right column of the SOM matrix (nodes 4, 8, 12), but the differences between high and low tornado LSR frequency is small. Additionally, the highest impact tornado LSR nodes do not align symmetrically to the highest impact total LSR nodes and all nodes have small median number of tornado LSRs. Despite this, comparing Figure 3.3a-b shows a trend (albeit not consistent throughout) in the type of synoptic patterns associated with higher-impact nodes for both total LSRs and tornado-only LSRs: they are generally defined by stronger surface cyclones and north-to-south oriented pressure troughs (and implied associated fronts, baroclinicity, and shear).

As discussed in section 2.1, we also identified a set of “null” events that were characterized by high shear and non-zero CAPE in the same location. Comparing Fig. 3.2 to the null SOM case, also trained on SLP (Figure 3.4), allows us to contrast the surface synoptic patterns of severe vs non-severe HSLC events. The same concept of organization by trough orientation and cyclone intensity is present here: this SOM is organized by surface cyclone location and trough orientation on the x-axis and cyclone intensity on the y-axis. The nodes in the right column of this SOM matrix more closely resemble the pressure patterns depicted in the severe event SOM, while those in the left column are characterized by anticyclonic flow, and an eastward shifted surface cyclone with a pressure trough extending westward from the low-pressure center. Despite the sea-level pressure patterns on the right side of the null event SOM resembling some of the patterns defined in the severe event SOM, the instability environments (as defined by MUCAPE) of the null event nodes are less enhanced. A tongue of northward extending instability is more prominent in the severe event SOM, while the greatest values of northward extending instability in the null event SOM are located near the eastern edge of the domain over ocean waters. In retrospect, requiring a higher minimum CAPE threshold over land areas may have produced a more convectively favorable set of null events.

A similar comparison between severe and null HSLC events was performed by Sherburn et al. (2016, their Figs. 9j-l). They found that a composite of HSLC severe events was defined by a surface cyclone just north of the domain center, with a tongue of enhanced 0-3 km surface-based CAPE (SBCAPE) extending northward, while composites of HSLC null events (defined as
warning or radar-based nulls) featured weaker composite surface cyclones and a less prominent region of northward extending SBCAPE in the radar-based null composite. On average, the severe event SOM and null event SOM demonstrate similar results to Sherburn et al. (2016). However, the novelty of the SOM output lies in the definition of individual nodes that show a variety of synoptic patterns: the severe event SOM shows varying surface cyclone intensity and pressure trough orientations that are capable of producing HSLE severe convection. The null event SOM shows similar surface cyclone patterns in some nodes, but on average, weaker surface cyclones, more limited instability, and patterns that are less conducive to upward vertical motion.

3.2.2 Potential Instability Patterns

In addition to traditional severe convective ingredients, the release of potential instability (PI) was determined to be an important factor in the rapid destabilization of environments in some HSLE severe events (Sherburn et al. 2016; King et al. 2017). This motivated the generation of a SOM trained on 700-1000 hPa theta-E difference (TE710) to distinguish the differences in PI environments across HSLE severe events (Figure 3.5). The PI trained SOM is organized by the orientation of the TE710 gradient on the x-axis and the strength of the TE710 gradient on the y-axis. The top row exhibits weak, zonal TE710 gradients, while the bottom row is defined by stronger, more meridional TE710 gradients that resemble a frontal boundary. Additionally, the amount of negative TE710 is limited in left column of the SOM matrix while the right column is characterized by northward extent of negative TE710 values, representing more ample PI environments. This SOM shows that some amount of PI is present in HSLE severe events, but its strength and distribution vary considerably across the SOM matrix.

The composite SLP environments for each TE710-trained node show minimal difference in the intensity of the surface cyclone, but orientation of a pressure trough across nodes matches TE710. The nodes with well-defined TE710 gradients are well represented by a strong pressure trough, while the nodes with a broader region of PI are represented by a pressure trough behind the northward extending region of PI. The regions of PI in all nodes are encompassed by the warm sector of the surface cyclone. The orientation of the TE710 gradient aligns with the distribution of LSR density as well. Nodes with broader regions of PI have a slightly less defined
area of LSR density located in the tongue of northward extending PI, but nodes with sharp
gradients of TE710 exhibit LSR density slightly more narrow regions of LSR density directly
along the gradient. When overlaying 500-hPa vertical velocity to represent the release of PI
(Figures 3.6), there is generally stronger rising motion when a strong TE710 gradient is present,
while the areas of broader PI have weaker rising motion. TE710 patterns are compared to the
remaining composite variables (Table 2.2) in Appendix C.

For nodes in the TE710 trained SOM with a clearly defined TE710 gradient, the LSR
frequency is located on and oriented with respect to the gradient in a region of maximum 500-
hPa upward vertical velocity (Figures 3.5 and 3.6). Nodes with weaker gradients have the highest
LSR frequency where there is a northward extent of PI and are also aligned with the region of
maximum 500-hPa upward vertical velocity. Nodes with the highest average number of LSRs are
generally confined to the right columns of the SOM matrix where greater and more northward
extending amounts of PI are present. However, nodes 8 and 12 serve as an exception as shown
by the mean number of LSRs (Figure 3.7a). There is greater impact, in terms of LSRs, in most of
the higher PI nodes in regions of strong 500-hPa upward vertical velocity (ex. nodes 3, 7, and
11).

Similar to the SLP SOM, while a general trend may be present between meteorological
patterns and the number of LSRs, individual node comparisons can produce varying results.
Comparing node 1 and node 16 as two of the most unique patterns shows a difference in average
LSRs (24) that is nearly double the standard deviation of average LSRs across all nodes (12.5)
and suggestive of a meaningful separation of severe weather occurrence. However, the adjacent
node 12 depicts a similar pattern (albeit with a weaker TE710 gradient) and has nearly the same
number of average LSRs (20) and node 1 (21), contradicting the previous suggestion. The
highest impact nodes are generally not confined to a single edge of the SOM where a single
pattern type is more distinct. Rather, the nodes defined by broad areas of PI with weaker TE710
gradients (nodes 3, 7, and 11) are the highest-impact, but do not contain the most PI or the
weakest TE710 gradients of all nodes. A similar pattern is present with the tornado-only LSR
density (Figure 3.7b) where the majority of high-impact tornado nodes are confined to the top
SOM rows, which are defined by broader areas of PI and weak TE710 gradients of varying
magnitude. The nodes with stronger gradients and relatively few tornado LSRs suggest those
environments are more likely to be severe wind events rather than severe tornado events.
Ultimately, a particular TE710 pattern is not consistently associated with higher-impact HSLC severe convection.

