

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Dissertations & Theses in Earth and
Atmospheric Sciences

Earth and Atmospheric Sciences, Department
of

Fall 12-6-2013

ENVIRONMENTAL ANALYSIS OF GOES-R PROVING GROUND CONVECTION INITIATION FORECASTING ALGORITHMS

Jason M. Apke

University of Nebraska-Lincoln, jason.apke@gmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/geoscidiss>



Part of the [Atmospheric Sciences Commons](#)

Apke, Jason M., "ENVIRONMENTAL ANALYSIS OF GOES-R PROVING GROUND CONVECTION INITIATION FORECASTING ALGORITHMS" (2013). *Dissertations & Theses in Earth and Atmospheric Sciences*. 46.
<https://digitalcommons.unl.edu/geoscidiss/46>

This Article is brought to you for free and open access by the Earth and Atmospheric Sciences, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Dissertations & Theses in Earth and Atmospheric Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

ENVIRONMENTAL ANALYSIS OF GOES-R PROVING GROUND CONVECTION
INITIATION FORECASTING ALGORITHMS

By

Jason M. Apke

A THESIS

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Master of Science

Major: Earth and Atmospheric Sciences
Under the Supervision of Professor Mark R. Anderson

Lincoln, Nebraska

November, 2013

ENVIRONMENTAL ANALYSIS OF THE GOES-R PROVING GROUND
CONVECTION INITIATION FORECASTING ALGORITHMS

Jason M. Apke, M.S.

University of Nebraska, 2013

Adviser: Mark R. Anderson

The enhanced temporal and spatial resolution of the GOES-R series will allow for the use of cloud top cooling based convection initiation (CI) forecasting algorithms. Two such algorithms have been created on the current generation of GOES satellites: the University of Wisconsin cloud top cooling algorithm (UWCTC) and the University of Alabama-Huntsville's satellite convection analysis and tracking algorithm (SATCAST). Preliminary analysis of algorithm products has led to speculation over pre-convective environmental effects on algorithm performance, which this study aims to examine. CI indications are used with objective segmentation tools to identify and cluster radar objects over the Great Plains based on reflectivity quantitative and spatial thresholds. The identified clusters are tracked objectively to identify points of CI. Any SATCAST or UWCTC indication that corresponds with (without) an evaluated initiation point within an hour is considered a positive (false) indication. The objective approach is compared to a small-scale hand validation for optimal results. 17 pre-convective environmental variables are examined for the positive and false indications to improve algorithm output. The total dataset consists of two time periods, one in the late convective season of 2012 and one in the early convective season of 2013. Data are examined for environmental relationships using principal components analysis (PCA) and quadratic discriminant

analysis (QDA). Significant differences are determined for pre-convective environmental variables between positive and false indications. Data fusion by QDA is tested for SATCAST and UWCTC on five separate case study days to determine if application of environmental variables improves satellite-based CI forecasting. PCA and significance testing revealed that positive indications favored environments with greater instability (CAPE), less stability (CIN) and more low-level convergence. The QDA improved both algorithms on all five case studies using significantly different variables. This study provides a preliminary examination of environmental effects on the performance of GOES-R proving ground CI forecasting algorithms, and shows that probability-based discrimination on the algorithms using environmental variables will ultimately help the situational awareness of a nowcaster.

ACKNOWLEDGEMENTS

I personally wish to thank Dr. Mark Anderson and Dan Nietfeld for the wonderful direction and operational expertise they provided. I would also like to thank Dr. Adam Houston and Dr. Matthew Van Den Broeke for the advice they have provided as my committee members along the way. This project was also greatly aided by the cooperation from multiple institutions. I personally would like to thank Dr. John Mecikalski, John Walker, and Chris Jewett from the University of Alabama-Huntsville and Justin Sieglaff and Wayne Feltz from the University of Wisconsin at Madison. This project also benefitted from the direction and feedback provided by forecasters at the National Weather Service in Omaha/Valley, Nebraska. I would also like to thank the many undergraduate students who have provided assistance towards the completion of this project, including Adam Taylor, Casey Griffin, Nathan Jurgensen and Katie Rourke. I am also very grateful for the help provided by Noah Lock, who assisted with objective segmentation and tracking for radar and satellite data.

I am also thankful to the University of Nebraska-Lincoln Earth and Atmospheric Sciences department, who provided the facilities and computational power which made this project possible. Finally, I am indebted to COMET for two years of funding throughout the pursuit of my master's degree. This project was funded in part by a subaward with UCAR under the sponsorship of NOAA/DOC as part of the COMET Cooperative project Z11-91842.