The null event SOM provides valuable context as to the importance of PI release compared to PI presence alone (Figure 3.8). The PI environments are organized by the strength of the TE710 gradient on the x-axis and the orientation of the TE710 gradient on the y-axis. Nodes in the bottom row of the SOM matrix are defined by minimal PI, while a signature of northward extending PI (especially along the eastern half of the domain) is more prevalent in the top row. Despite an increase in PI in nodes in the top half of the SOM, there is minimal mid-level ascent in these environments, and in fact descent in many. This is supported by a similar result in previous literature (Sherburn et al. 2016; Figure 9d-f), where composites of 700-hPa vertical velocity with respect to pressure were significantly weaker in radar-based HSLC null events as compared to HSLC severe events. These TE710 trained SOMs show that some amount of PI is present in all HSLC severe events and in some HSLC null events. The biggest discriminator is the associated mid-level ascent, which acts as a mechanism for PI release - a key ingredient in HSLC severe convection that is missing in the null SOM.

3.2.3 MOSH Patterns

To discern the HSLC environmental patterns associated with varying predictability as determined by environmental parameters used in operational HSLC forecasting, an event-relative SOM was trained using MOSH as the input variable (Figure 3.9). To stay concise, the MOSH-trained SOM was selected for in-depth discussion over the SHERBS3, as the MOSH represents an improvement over the SHERBS3 by eliminating the amount of false alarms through the addition of a term designed to represent the release of PI (Sherburn et al. 2016). Along the y-axis of the SOM matrix, greater MOSH values are located on the bottom row, while lower MOSH values are located on the top row. The x-axis is defined by the zonal location of the MOSH maxima relative to the domain center; the left column contains nodes with MOSH maxima slightly west of the domain center, and the right column features MOSH maxima east of the center. The MOSH maxima in the right column in some cases exhibits substantial displacement from the domain center; the MOSH maxima in nodes 11-12 and 15-16 are located in the most eastward zonal quarter of the domain. These nodes are less populated and exhibit some
combination of a more eastward located surface cyclone and hint at the presence of a warm front than several other nodes that feature centralized MOSH.

The variability in MOSH across nodes can be assessed using node composite parameters that reflect the component terms of MOSH: 0-1 km vertical wind shear (similar to S15MG) and TE710 with 500-hPa vertical velocity (similar to MAXTEVV). Some amount of node composite 0-1 km vertical wind shear exceeding 30 kts is present in all nodes (Figure 3.10). The threshold of 0-1.5 km vertical wind shear required to contribute to non-zero MOSH is 8 m/s (15.6 kts), so 0-1 km vertical wind shear is apparently not driving the magnitude differences in MOSH across SOM nodes. Despite this, nodes with spatial displacement of MOSH centers are associated with spatial displacement in the 0-1 km vertical wind shear maximum and many SOM nodes with prominent MOSH are also aligned closely to the 0-1 km vertical wind shear maximum. The presence and release of PI appears to have a stronger influence on the node averaged MOSH (Figure 3.11). Nodes with higher MOSH values have a more northward extending tongue of PI, less negative PI, and regions of vertical velocity and PI that overlap relative to those with little or no MOSH. This is exemplified by nodes 1 and 2, which exhibit areas of upward vertical velocity primarily where TE710 is positive (a situation that is not conducive to release of PI). Nodes 11-12 and 15-16 also reveal the sensitivity of the spatial distribution of MOSH to upward motion, as a broad area of PI extends much farther northward than in other nodes in regions defined by some upward vertical velocity. While PI release appears to be the key factor at these fixed levels, it is important to note that MAXTEVV is defined as is the maximum $d\theta/dz \times \omega$ product calculated from the 0-2 km layer through the 0-6 km layer at 0.5 km intervals. While the variables at the fixed levels provide some insight into MOSH variability, the same conclusions cannot be made for all levels across which MAXTEVV is calculated. These MOSH patterns are compared to the remaining composite variables (Table 2.2) in Appendix D.

The distribution of LSR frequency across nodes in the MOSH trained SOM establishes the usefulness of MOSH in significant severe environments, as the highest impact nodes are defined by greater maximum values of MOSH in LSR density centers (Figure 3.9). While many of the cases in this dataset are included in minimal MOSH nodes (nodes 2 and 4), node 13 serves as an example of a reasonable group of cases (60) defined by high MOSH values. As discussed previously, the reasons for variability in maximum MOSH values are associated with the variability in vertical velocity in PI environments. The difference in average LSRs between the
nodes with the lowest maximum MOSH (nodes 1-4) and 16 are 2-3x the standard deviation of average LSRs across all nodes (17.3), which reinforces the usefulness of MOSH in determining the impact of HSLC severe events. However, there are nodes with a large eastward extent of high MOSH values that are the least impactful, as determined by the mean and median number of LSRs in each node (Figure 3.12a). This could be due to the lack of LSRs in offshore regions, which tend to characterize the lower and right portions of these LSR-relative domains. Node 1 is interesting in that it features a relatively high average LSR count, but shows very weak composite MOSH. This suggests the dataset includes severe events that do not rely on the release of PI in the convective environment, although they are characterized by similar ambient CAPE values as other nodes (not shown).

With the exception of the top row, a pattern also emerges where the left column of the SOM generally contains the highest impact node of a given row (Figure 3.9). These patterns are also true for the tornado-only LSRs (Figure 3.12b) where the highest tornado impact cases are mostly confined to the bottom left corner SOM nodes. Node 16 exhibits an unusually high frequency of average tornado LSRs compared to the low average of total LSRs, but this is heavily influenced by outliers, partially demonstrated by 0 median LSRs per case (Figure 3.12b). It is noteworthy that regardless of the amount of MOSH present, all nodes contain some amount of tornadic activity, and many include outlier events of 10+ tornado LSRs. Sherburn and Parker (2014) and Sherburn et al. (2016) describe the SHERBS3 and MOSH as parameters being most skillful in discriminating between significant HSLC severe reports from non-severe HSLC storms, where significant severe weather is defined as “tornadoes rated EF2 or higher, wind reports of 65 knots or higher, and hail reports of 2 in. diameter or larger.” The LSRs used in this study did not filter for events meeting these criteria only, and thus some non-significant severe events defined by MOSH<1 are included, which could be the reason for the significant variability in maximum MOSH between nodes. Forecasters using these composite parameters would not know a priori whether a given event will be severe or significant severe.

The null event SOM provides another test of MOSH as a predictor of HSLC severe convection (Figure 3.13). While most events belong to nodes that are defined by zero MOSH, as shown by the number of cases assigned to each pattern, other nodes exhibit areas of MOSH>1 or broad areas of limited MOSH. Although this is mostly explained by the presence of PI in areas of weak upward vertical velocity along the eastern domain boundary (Figure 3.13), nodes in the left
column of the SOM matrix exhibit broader areas of weak PI, and minimal vertical velocity. Examination of individual cases within each node (not shown) provides some insight: these nodes (particularly nodes 9 and 13) contain outlier events where a prominent region of PI extends northward into a region of moderate upward vertical velocity and enhanced 0-1 km vertical wind shear that results in high values of MOSH. The high MOSH values from these outlier events in nodes with limited numbers of cases leaves some residual MOSH in the node composites, and can be seen overlaid on regions of small, but negative TE710 that extend northward. Non-zero MOSH is also not entirely unexpected: Previous literature (Sherburn et al. 2016; Figure 16d-f) show non-zero MOSH in the null composites, much more so in warning nulls. However, the null events defined in Sherburn et al. 2016 required a convection prerequisite that the null events in this study were not subject to.

The MOSH trained SOMs show variability in the magnitude and spatial orientation of MOSH values across the HSLC severe event dataset, often associated with the variability in vertical velocity in PI environments. This is also true for null events, but on a much smaller scale. Most null events are defined by little MOSH, but some outlier events with ample PI environments are still present.

### 3.3 LSR Quartile SOMs

In section 3.2, we saw pronounced differences between patterns of HSLC severe events and null events. While there were general trends in the mean and median values of total and tornado LSRs in the HSLC severe event environmental patterns, the distribution of severe weather would sometimes vary across individual nodes defined by similar patterns. This still prompts the question: are there distinct patterns accompanying high and low LSR activity within the set of HSLC severe events? In order to examine this, the upper- and lower-quartile of LSR events were used to generate SLP-trained SOMs. While the average cyclone intensity is lower across nodes in the upper-quartile SLP SOM (Figure 3.14) as opposed its lower-quartile counterpart (Figure 3.15), there is not a substantial difference in the structure of SLP patterns. In particular, node 7 of the lower-quartile SLP SOM depicts a node-averaged surface cyclone of 996 hPa with a prominent north-to-south pressure trough. This is quite similar to both node 7 of the upper-quartile SLP SOM, and the higher impact nodes of the previously discussed complete
dataset SLP SOM (Figure 3.2). While on average higher-impact HSLC severe events appear to occur in environments defined by stronger surface cyclones and north-to-south oriented pressure troughs, as expected, the variability in SLP patterns across individual events is minimal.

The accompanying instability environments and composite MOSH provide a more illuminating representation of the differences between high and low impact events. On average, there was more instability extending farther northward in the upper-quartile SOM nodes (Figure 3.14) than the lower-quartile SOM nodes (Figure 3.15), and this is represented by a total area of MUCAPE > 500 J/kg that is 92% greater in the upper-quartile SOM. There are also greater MOSH values on average in the upper-quartile SOM, in which the maximum values overlap with the tongue of northward extending instability. MOSH is not absent in the lower-quartile SOM, however, but the average values in most nodes are well below the threshold of significance. The difference in both 0-1 km vertical wind shear and near-surface theta-E advection patterns associated with these SLP environments between quartiles was minimal as similar patterns were present in both (not shown). To better assess if differences in low-level shear or near-surface advection processes are present with varying event impact, LSR quartile SOMs were trained for both variables.

Despite being separated by event impact, the upper-quartile 0-1 km vertical wind shear SOM (Figure 3.16) shows a similar distribution of low-level wind shear patterns as its lower-quartile counterpart (Figure 3.17) as represented by only a 9% greater area of shear > 20 kts in the upper-quartile SOM. Both SOMs contain nodes with low-level wind shear exceeding 40 kts. However, in the lower-quartile SOM, the strong shear nodes (nodes 1 and 4) are defined by a low-level jet east of the domain center with an associated eastward located surface cyclone whereas in the upper-quartile SOM, nodes with similar shear values (nodes 3, 6, 9) are represented by a centrally located low-level jet and surface cyclone. Outside of these nodes, the lower-quartile SOM contains many nodes of similar low-level wind shear and SLP patterns (ex. Node 5 in each SOM). The upper-quartile SOM also contains a node of relatively low shear (node 7; 25 kts) similar to the lower-quartile SOM (node 9; 20 kts). Here, the distribution of low-level wind shear does not show significant variation with event impact.

When comparing the upper-quartile near-surface theta-E advection SOM (Figure 3.18) to its lower-quartile equivalent (Figure 3.19), the well-defined dominance of post-frontal negative theta-E advection is prevalent in both SOMs. However, this is unsurprising given the number of
HSLC severe events that are defined by a QLCS. The upper-quartile SOM contains more nodes with significant positive theta-E advection present (nodes 1, 4, 7, 8) than the lower-quartile SOM (node 9), despite most nodes in both SOMs defined with weak positive theta-E advection. There is always a region of positive theta-E advection that extends into the domain center in the upper-quartile SOM regardless of its strength, whereas the lower-quartile SOM contains nodes defined predominately negative theta-E advection. This is characterized by a 144% greater area of theta-E advection > +10 K/day, but as previously mentioned the dominant feature in these nodes is the negative theta-E advection. The associated surface cyclones and pressure troughs in the upper-quartile SOM are consistently stronger as well. Environments defined by a region of marginal positive theta-E advection ahead of a strong surface cyclone, associated pressure trough, and strong negative theta-E advection are shown to be generally higher impact in this HSLC severe event dataset.

Both the upper- and lower-quartile LSR TE710-trained SOMs show a similar distribution in the orientation of PI environments across all nodes (Figure 3.20 and 3.21). The differences between these SOMs lie in the magnitude of PI and the accompanying 500-hPa upward vertical velocity fields. The upper-quartile TE710 SOM exhibits more positive TE710, less negative TE710, and greater 500-hPa upward vertical velocity values on average across all nodes than the lower-quartile TE710 SOM. PI is present in the lower-quartile TE710 SOM nodes, and is even ample in nodes 6 and 9, but the lack of accompanying upward vertical velocity to represent the release of PI on average is likely a contributing factor to lower-impact severe events. This is demonstrated by comparing the total area of PI release, as defined by negative theta-E difference and upward vertical velocity > 3 -μb/s, where a 209% greater area of PI release is present in the upper-quartile SOM. Variables representing PI release are more prevalent in higher-impact events across the HSLC severe event dataset.