Contents	
Acknowledgements	i
Abbreviations	iii
List of Tables	v
List of Figures	viii
Chapter 1. Introduction	1
Chapter 2. Background	4
a. Algorithms	5
b. Indication Validation	10
c. Environmental Analysis	12
Chapter 3. Methodology	14
a. National Mosaic Quantitative Precipitation Estimation	16
b. Rapid Refresh	17
c. Defining Convection Initiation	18
d. Hand Validation Technique	22
e. Algorithm Validation	23
f. Environmental Analysis	30
g. Data Fusion	34
h. Case Studies	42
Chapter 4. Hand Validation	48
Chapter 5. Results	56
a. Principal Components Analysis	56
b. Variable Contribution	69
c. Case Studies	89
Chapter 6. Summary and Conclusions	107
Chapter 7. References	114
Chapter 8. Data Sources	119

ABBREVIATIONS

3DVAR	Three-dimensional variational assimilation	GSI	Grid point statistical interpolation
ANC	Auto-nowcast system	HWT	Hazardous Weather Testbed
ANCOVA	Analysis of covariance	IR	Infrared
ANOVA	Analysis of variance	LCL	Lifting condensation level
CAPE	Convective available potential energy	LFC	Level of free convection
CAPPI	Constant Altitude Plan Position Indicator	LOWLR	0-3 km lapse rate
CCL	Convective condensation level	LPL	Lifted parcel level
CI	Convection initiation	LPLD	Lifted parcel level divergence
CIN	Convective inhibition	MIDLR	700 – 500 hPa lapse rate
COLIDE	Convergence line detection	MW	Mann-Whitney
CSI	Critical success index	MWD	Mean wind differential
CTC	Cloud top cooling	NCAPE	Normalized convective available potential energy
dBZ	Decibels relative to Z	NCIN	Normalized convective inhibition
EL	Equilibrium level	NMQ	National mosaic quantitative precipitation estimation dataset
FAR	False alarm ratio	NWP	Numerical weather prediction
GOES-R	Geostationary Operational Environmental Satellite - R Series	PC1	First principal component

PC2	Second principal component	SoS	Strength of signal
PC3	Third principal component	Tb	Brightness temperature
PC4	Fourth principal component	ThOR	Thunderstorm observations by radar tracking algorithm
PC5	Fifth principal component	UWCTC	University of Wisconsin cloud top cooling algorithm
PCA	Principle components analysis	WDSS-II	Warning decision support services-integrated information
POD	Probability of detection	WSR-88D	Weather Surveillance Radar 1988 Doppler
QDA	Quadratic discriminant analysis	Z _{EL-LFC}	Depth from LFC to EL
RAP	Rapid refresh model	Z _{LFC-LCL}	Depth from LCL to LFC
RH	Layer relative humidity		
RUC	Rapid update cycle model		
SATCAST	Satellite convection analysis and tracking algorithm		
SM	Storm motion		

List of Tables

Table 3.1	List of variables used with respective abbreviations and descriptions.	31
Table 3.2	Transformations done to normalize convective variable distributions. These transformations are used in the statistical tests and results are reported in the original forms.	37
Table 3.3	Example confusion matrix. Columns represent QDA classification, rows represent actual classification. Percentage of improvement is found using $j = 100\% * ((\frac{a}{e}) - (\frac{c}{d}))$.	41
Table 4.1	Table 4.1. Validation of WDSS-II/ThOR CI clusters compared to subjectively identified truth CI clusters. WDSS-II settings presented as low smooth (50th percentile smoothing over a 5x5 box), medium smooth (50th percentile smoothing over a 7x7 box) and high smooth (50th percentile smoothing over a 9x9 box).	50
Table 5.1	July UWCTC PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.	57
Table 5.2	April UWCTC PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.	60
Table 5.3	July SATCAST PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.	64
Table 5.4	April SATCAST PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.	66

Table 5.5	P-values for UWCTC tests for significant difference between positive and false indications for the total dataset (combined July and April). Variables with 99.5% significance are shown in bold and stars indicate that the variable was omitted from the test.	80
Table 5.6	P-values for SATCAST tests for significant difference between positive and false indications for the total dataset (combined July and April). Variables with 99.5% significance are shown in bold.	82
Table 5.7	P-values for UWCTC tests for significant difference between positive and false indications for the July study period. Variables with 99.5% significance are shown in bold.	83
Table 5.8	P-values for SATCAST tests for significant difference between positive and false indications for the July study period. Variables with 99.5% significance are shown in bold.	84
Table 5.9	P-values for UWCTC tests for significant difference between positive and false indications for the April study period. Variables with 99.5% significance are shown in bold.	86
Table 5.10	P-values for SATCAST tests for significant difference between positive and false indications for the April study period. Variables with 99.5% significance are shown in bold.	87
Table 5.11	20 May 2013 UWCTC performance of QDA removing insignificant variables diagnosed by ANOVA, ANCOVA and MW.	90
Table 5.12	20 May 2013 SATCAST performance of QDA removing insignificant variables diagnosed by ANOVA, ANCOVA and MW.	91
Table 5.13	UWCTC and SATCAST percentage of improvement values for all case studies using all three statistical approach solutions for QDA.	93
Table 5.14	Table 5.14. 9 April 2013 confusion matrices for UWCTC and SATCAST performance of QDA removing insignificant variables diagnosed by MW.	94
Table 5.15	20 June 2013 – 22 June 2013 confusion matrices for UWCTC and SATCAST performance of QDA removing insignificant variables diagnosed by MW.	100

Table 5.16	25 June 2013 confusion matrices for UWCTC and SATCAST performance of QDA removing insignificant variables diagnosed by MW.	101
------------	--	-----