When comparing upper- and lower-quartile number of LSR MOSH SOMs, there is a substantial spread in maximum MOSH values across nodes of the respective SOMs (Figure 3.22 and 3.23). Some pronounced spatial variability of MOSH in nodes 2 and 3 of the upper-quartile SOM are also present, in which the associated PI environment resembles the inter-node event variability presented by nodes in the MOSH null event SOM (Figure 3.13). Nodes 3, 6, and 9 of the upper-quartile MOSH SOM (Fig. 3.22) exhibit mis-match between MOSH maxima and LSR density. However, on average across all nodes, the MOSH values associated with the upper-
quartile MOSH SOM are greater than the MOSH values associated with the lower-quartile MOSH SOM, which can be partially exemplified by a 148% greater area of MOSH > 0.5. Between SOMs, this difference is associated with the greater amounts of PI and 500-hPa upward vertical velocity acting to release PI in the upper-quartile MOSH SOM as compared to the lower-quartile MOSH SOM. The reason for variability across nodes in each individual SOM is explained by the variability between individual events in a given node (not shown). Node 9 in the upper-quartile MOSH SOM is defined by limited MOSH in the node composite because it comprises events exhibiting narrow regions of significant upward vertical velocity in PI environments that vary spatially across events. This results in equally narrow regions of high MOSH values that are smoothed out in the node composite. Nodes 1 and 2 in the lower-quartile MOSH SOM have broad regions of high MOSH values simply because they contain events with regions of significant upward vertical velocity in PI environments over a broad area. There is also some variability in 0-1 km vertical wind shear across nodes in this SOM (not shown), but the nodes with highest values of 0-1 km vertical wind shear are not consistently associated with nodes of high MOSH values. Over the entire HSLC severe event dataset, there are few low-impact HSLC severe events that are defined by ample PI environments with sufficient vertical velocity forcing, thus contributing to high values of MOSH. These events could warrant individual examination in future studies. On average, high values of MOSH are most associated with higher-impact HSLC severe events. However, upon further examination of individual events, examples of outliers are present, and the inter-node variability of events is greater than that of SLP or TE710 (which was represented by higher quantization error and lower inter-node correlation coefficients).
Figure 3.1 Histogram plots depicting the correlation coefficients for all cases within a node with respect to the node composite for A. a SLP trained SOM, B. a 700 - 1000 hPa theta-E difference trained SOM, and C. a MOSH trained SOM. The mean correlation coefficient in a node is indicated in red and the median correlation coefficient in a node is indicated in blue, as in legend.
C

node 1
# of cases: 53, avg LSRs: 26

node 2
# of cases: 100, avg LSRs: 11

node 3
# of cases: 20, avg LSRs: 20

node 4
# of cases: 104, avg LSRs: 28

node 5
# of cases: 29, avg LSRs: 45

node 6
# of cases: 10, avg LSRs: 20

node 7
# of cases: 32, avg LSRs: 34

node 8
# of cases: 51, avg LSRs: 25

node 9
# of cases: 28, avg LSRs: 57

node 10
# of cases: 36, avg LSRs: 34

node 11
# of cases: 13, avg LSRs: 42

node 12
# of cases: 6, avg LSRs: 7

node 13
# of cases: 60, avg LSRs: 68

node 14
# of cases: 37, avg LSRs: 45

node 15
# of cases: 18, avg LSRs: 11

node 16
# of cases: 11, avg LSRs: 17
Figure 3.2 SOM trained on sea level pressure (hPa) from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Pressure is contoured in dashed black every 2 hPa, node-composite MUCAPE is shaded every 500 J/kg, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
**Figure 3.3** Histogram plots depicting the amount of LSRs for all cases within each SLP trained SOM node for **A.** all LSRs and **B.** tornado-only LSRs. Outliers are excluded to better represent node-to-node LSR spread. The mean number of LSRs in a node is indicated in red and the median number of LSRs in a node is indicated in blue, as in legend.
Figure 3.4 SOM trained on sea level pressure (hPa) from HSLC null events. Data are event-relative 2000 km x 2000 km grids centered on 34.5 °N and -84.5 °W. Sea-level pressure is contoured in black every 2 hPa and node-composite MUCAPE is shaded, as in legend. Node number and total number of events listed above each node.
Figure 3.5 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is shaded every 6 K, as in legend, node composite sea-level pressure is contoured in black every 2 hPa, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Node number and total number of events listed above each node.
Figure 3.6 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is contoured in black and shaded in K, as in legend, and node composite 500-hPa vertical velocity is contoured in dashed black every 3 -μb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
**Figure 3.7** Histogram plots depicting the amount of LSRs for all cases within each 700-1000 hPa theta-E difference trained SOM node for A. all LSRs and B. tornado-only LSRs. Outliers are excluded to better represent node-to-node LSR spread. The mean number of LSRs in a node is indicated in red and the median number of LSRs in a node is indicated in blue, as in legend.
Figure 3.8 SOM trained on 700-1000 hPa theta-E difference from HSLC null events. Data are event-relative 2000 km x 2000 km grids centered on 34.5 °N and -84.5 °W. 700-1000 hPa theta-E difference is contoured in black and shaded, as in legend, and node composite 500-hPa vertical velocity is contoured in dashed black every 2 µb/s. Node number and total number of events listed above each node.
Figure 3.9 SOM trained on MOSH from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red every 0.25 node, composite SLP is contoured every 2 hPa, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Node number, total number of events, and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.10 SOM trained on MOSH from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red every 0.25 and node-composite 0-1 km vertical wind shear is shaded, as in legend. Node number, total number of events, and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.11 SOM trained on MOSH from HSLC severe events. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red 0.25, node composite theta-E difference is shaded, as in legend, and node composite 500-hPa vertical velocity is contoured in dashed black every 3 µb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.12 Box and whisker plots depicting the amount of LSRs for all cases within each MOSH trained SOM node for A. all LSRs and B. tornado-only LSRs. Outliers are excluded to better represent node-to-node LSR spread. The mean number of LSRs in a node is indicated in red and the median number of LSRs in a node is indicated in blue, as in legend.
A