List of Figures

Figure 1.1	Conceptual model of cloud top temperature with respect to convective mode and composite radar from time t to $t + 3\Delta t$ (Adapted from Hartung et al. 2013).	2
Figure 2.1	Schematic of boxed average tracking. 7×7 pixel box drawn around a pixel of interest outlined in bold with example brightness temperatures (T_b) in $^{\circ}\text{C}$ at time t (T_{b1}) and time $t + \Delta t$ (T_{b2}). <i>(Top)</i> Brightness temperature grid with a pixel to be box averaged in bold. The same box is superimposed on time $t + \Delta t$. <i>(Bottom)</i> Box averaged values shown for each time, with a box averaged CTC $T_{b1} - T_{b2} = -1^{\circ}\text{C}$ from t to $t + \Delta t$.	7
Figure 2.2	Schematic of SATCAST object tracking methodology. Objects 1, 2, and 3 are resolved at time t . Objects 4, 5 and 6 are resolved at time $t + \Delta t$. Arrows represent horizontal atmospheric motion vectors. Dashed lines represent extrapolated forward using horizontal atmospheric motion vectors. Step 2 has objects reassigned from step 1 based on the position estimates, where dotted lines represent previous object location.	10
Figure 3.1	A) Total area of study (solid outline). B) Parallax and track-corrected area of study (dashed outline ranging from $91 - 104^{\circ}\text{W}$, $34 - 48^{\circ}\text{N}$). C) CI truth cluster validation domain (dotted outline).	15

Figure 3.2	Schematic of ThOR tracking. Light gray arrows represent NARR storm motion data. Solid shapes represent current cluster locations. The first two points of track are represented with black dots, in addition to three possible tracks shown as dashed lines. The ThOR track guess shown is based on NARR storm motion and previous direction travelled from point 1 to point 2. ThOR chooses between the optional tracks based on the distances (error) from the guess centroid to the optional observed centroids (options 1, 2 and 3). In this figure, ThOR would test all three possible tracks, and choose the track with the least overall mean error (adapted from updraft.unl.edu/thor/wiki).	21
Figure 3.3	Schematic of nearest neighbor grid remapping. Values of the old grid (large squares) are reassigned to the closest new grid pixels (small squares).	24
Figure 3.4	Three different environments (notated A, B and C) of validation areas (solid polygons) and a validation area that is advected forward (dotted polygons). Solid circles indicate locations of CI, and dashed circles indicate original cloud object location.	27
Figure 3.5	Sample variable distribution comparison performed by QDA. Two group distributions are presented, Group A and Group B. QDA will use these distributions on and classify new data into groups based on probability (area under the curve). For example, if these distributions are used to classify a new point with a variable value of 100, that variable would be assigned to group A given a higher probability. In this paper this technique is performed on a multivariate basis.	39
Figure 3.6	1800 UTC 9 April 2013 RAP MSLP and surface temperature with surface boundaries.	43
Figure 3.7	1800 UTC 20 May 2013 RAP MSLP and surface dew point with surface boundaries.	44
Figure 3.8	0000 UTC 22 June 2013 RAP MSLP and MUCAPE with surface boundaries.	46
Figure 3.9	0200 UTC 25 June 2013 RAP MSLP and surface dew point.	47

Figure 4.1	a) Example of unsmoothed composite NMQ data over the Texas panhandle and southwest Oklahoma. b) Light smoothing applied, 50 th percentile filter over a 5x5 box. c) Medium smoothing applied, 50 th percentile filter over a 7x7 box. d) Heavy smoothing, 50 th percentile filter over a 9x9 box. Colors hotter than yellow are above 35 dBZ and considered convective.	50
Figure 4.2	Radar shown in dBZ at three time steps, where initiation occurs at time t . Black outlined polygons represent CTC indications. a) Example of a positive post-CI CTC indication, where reflectivity values intensify after cooling. b) Example of a false CTC detection occurring after CI, with no nearby ongoing convection.	52
Figure 4.3	1910 UTC 20 May 2013 example of validated UWCTC and SATCAST data one hour before the Moore, OK, tornado with contoured MSLP (hPa). Positive indications are shown in green (blue) for SATCAST (UWCTC). False indications are red (yellow) for SATCAST (UWCTC).	54
Figure 5.1	July UWCTC study period plot of the first (x axis) and second (y axis) principal components. Positive (false) UWCTC indications shown in green (red).	58
Figure 5.2	April UWCTC study period plot of the first (x axis) and second (y axis) principal components. Positive (false) UWCTC indications shown in green (red).	62
Figure 5.3	July SATCAST study period PCA with the first principal component on the x axis and the second principal component on the y axis. Positive (false) SATCAST indications are shown in green (red).	63

Figure 5.4	April SATCAST study period PCA with the first principal component on the x axis and the second principal component on the y axis. Positive (false) SATCAST indications are shown in green (red).	68
Figure 5.5	July, April and total UWCTC (left) and SATCAST (right) study period RAP values for positive (green) and negative (red) indications. Boxes represent the 25 th to 75 th percentiles and whiskers are from the 10 th to 90 th percentiles. Medians are shown as solid lines within the boxes.	70
Figure 5.6	July, April and total SATCAST study period RAP values for positive (green) and negative (red) indications. Boxes represent the 25 th to 75 th percentiles and whiskers are from the 10 th to 90 th percentiles. Medians are shown as solid lines within the boxes.	88
Figure 5.7	1832 UTC 9 April 2013 unfiltered SATCAST (top left) and QDA-filtered SATCAST (top right). The green circle displays an area where QDA successfully removes non convective SATCAST indications. 1835 UTC 9 April 2013 composite NMQ radar reflectivity data (dBZ) inside the rectangle (bottom left) and 1935 UTC 9 April composite radar reflectivity (dBZ) displaying several CI events (>35 dBZ) successfully forecast by SATCAST (bottom right).	95
Figure 5.8	1825 UTC 20 May 2013 unfiltered (top left) and QDA-filtered (top right) SATCAST data. The green circle highlights an area where false detections are correctly removed. 1910 UTC 20 May 2013 NMQ reflectivity data (dBZ) over central Oklahoma (bottom left) and one hour later (bottom right).	96