node 1

# of cases: 53, avg LSRs: 26

node 2

# of cases: 100, avg LSRs: 11

node 3

# of cases: 20, avg LSRs: 20

node 4

# of cases: 104, avg LSRs: 28

node 5

# of cases: 29, avg LSRs: 45

node 6

# of cases: 19, avg LSRs: 20

node 7

# of cases: 22, avg LSRs: 34

node 8

# of cases: 51, avg LSRs: 25

node 9

# of cases: 28, avg LSRs: 57

node 10

# of cases: 36, avg LSRs: 34

node 11

# of cases: 13, avg LSRs: 42

node 12

# of cases: 6, avg LSRs: 7

node 13

# of cases: 60, avg LSRs: 68

node 14

# of cases: 37, avg LSRs: 45

node 15

# of cases: 10, avg LSRs: 11

node 16

# of cases: 11, avg LSRs: 17
Figure 3.13 SOM trained on MOSH from HSLC null events. Data are event-relative 2000 km x 2000 km grids centered on 34.5 °N and -84.5 °W. MOSH is contoured in red every 0.25 node composite 500-hPa vertical velocity is contoured in dashed black every 3 µb/s and node-composite 700-1000 hPa theta-E difference is shaded, as in legend. Node number and total number of events listed above each node.
Figure 3.14 SOM trained on sea level pressure (hPa) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Pressure is contoured in dashed black every 2 hPa, MUCAPE is shaded, as in legend, and MOSH is contoured in red every 0.25. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.15 SOM trained on sea level pressure (hPa) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Pressure is contoured in dashed black every 2 hPa, MUCAPE is shaded, as in legend, and MOSH is contoured in red every 0.25. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.16 SOM trained on 0-1 km vertical wind shear (kts) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Vertical wind shear is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.17 SOM trained on 0-1 km vertical wind shear (kts) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Vertical wind shear is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.18 SOM trained on near-surface theta-E advection (K/day) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E advection is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.19 SOM trained on near-surface theta-E advection (K/day) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E advection is contoured and shaded, as in legend, and node-averaged sea-level pressure is contoured in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.20 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is contoured and shaded, as in legend, and node-averaged 500-hPa vertical velocity is contoured in black every 3 µb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.21 SOM trained on 700-1000 hPa theta-E difference (K) from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. Theta-E difference is contoured and shaded, as in legend, and node-averaged 500-hPa vertical velocity is contoured in black every 3 µb/s. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.22 SOM trained on MOSH from HSLC severe events in the upper-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red 0.25, node composite theta-E difference is contoured in dashed black and shaded every 6 K, node composite 500-hPa vertical velocity is contoured in dashed black every 3 -µb/s, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure 3.23 SOM trained on MOSH from HSLC severe events in the lower-quartile number of LSRs. Data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude. MOSH is contoured in red 0.25, node composite theta-E difference is contoured in dashed black and shaded every 6 K, node composite 500-hPa upward vertical velocity is contoured in dashed black every 3 -μb/s, and LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
CHAPTER 4

4. Conclusions

Severe convection in HSLC environments is a common threat to the southeastern U.S. and causes a great operational challenge due to high false alarm rates and low probabilities of detection. Previous studies have assessed the patterns and ingredients of HSLC severe convection and verified forecasts of UH in these environments (Sherburn and Parker 2014; Sherburn et al. 2016; King et al. 2017; Graham 2021). However, to the author’s knowledge, little work has been done to objectively classify the different synoptic and meso-beta scale patterns that accompany HSLC severe convection. This study was performed to objectively classify the patterns that comprise southeastern U.S. HSLC severe convection, assess the distribution of severe weather occurrence across these patterns, and determine if these findings are consistent with previously developed HSLC composite parameters.

An automated, objective algorithm was developed to identify HSLC events using historical data from 2008 through 2021. Data sources were local storm reports along with Rapid Refresh and Rapid Update Cycle analyses, obtained from (NOAA/SPC 1999, NOAA NCEP 2005). This large HSLC dataset was cross-checked against earlier HSLC event lists compiled by Keith Sherburn and Chase Graham, and was found to provide an adequate match. Owing to the extensive sample size, a neural-network-based clustering algorithm, self-organizing maps (SOMs), were used to identify characteristic patterns accompanying the HSLC events.

4.1 Summary of Findings

In this study, SOMs were implemented to objectively classify the environmental patterns of 587 HSLC severe events over 13 cool seasons. SOMs were trained on a set of 13 meteorological variables, selected based on the findings of previous HSLC literature (Sherburn and Parker 2014; Sherburn et al. 2016; King et al. 2017) to represent the variability in synoptic and mesoscale ingredients deemed influential to the production of HSLC severe convection and its respective impacts. The main findings of this study are:
• On the average, the objectively defined meteorological patterns are consistent with previous HSLC environmental pattern studies. However, the use of SOMs provided some insight into the variations on these patterns:
  o The most common synoptic structure was a strong surface cyclone, with a north-to-south oriented surface pressure trough and was accompanied by a deep, negatively-tilted mid-level trough. Few events were defined by weaker surface cyclones west of the observation domain with an eastward extending pressure trough (and an implied surface warm front). Events of all synoptic structure were generally characterized by a region of instability extending northward toward the center of LSR density with intensity varying from < 500 J/kg to nearly 1000 J/kg (Figure 3.1).
  o Nearly all events were defined by the presence of potential instability with upward some vertical velocity > 3 -μb/s, that act as release mechanism. Upward vertical velocity was generally strongest in the presence of a strong potential instability gradient, which was aligned with an implied surface cold front (Figure 3.5).
  o Substantial deep-layer wind shear (>40 kts) and lower-tropospheric wind shear (>20 kts) were present in all cases, but the amount of lower-tropospheric shear exhibited greater variability, ranging from limited areas of shear > 20 kts to widespread regions of shear > 35 kts (Figures 3.15 and 3.16).
  o Near-surface negative theta-E advection was a dominant surface process, occurring behind an implied surface cold front, while positive theta-E advection was weaker in comparison, often occurring east of the LSR density centers, and was present in limited amounts in some nodes (Figures 3.17 and 3.18).
  o A high degree of variability in the representation of MOSH across events was observed. Some MOSH nodes were defined by a shift in MOSH maximum values from the LSR density center, or limited depiction of any MOSH values altogether (Figures 3.10, 3.21, 3.22).
• Comparing the patterns of HSLC severe events to HSLC null events revealed pronounced differences on average for all variables (Figures 3.3 and 3.7).
• MOSH was able to discriminate between HSLC severe events and HSLC null events in on average. Instances of enhanced MOSH values in non-severe events were associated with strong synoptic forcing in marginal potential instability environments, which often coincided with coastal regions (Figure 3.12).

• While SOMs trained on some variables showed a general distinction in the distribution of severe weather occurrence across patterns, no one variable was consistently able to distinguish high-frequency LSR events from low-frequency LSR events.

• When separating the dataset into upper and lower quartile LSR events, the release of potential instability was found to show the greatest variation across event impact (Figures 3.19 and 3.20).

• On average, higher-impact HSLC severe events were found to be associated with a strong surface cyclone and an ample potential instability environment with sufficient upward vertical motion overlapping (Figures 3.13, 3.14, 3.19, 3.20).

• Differences in lower-tropospheric wind shear were intensity between LSR quartile SOMs were not prevalent, while positive near-surface theta-E advection was slightly more prominent in high-impact severe events, but negative near-surface theta-E advection was once again the dominant process (Figures 3.15, 3.16, 3.17, 3.18).

• On average, MOSH was able to distinguish between event impact (Figures 3.21, 3.22). Events where MOSH signaled a false alarm, missed, or varied spatially from the center of LSR activity were most associated with variations in the potential instability environment as opposed to lower-tropospheric wind shear.

Ultimately, the key findings of this study are that no one variable was able to consistently demonstrate differences in the distribution of severe weather occurrence across patterns, but processes representing the release of potential instability were more prevalent in higher-impact events on average. Assessment of the variations in meteorological patterns across SOMs mostly show slight variations on the patterns defined in previous literature (Sherburn and Parker 2014, Sherburn et al. 2016, King et al. 2017), but some key points are gleaned. No nodes representing HSLC severe convection along a prominent surface warm front are present and LSR density is nearly always centered in or just ahead of a surface pressure trough, demonstrating the importance of the surface cold front in producing HSLC severe events. The importance of PI
release is also documented, as regions with upward vertical velocity and negative theta-E difference are clearly aligned with the LSR density centers. When directly comparing SOMs of high-impact events to low-impact events, the difference in total area representing PI release between SOMs is greater than similar comparisons of other variables. Additionally, PI release is a critical component in the MOSH composite parameter. As presented in Sherburn et al. (2016), MOSH is most skillful in distinguishing HSCLC significant severe events from HSCLC non-severe events. While assessment of the MOSH is confined to the limitations of this study, there are instances of lower-impact events where MOSH values do not represent severe convection due to either spatial variations or lower parameter values (Figures 3.10 and 3.22). Despite this, MOSH was able to accurately depict HSCLC severe convection in high-impact events on average. It is important to analyze the cases comprising SOM nodes where the MOSH parameter did not match well with LSR density.

4.2 Future Work

There are several methods for improving and expanding upon the work completed in this study. Identifying HSCLC severe events based on LSRs alone and requiring no minimum severe wind or tornado thresholds is limitation in this study. Stricter wind speed or tornado severity requirements could have been implemented to overcome the sources of error associated with LSRs (documented in Section 2), or additional metrics such as radar reflectivity signatures or NWS warnings could have been used to supplement LSRs. In addition, the process of HSCLC severe event identification was automated, and only the local CAPE and wind shear environments around LSRs were examined. Treating these local environments as the method identification could both include or exclude events that a human forecaster may determine to be HSCLC or non-HSCLC respectively. As HSCLC null events contained no LSRs, an adjusted domain was used for examination of the CAPE and shear environments. In retrospect, requiring a higher minimum CAPE threshold over land areas may have produced a more accurate representation of null events. This involves a similar limitation of including or excluding events differently than a human forecaster. Additionally, while classifying events as six-hour windows allows for an objective event definition process, there are some limitations to classifying a six-hour window of severe reports with a single environmental analysis at the event onset time. The rapid evolution
of HSLC environments has been documented (King et al. 2017) and a severe report at 05 Z, for example, may not be well represented by the environment at 00 Z. Future work could involve performing the same SOM analysis, but separating events into hourly increments as LSRs allow rather than characterizing the entire event by window onset time.

NWS warning data could also allow for an improved aspect of predictability in this study. MOSH and SHERBS3 are useful tools in forecasting HSLC severe convection but are documented as composite parameters that do not forecast whether severe convection will occur, rather serving as a discriminator between significant severe and non-severe events. Using warning data in combination with LSRs and radar reflectivity signatures would allow for a binary assessment of an event as a successful forecast. Additionally, this study did not assess the SHERBE or MOSHE composite parameters, where effective bulk shear vector magnitude is incorporated into parameter calculation. The inclusion of these parameters may have provided additional distinction in environmental patterns or distribution of severe weather occurrence.

For specific “interesting” SOM nodes, such as those where there was large composite MOSH and very low LSR density, or vice versa, it would be worthwhile to examine individual cases in order to learn what meteorological conditions were responsible for the mismatch. There are also limitations to the study that are due to limitations in SOMs. We worked with a large sample size, and each case was matched with a “winning” SOM node. While using SOMs allowed for increased variety in the patterns of each meteorological variable within the HSLC subclass of convection, each node is still a composite of the cases assigned to it. Variations in the dimension of the SOM matrix may have revealed additional important patterns that are somewhat rare. At one point, geographically fixed SOMs were constructed, but these were not shown for the sake of brevity, and because the event-relative SOMs produced more interpretable results.
REFERENCES


APPENDICES
Appendix A: SOM Training Variable Output

This appendix documents the event-relative SOM output for the 12 remaining training variables listed in Table 2.2 that were not discussed in Section 3. All output is presented as in Section 3, where data are event-relative 2000 km x 2000 km grids of the respective training variables centered on the event averaged storm report latitude and longitude. LSRs are plotted as the percentage of cases where wind and tornado LSRs occur at a given grid point, shaded as in legend. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.
Figure A1. 500 hPa geopotential height trained SOM. Geopotential height is contoured in solid black every 6 dm.
Figure A2. 500-hPa geopotential height and sea-level pressure trained SOM. Geopotential height is contoured in solid black every 6 hPa and sea-level pressure is contoured in dashed black and shaded every 2 hPa.
Figure A3. CAPE trained SOM. CAPE is contoured and shaded every 500 J/kg.
Figure A4. 0-6 km vertical wind shear trained SOM. Vertical wind shear is contoured and shaded every 5 kts.
Figure A5. 0-1 km vertical wind shear trained SOM. Vertical wind shear is contoured and shaded every 5 kts.
Figure A6. 0-1 km storm relative helicity trained SOM. Storm relative helicity is contoured and shaded in m²/s².
Figure A7. Near-surface temperature advection trained SOM. Temperature advection is contoured and shaded 5°C/day.
Figure A8. Near-surface dew point temperature advection trained SOM. Dew Point temperature advection is contoured and shaded every 6°C/day.
Figure A9. Near-surface theta-E advection trained SOM. Theta-E advection is contoured and shaded 10 K/day.
Figure A10. 500-hPa vertical velocity trained SOM. Vertical velocity is contoured every 3 \( \mu \text{b/s} \).
Figure A11. 700-1000 hPa theta-E difference and 500-hPa vertical velocity trained SOM. Theta-E difference is shaded every 6K and vertical velocity is contoured every 3 µb/s.
Figure A12. SHERBS3 trained SOM. SHERBS3 is contoured in dashed red every 0.25.
Appendix B: Sea-Level Pressure Trained SOM Composite Variables