Figure 5.9	1855 UTC 9 April 2013 UWCTC product with one positive detection in northern Kansas and 3 false detections (top left) with a QDA filter applied to the product (top right). 1855 UTC 9 April 2013 NMQ reflectivity data (dBZ) (bottom left) and one hour later (bottom right). Two false detections are correctly dropped (green circle) at the cost of one positive (red circle).	97
Figure 5.10	1910 UTC 20 May 2013 UWCTC product with QDA filtering and the maximum value of measured cooling for the Moore, Oklahoma, tornadic supercell (left) and 1815 UTC 20 May 2013 example of correctly QDA-filtered UWCTC product circled in red (right). For radar refer to Figure 5.9.	98
Figure 5.11	0055 UTC 22 June 2013 unfiltered (top left) and QDA-filtered (top right) SATCAST data. The green circle highlights an area where false detections are correctly removed. 0125 UTC 22 June 2013 NMQ reflectivity data (dBZ) over central Nebraska (bottom left) and one hour later (bottom right).	102
Figure 5.12	0125 UTC 22 June 2013 QDA-filtered UWCTC product over northern Nebraska (left) and 0155 UTC 22 June 2013 example of two correctly removed products in the non-convective area of northern Missouri (right).	103
Figure 5.13	0215 UTC 25 June 2013 unfiltered (top left) and QDA-filtered (top right) SATCAST data. The green circle highlights an area where false detections are correctly removed. 0215 UTC 22 June 2013 NMQ reflectivity data (dBZ) over southern Nebraska (bottom left) and one hour later (bottom right).	104
Figure 5.14	0815 UTC 25 June 2013 QDA-filtered UWCTC product over Iowa (left) and 0945 UTC 25 June 2013 example of a UWCTC removed value.	105

1. Introduction

Nowcasting convection initiation (CI) by satellite is an established and skillful new technique that will improve with the launch of new instrumentation (Mecikalski and Bedka 2006; Sieglaff et al. 2011). The use of cloud top cooling (CTC) brightness temperature changes to forecast CI was explored by Roberts and Rutledge (2003) who found through comparisons of Geostationary Operational Environmental Satellite (GOES) imagery and Weather Surveillance Radar-1988 Doppler (WSR-88D) that the first 35 dBZ echo occurred approximately 30 min after large cooling rates were observed at the cloud tops of subfreezing ($>0^{\circ}\text{C}$) quasi-stationary cloud pixels. As immature cumulus builds vertically, temperature observed at the cloud top decreases long before the first radar echo is observed (Figure 1.1). Roberts and Rutledge (2003) also observed that CTC can discriminate between storms with weak precipitation (<35 dBZ) and storms with strong precipitation (>35 dBZ).

The discovery of nowcasting using CTC led to the development of two GOES infrared (IR) based CI forecasting algorithms. The University of Wisconsin-Madison cloud top cooling algorithm (UWCTC) uses a computationally inexpensive boxed averaging method to track cloud objects with complex logic to filter out horizontal motion based cooling (Sieglaff et al. 2011). The University of Alabama at Huntsville satellite convection analysis and tracking algorithm (SATCAST) uses mesoscale atmospheric motion vectors to track cloud objects of interest and monitor changes in several spectral trends (Mecikalski and Bedka 2006; Walker et al. 2012).

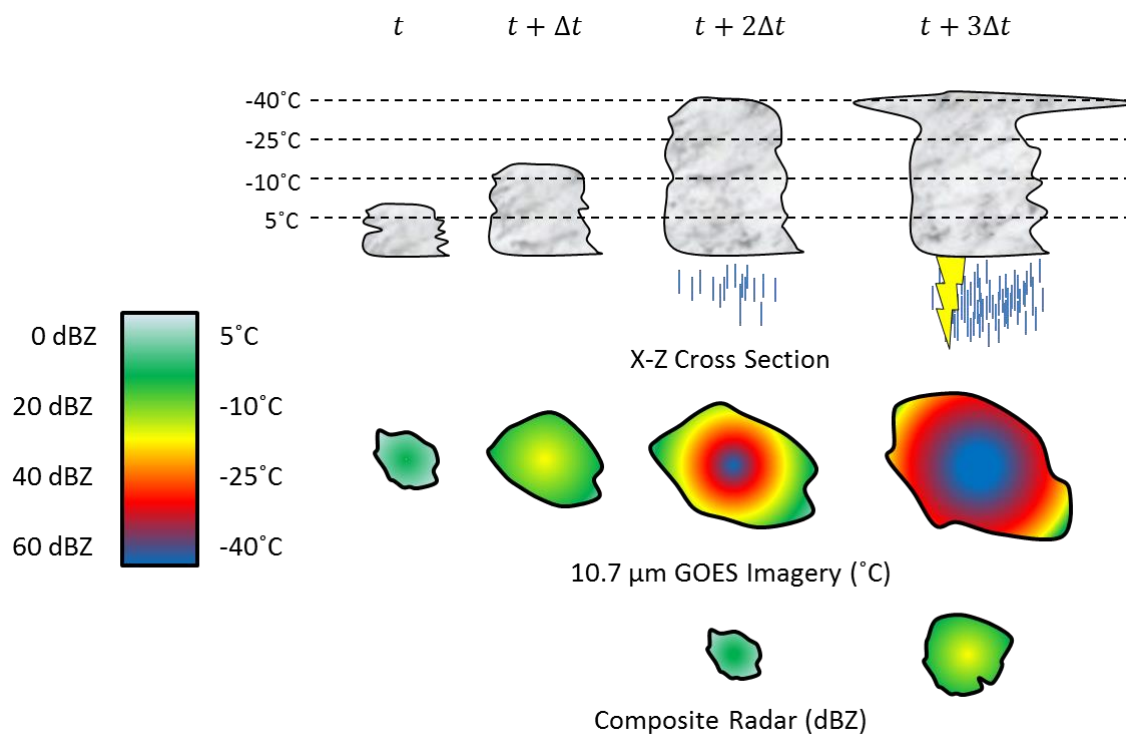


Figure 1.1. Conceptual model of cloud top temperature with respect to convective mode and composite radar from time t to $t + 3\Delta t$ (Adapted from Hartung et al. 2013).

While previous studies have done basic validations to test the skill of CI products (Mecikalski et al. 2008; Hartung et al. 2013), only a few have speculated on the impact of an unfavorable convective environment (Mecikalski et al. 2008; Sieglaff et al. 2011; Walker et al. 2012). The goal of this work is to consider these pre-convective environments and what effect they could have on algorithm performance. This study is designed to answer three questions: Do pre-convective environments have an impact on satellite-based CI forecasting algorithms? Can pre-convective environmental variables be applied to new satellite products through a data fusion process? Will the resulting product help the situational awareness of a nowcaster? To answer these questions, an objective validation methodology is developed to examine large samples of both SATCAST and UWCTC products. Positive and false indications of both algorithms are examined with respect to pre-convective environments derived through numerical weather prediction (NWP) models using principal components analysis (PCA) and tests for statistically significant differences. Environmental variables identified to hinder satellite-based CI forecasting are examined for their potential in a data fusion process. Data fusion is tested using a quadratic discriminant analysis (QDA) approach on select case studies during the 2013 convective season. It is the goal of this work to explore the utility of data fusion and if it will ultimately help the situational awareness of a nowcaster.

2. Background

Convection initiation occurs as a parcel of air reaches its level of free convection (LFC) and achieves and maintains positive buoyancy under a significant upward vertical excursion (Markowski et al. 2006). Johns and Doswell (1992) explain that in order for convection initiation to occur, there must be a moist layer of sufficient depth in the low to mid troposphere, a sufficient layer of instability, and sufficient lifting for a parcel of air in the moist layer to reach its LFC. Failure of CI, however, is not dependent on the lack of these variables. Several studies have been implemented with the goal of finding the CI on/off switch (Markowski et al. 2006). While NWP models offer substantial guidance towards forecasting CI, lack of spatial and temporal resolution of complex nonlinear processes lead to significant errors in timing and location of forecasts (Browning et al. 2007).

Nowcasting the onset of convection is normally performed by first identifying convergent boundaries. While mesonets provide a useful dataset to identify areas of localized surface-based convergence, they only provide marginal utility to forecasting the timing and specific locations of CI (Mueller et al. 1993). Convergent boundaries can be located as fine lines by radar detection (Roberts and Rutledge 2003). However, automated detection of radar fine lines by algorithms often fails over multiple radar sights (Roberts et al. 2012). Algorithms such as the auto-nowcaster (Mueller et al. 2003) benefit from the input of convergent boundaries by forecasters. However, the lack of a robust observation network for moisture and flow through the depth of the boundary layer can make nowcasting convection a challenge when boundaries are not obvious

through WSR-88D detection (Mueller et al. 2003; Mecikalski and Bedka 2006). Thus a need for a reliable detection of both kinematics and low-level moisture arises.

Mecikalski and Bedka (2006) discuss that complex CI processes can be resolved well by satellite throughout evolution from cumulus to cumulonimbus, and that the amount of available data in near real time makes satellite-based CI nowcasting a reasonable new method. The value of satellite data combined with new high temporal resolution data available after the launch of GOES-R led to the development of UWCTC and SATCAST.

a. Algorithms

The UWCTC algorithm uses GOES 13 data and classifies the cloud types using methodology from Pavlonis (2010). The typing algorithm uses a clear sky correction and a cloud mask derived in Heidinger (2010). The cloud typing algorithm produces clear, liquid water, supercooled liquid water, mixed phase, opaque ice, non-opaque ice, and multilayered ice pixel categories. The cloud types are independent of satellite and solar zenith angle, given that pixels are derived from IR data only with the clear sky correction. Since the goal of the UWCTC algorithm is to generate cloud top cooling values of vertically growing immature cumulus, the ice cloud types are considered to be pixels of interest (Sieglaff et al. 2011).