This appendix documents the 10 remaining composite variables listed in Table 2.2 that accompany sea-level pressure trained SOM but were not presented in Section 3.2.1, as well as low-level lapse rate (LLLR). All output is presented similarly to SOM output presented in section 3, where data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude, but of the composite variables instead of the training variable. Sea-level pressure is plotted in black every 2 hPa. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

The composite 500-hPa geopotential height field (Figure B1) displays the expected three-dimensional synoptic structure associated with a given SLP pattern: strong surface cyclones and pressure troughs are accompanied by deeper, negatively tilted mid-level trough and weaker surface cyclones and pressure troughs are associated with more zonal mid-level flow. The pressure troughs are also defined by a wind shift and near-surface theta-E gradient signifying a surface cold front (Figure B2). All nodes are defined by the presence of potential instability (PI), as defined by 700-1000 hPa theta-E difference, where larger negative values are indicative of more PI (Figure B3). When combined with 500-hPa vertical velocity to represent PI release, no notable trend is observed. Nodes with greater vertical velocity values are not associated with the strongest surface cyclones or a significant PI pattern.

The distribution of maximum 0-6 km vertical wind shear values (Figure B4) is scattered across nodes whereas the nodes with the greatest 0-1 km vertical wind shear (Figure B5) generally correspond to the strongest surface cyclones and mid-level troughs. The same is true for 0-1 km storm relative helicity (Figure B6), where higher values are associated with the strongest surface cyclones and mid-level troughs. Near-surface temperature advection (Figure B7), near-surface dew point temperature advection (Figure B8), and near-surface theta-E advection (Figure B9) exhibit similar patterns across all nodes. Cold- and dry-air advection was more prevalent than warm- and moist-air advection across all nodes, and the nodes with slightly greater values of warm- and moist-air advection were generally those with stronger surface cyclones. However, this trend is not consistent across all nodes.

The node averaged MOSH (Figure B10) and SHERBS3 (Figure B11) provides an aspect of HSLC predictability to SOM analysis, where values >1 are indicative of significant severe weather being more likely to occur (Sherburn et al. 2016). Maximum values of both are centered...
in the domain but are rarely >1 across all nodes. Examination of individual cases showed that while in many events the MOSH and SHERBS3 maxima was in the domain center, there were enough instances of maximum values being displaced from the center that many values >1 were smoothed out. Ultimately, the usefulness of MOSH and SHERBS3 as a composite variable in this study was determined by comparing values across nodes instead of assessing whether the threshold of 1 was met. Although the MOSH showed some variability in maximum values between SLP patterns, was not a significant difference in MOSH or SHERBS3 values across nodes. The nodes with strong surface cyclones are generally defined by relatively high values of MOSH and SHERBS3, but the remaining nodes do not exhibit a significant departure from these values. Additionally, the difference between the highest and lowest maxima do not correspond to a pattern in SLP. The one distinction in these parameters, greater values of MOSH in the right column of the SOM matrix (Figure B10), can be traced back to the 0-1 km vertical wind shear composite parameter (Figure B5), which closely resembles of the variables used in calculating MOSH (0-1.5 km vertical wind shear). Additionally, these areas of greater MOSH values coincide with the area of steepest low-level lapse rates (Figure B12), which are most commonly found ahead of a southward extending pressure trough.
Figure B1. Composite 500-hPa geopotential height field for sea-level pressure trained SOM. Geopotential height is contoured and shaded in dm, as in legend.
Figure B2. Composite surface theta-E field and surface wind barbs for sea-level pressure trained SOM. Theta-E is contoured and shaded in K, as in legend.
Figure B3. Composite 500-hPa vertical velocity and 700-1000 hPa theta-E difference fields for sea-level pressure trained SOM. Theta-E difference is shaded in K, as in legend, and vertical velocity is contoured every 3 µb/s.
**Figure B4.** Composite 0-6 km vertical wind shear field for sea-level pressure trained SOM. Vertical wind shear is contoured and shaded in kts, as in legend.
Figure B5. Composite 0-1 km vertical wind shear field for sea-level pressure trained SOM. Vertical wind shear is contoured and shaded in kts, as in legend.
**Figure B6.** Composite 0-1 km storm-relative helicity field for sea-level pressure trained SOM. Storm relative helicity is contoured and shaded in m²/s², as in legend.
Figure B7. Composite near-surface temperature advection field for sea-level pressure trained SOM. Temperature advection is contoured and shaded in °C/day, as in legend.
Figure B8. Composite near-surface dew point temperature advection field for sea-level pressure trained SOM. Dew point temperature advection is contoured and shaded in °C/day, as in legend.
Figure B9. Composite near-surface theta-E advection field for sea-level pressure trained SOM. Theta-E advection is contoured and shaded in K/day, as in legend.
Figure B10. Composite MOSH field for sea-level pressure trained SOM. MOSH is contoured in dashed red every 0.25.
Figure B11. Composite SHERBS3 field for sea-level pressure trained SOM. SHERBS3 is contoured in dashed red every 0.25.
Figure B12. Composite low-level lapse rate field for sea-level pressure trained SOM. Lapse rate is contoured in K, as in legend.
Appendix C: 700-1000 hPa Theta-E Difference Trained SOM Composite

Variables

This appendix documents the 10 remaining composite variables listed in Table 2.2 that accompany the 700-1000 hPa theta-E difference trained SOM, as well as low-level lapse rate (LLLR). All output is presented similarly to SOM output presented in section 3, where data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude, but of the composite variables instead of the training variable. 700-1000 hPa theta-E difference is contoured in black every 6 K. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

The HGT500 patterns associated with the TE710 trained SOM show similar orientation and strength of a lifting, negatively-tilted mid-level trough across all nodes (Figure C1). The near-surface theta-E field is unsurprisingly similar to TE710, and the surface wind shift along the theta-E gradient, with increased surface convergence present with a stronger theta-E gradient, indicates a surface cold front (Figure C2). The composite CAPE fields (Figure C3) follow the patterns of PI as expected, where the maximum values are associated with greater PI and the orientation of higher CAPE values follow the northward extent of PI and TE710 gradient. Many of the other variables show little distinction between nodes, including 0-6 km vertical wind shear (Figure C4), 0-1 km vertical wind shear (Figure C5), and 0-1 km storm-relative helicity (Figure C6). The distribution of nodes with the greatest values for each of these variables is generally along the left column and bottom row of the SOM matrix where there are limited amounts of PI and stronger TE710 gradients. The near-surface advection variables (Figures C7-C9) are strongest in terms of both positive and negative advection when the TE710 gradient is strongest. However, the cold- and dry-air advection is more dominant and the warm- and moist-air advection is generally displaced from the center of the domain even when stronger than average.