The main innovation behind the UWCTC algorithm is its use of a box averaging technique to track cloud objects. The identification of pixels of interest with the cloud-typing algorithm is the first step. The challenge to measuring vertical cloud top cooling trends is to track the cloud objects of interest through time (Sieglaff et al. 2011;

Walker et al. 2012). The box average method creates a 7x7 pixel box centered on the pixel of interest (Figure 2.1). The pixel of interest is then assigned the average 10.7 μm IR brightness temperature (T_b) inside that 7x7 pixel box to be denoted T_{b1} . The same calculation is done on the same pixel area for the next time frame to calculate T_{b2} . Thus, an unfiltered cloud top cooling rate is established using $T_{b2} - T_{b1}$. UWCTC then uses a larger 13x13 box and the smaller box with several cooling logic filters to remove false temporal cooling trends seen with box averaging caused by horizontal advection of clouds. The final products are forecast indications with units of measured vertical cooling ($\text{K } 15 \text{ min}^{-1}$) using IR only satellite trends (Sieglaff et al. 2011). The UWCTC data are obtained from the University of Wisconsin Space Science and Engineering Center.

SATCAST differs in several ways from the UWCTC algorithm. Instead of using box averaging, SATCAST uses mesoscale atmospheric motion vector calculations (Mecikalski and Bedka 2006). SATCAST uses a cloud mask generation technique that differs from day to night depending on the availability of visible satellite data. Daytime SATCAST products initially identify objects using cloud typing described in Bernedes et al. (2008). The discrimination between ‘immature’ and ‘mature’ cumulus is heavily dependent on observed visible texture (Walker et al. 2012). Use of the visible spectrum produces CI forecasts on a $\sim 1 \text{ km}$ resolution. Nighttime SATCAST products identify cloud types using methodology described in Jedlovec et al. (2008). Nighttime SATCAST has a $\sim 4 \text{ km}$ spatial resolution and utilizes a 3x3 kernel on the resolved cloud pixels through the 10.7 μm channel to classify clouds into cumulus and stratus types,

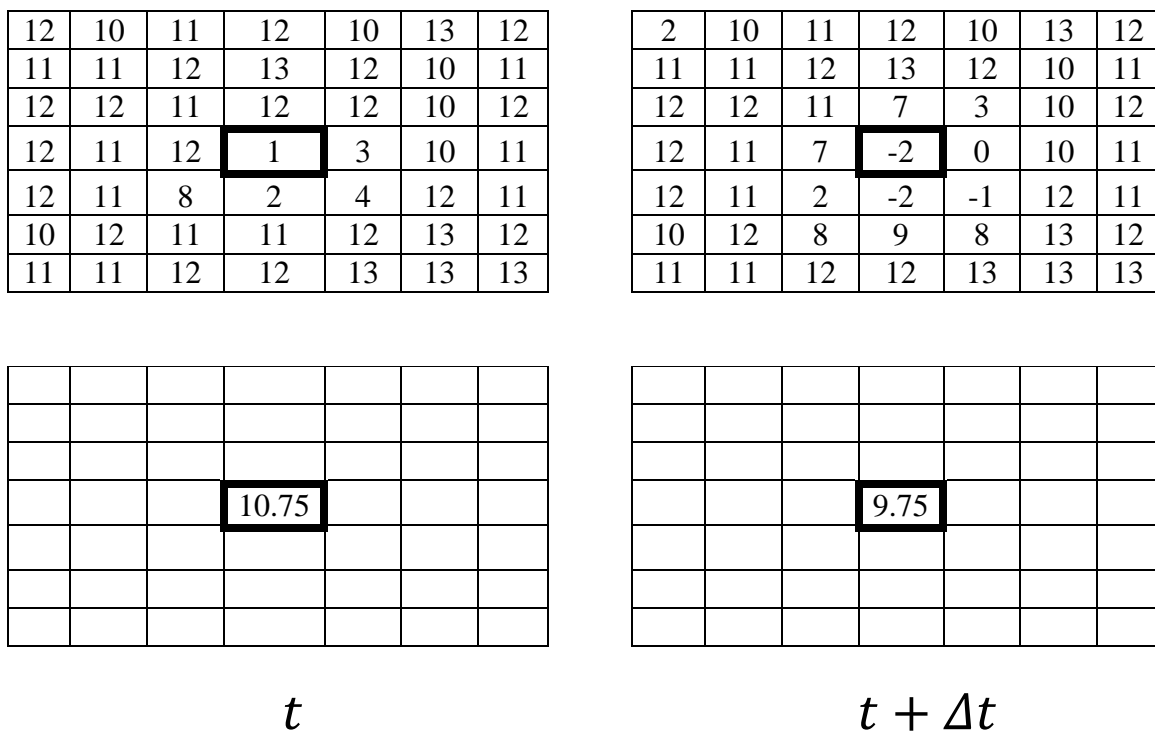


Figure 2.1. Schematic of boxed average tracking. 7x7 pixel box drawn around a pixel of interest outlined in bold with example brightness temperatures (T_b) in $^{\circ}\text{C}$ at time t (T_{b1}) and time $t + \Delta t$ (T_{b2}). (*Top*) Brightness temperature grid with a pixel to be box averaged in bold. The same box is superimposed on time $t + \Delta t$. (*Bottom*) Box averaged values shown for each time, with a box averaged CTC $T_{b1} - T_{b2} = -1^{\circ}\text{C}$ from t to $t + \Delta t$.

eliminating large uniform surfaces. SATCAST then eliminates clouds with a $10.7\text{ }\mu\text{m}$ $T_b < -20^\circ\text{C}$, assuming these objects are either mature cumulonimbus or thick upper level cirrus (Mecikalski and Jewett, personal communications: 2012). IR-based cloud masks are an improvement over previous versions, which were day/night dependent (Walker et al. 2012).

The main innovation of the SATCAST algorithm is through its object tracking methodology (Figure 2.2). SATCAST version 1 originally tracked individual pixels through atmospheric motion vectors (Mecikalski and Bedka 2006). Version 2 tracks identified cloud objects rather than individual pixels (Walker et al. 2012). The atmospheric motion vectors are used to extrapolate derived cumulus objects forward in time. Extrapolated objects are compared to the next time frame. A mask is then generated again for the next time to find cloud objects of interest. Extrapolated object IDs are passed to new convective objects that share the most spatial overlap with the projected area. Cloud object tracking is successfully achieved using the object overlap method (Walker et al. 2012). SATCAST calculates several interest fields, including $10.7\text{ }\mu\text{m}$ values/trends and $6.7\text{ }\mu\text{m}$ water vapor trends (Mecikalski et al. 2008). Version 2 of SATCAST used several interest fields to create a binary yes/no product. Through results of the 2012 Hazardous Weather Testbed (HWT), the SATCAST product was updated to version 2.2, which produces a strength of signal (SoS) product indicating a likelihood of CI, rather than the yes/no approach which at the time of this writing is still in an experimental phase (Mecikalski, personal communication: 2013). The SoS value is calculated through logarithmic regression using interest fields from previous versions of

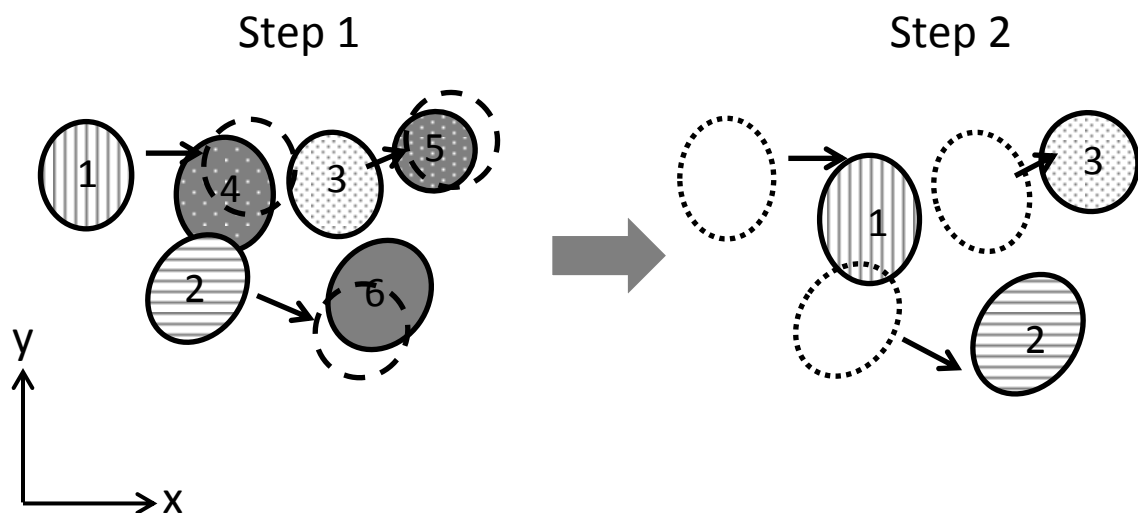


Figure 2.2. Schematic of SATCAST object tracking methodology. Objects 1, 2, and 3 are resolved at time t . Objects 4, 5 and 6 are resolved at time $t + \Delta t$. Arrows represent horizontal atmospheric motion vectors. Dashed lines represent extrapolated forward using horizontal atmospheric motion vectors. Step 2 has objects reassigned from step 1 based on the position estimates, where dotted lines represent previous object location.

SATCAST (Mecikalski and Jewett, personal communications: 2013). More recent versions of SATCAST (version 2.3, after April 2013) are beginning to incorporate pre-convective environmental variables in the logarithmic regression equation for SoS (Mecikalski, personal communication: 2013). The SATCAST data are obtained from the University of Alabama at Huntsville.

b. Indication Validation

Several approaches have been taken to validate CI indication datasets. Mecikalski and Bedka (2006) used a composite radar dataset with a 35 dBZ criterion. They found using composite reflectivity and a detailed visual comparison with the products, the binary yes/no version of SATCAST displayed a ~65% probability of detection (POD). Walker et al. (2012) also used a subjective SATCAST radar comparison methodology, only using radar data within a ~75 km range to capture all low-topped convection and mitigate degradation in resolution as a result of beam spreading. Mecikalski et al. (2008) expanded into radar comparison objectively using WSR-88D data from Hytop, Alabama, and Topeka, Kansas, merged into a CAPPI 1 km resolution grid. Points above ~1-1.5 km below the sounding-derived freezing level were removed to account for possible bright band effects. Pixel objects are advected forward using derived mesoscale atmospheric motion vectors to objectively associate forecasts with radar values. Mecikalski et al. (2008) used the objective approach to explore the value of eight IR-based interest fields for application in the SATCAST algorithm. However, when using the optimal combination of interest fields, false alarms were abundant in the indication dataset (Mecikalski et al. 2008). Walker et al. (2012) also noted a modest false alarm ratio even

after filtering out missed detections resulting from an inability to track cloud objects across low temporal resolution datasets. An environmental NWP model dataset would therefore be helpful in reducing the over prediction issue in SATCAST (Mecikalski et al. 2008; Walker et al. 2012).