The distribution of MOSH (Figure C10) and SHERBS3 (Figure C11) across nodes shows similar maximum values for both composite parameters. MOSH is generally more centralized when a strong TE710 or well-defined tongue of northward extending PI is present, but the maximum values do not follow a particular trend in TE710. The same is true for SHERBS3, which shows broad areas of smaller SHERBS3 values (<0.75) and similar maximum values.
across all nodes. Similarly, broad areas of maximum low-level lapse rates (Figure C12) are found in the regions of maximum theta-E difference.
Figure C1. Composite 500-hPa geopotential height field for 700-1000 hPa theta-E difference trained SOM. Geopotential height is contoured and shaded in dm, as in legend.
**Figure C2.** Composite surface theta-E field and surface wind barbs for 700-1000 hPa theta-E difference trained SOM. Theta-E is contoured and shaded in K, as in legend.
Figure C3. Composite MUCAPE field for 700-1000 hPa theta-E difference trained SOM. MUCAPE is contoured and shaded in J/kg, as in legend.
Figure C4. Composite 0-6 km vertical wind shear field for 700-1000 hPa theta-E difference trained SOM. Vertical wind shear is contoured and shaded in m/s, as in legend.
Figure C5. Composite 0-1 km vertical wind shear field for 700-1000 hPa theta-E difference trained SOM. Vertical wind shear is contoured and shaded in m/s, as in legend.
Figure C6. Composite 0-1 km storm-relative helicity field for 700-1000 hPa theta-E difference trained SOM. Storm-relative helicity is contoured and shaded in m²/s², as in legend.
Figure C7. Composite near-surface temperature advection field for 700-1000 hPa theta-E difference trained SOM. Temperature advection is contoured and shaded in °C/day, as in legend.
Figure C8. Composite near-surface dew point temperature advection field for 700-1000 hPa theta-E difference trained SOM. Dew point temperature advection is contoured and shaded in °C/day, as in legend.
Figure C9. Composite near-surface theta-E advection field for 700-1000 hPa theta-E difference trained SOM. Theta-E advection is contoured and shaded in K/day, as in legend.
Figure C10. Composite MOSH field for 700-1000 hPa theta-E difference trained SOM. MOSH is contoured in dashed red every 0.25.
Figure C11. Composite SHERBS3 field for 700-1000 hPa theta-E difference trained SOM. SHERBS3 is contoured in dashed red every 0.25.
Figure C12. Composite low-level lapse rate field for 700 - 1000 hPa theta-E difference trained SOM. Lapse rate is contoured in K, as in legend.
Appendix D: MOSH Trained SOM Composite Variables

This appendix documents the 8 remaining composite variables listed in Table 2.2 that accompany the MOSH trained SOM, as well as low-level lapse rate (LLLR). All output is presented similarly to SOM output presented in section 3, where data are event-relative 2000 km x 2000 km grids centered on the event averaged storm report latitude and longitude, but of the composite variables instead of the training variable. MOSH is contoured in black every 0.25. Total number of events and the average number of wind and tornado LSRs for each node listed above each node.

The synoptic environment is similar in the mid-levels, where a slightly negatively-tilted trough with maximum MOSH values in the downstream side of the trough is present in all nodes (Figure D1). Regions of greater near-surface theta-E values extending northward with accompanying southerly surface winds are present in areas of contoured MOSH (Figure D2). The magnitude of CAPE across nodes do no vary substantially, but all exhibit similar patterns of a northward extending region of instability that matches the near-surface theta-E (Figure D3). The zonal location of northward extending instability matches the zonal location of MOSH, but do not explain the differences in maximum MOSH values. The differences in 0-6 km vertical wind shear (Figure D4) and 0-1 km storm-relative helicity (Figure D5) are relatively minimal and do not align with the differences in MOSH. There is a slight indication of strength in deep-layer shear aligning with greater values of MOSH, but it is not a consistent trend.

Near-surface advection variables (Figures D6-D8) show similar patterns: while there is weak warm- and moist-air advection, the cold- and dry-air advection in the western region of the domain is more dominant. Nodes with maximum MOSH values displaced to the east are defined by a more distinct region of maximum near surface warm- and moist-air advection than other nodes, also located in the eastern portion of the domain. Comparing MOSH and SHERBS3 (Figure D9) shows the difference in accounting for PI release. The amount of SHERBS3 across nodes doesn’t vary as significantly as MOSH and there are instances of notable composite SHERBS3 values where little or no MOSH is present. The area defined by the largest low-level lapse rates is collocated with the maximum MOSH values (Figure D10) and the overall broad regions of relatively steep lapse rates are aligned with the broad SHERBS3 contours.
Figure D1. Composite 500-hPa geopotential height field for MOSH trained SOM. Geopotential height is contoured and shaded in dm, as in legend.
Figure D2. Composite surface theta-E field and surface wind barbs for MOSH trained SOM. Theta-E is contoured and shaded in K, as in legend.
Figure D3. Composite MUCAPE field for MOSH trained SOM. MUCAPE is contoured and shaded in J/kg, as in legend.
Figure D4. Composite 0-6 km vertical wind shear field for MOSH trained SOM. Vertical wind shear is contoured and shaded in m/s, as in legend.
Figure D5. Composite 0-1 km storm relative helicity field for MOSH trained SOM. Storm-relative helicity is contoured and shaded in m²/s², as in legend.
Figure D6. Composite near-surface temperature advection field for MOSH trained SOM. Temperature advection is contoured and shaded in °C/day, as in legend.
Figure D7. Composite near-surface dew point temperature advection field for MOSH trained SOM. Dew point temperature advection is contoured and shaded in °C/day, as in legend.
Figure D8. Composite near-surface theta-E advection field for MOSH trained SOM. Theta-E advection is contoured and shaded in K/day, as in legend.
Figure D9. Composite SHERBS3 field for MOSH trained SOM. SHERBS3 is contoured in dashed red every 0.25.
Figure D10. Composite low-level lapse rate field for MOSH trained SOM. Lapse rate is contoured in K, as in legend.