Although most of the literature uses validation through the 35 dBZ criteria, several validation studies have been completed through lightning verification (Sieglaff et al. 2011; Mecikalski et al. 2013). Cloud objects with associated cloud-to-ground lightning were manually tracked and compared to UWCTC indications in Sieglaff et al. (2011). UWCTC was found to have a relatively small FAR, around 34.8% for the overall validation domain. The subjective lightning validation lacked the ability to quantize convective events and therefore was unable to determine amount of missed detections using only lightning data (Hartung et al. 2013). A methodology was later developed to track satellite cloud objects (groups of convective cloud GOES pixels) coupled to lightning data to quantize number of CI events (Sieglaff et al. 2013). The objective methodology produced relatively similar results to Sieglaff et al. (2011), with UWCTC POD about 22% for weak echoes (>35 dBZ) and increased to 62% for strong echoes (>60 dBZ). The small POD and FAR values were expected given UWCTC's conservative filtering process (Sieglaff et al. 2011; Hartung et al. 2013). Stronger cloud top cooling rates had higher validation skill scores (Hartung et al. 2013). While the objective validation methodology in Sieglaff et al. (2013) allows for the collection of large samples for validation datasets, low temporal resolution (> 20 min) in satellite data causes large tracking skill depreciation. Thus larger temporal resolution improves object

tracking and identification (Sieglaff et al. 2013; Hartung et al. 2013). Therefore, there is a need for an event identification dataset with high temporal resolution. Ideas from all previous validation studies are employed in this study to create an optimal form of objective validation for environmental analysis.

c. Environmental Analysis

There are several variables that can act to hinder convective development (i.e. Johns and Doswell 1992; Weckworth and Parsons 2006; Roberts et al. 2012; among many others) and possibly effect cloud top cooling based algorithms (Mecikalski and Bedka 2006; Mecikalski et al. 2008; Sieglaff et al. 2011; Walker et al. 2012). Mecikalski and Bedka (2006) and Sieglaff et al. (2011) speculated that SATCAST and UWCTC would have difficulty generating forecasts in regions containing high storm motion (SM) as assumptions inherent in box averaging and atmospheric motion vector-based spatial overlap begin to fail. Mecikalski et al. (2008) suggests that satellite-based nowcasting may also be problematic on convective objects of interest that have not yet reached their LFC and have been forced to ascend below that level. Walker et al. (2012) speculated that performance issues under a capping inversion occur when cloud objects are initially identified in the ~1 km visible spectrum and grow to the infrared horizontally on the next time frame, resulting in large SoS values when SATCAST compares a warm ground-based pixel to a cool cloud-based value. Mueller et al. (1993) found areas of localized convergence could help determine where a parcel is more likely to reach its LFC and less likely to be affected by convective inhibition (CIN).

Mueller et al. (2003) uses the convergence line detection (COLIDE; Roberts et al. 1999) algorithm to resolve areas of surface or elevated boundaries using radar and satellite data in the National Center for Atmospheric Research's auto-nowcast system (ANC). ANC also employs a fuzzy logic system to combine various amounts of convective available potential energy (CAPE), CIN and other boundary parameters to produce a statistical model based convective forecast. The ANC statistical model was created using several hundred case studies (Mueller et al. 2003). Incorporation of these types of variables through a similar stochastic system should improve satellite-based CI algorithm performance (Mecikalski et al. 2008).

3. Methodology

Several CTC and satellite interest field based studies are performed on the US Central Plains (Roberts and Rutledge 2003; Mecikalski and Bedka 2006; Sieglaff et al. 2011; Walker et al 2012). The present study uses a similar region over the Great Plains (Figure 3.1) to examine satellite-based CI forecasts. The area was chosen because of the radar coverage and availability of in situ measurements through surface observations and radiosonde. The measurement capability in the Central Plains improves model performance with correct data assimilation (Devenyi et al. 2007; Kleist et al. 2009). An environmental study may significantly differ based on location (i.e. the tropics compared to desert regions), so a study on a well measured region is ideal for initial insight on how CTC algorithms behave in changing conditions. Data collection is dependent on algorithm and satellite dependability as well as radar availability. For the environmental analysis, data are used from 20 July 2012 to 17 August 2012 (July study period) and from 17 April 2013 to 17 May 2013 (April study period).

A large-scale mosaic radar dataset is needed to validate satellite-based convective algorithms. The national mosaic quantitative precipitation estimation radar dataset (NMQ) offers multi-radar multi-sensor coverage over the study region. In addition, a NWP model is needed to resolve a full three-dimensional picture of the environment containing the observed CI indications. The rapid refresh model (RAP) is a high-resolution operational model available at the time of this writing. NMQ and RAP are described in detail in the following sections.

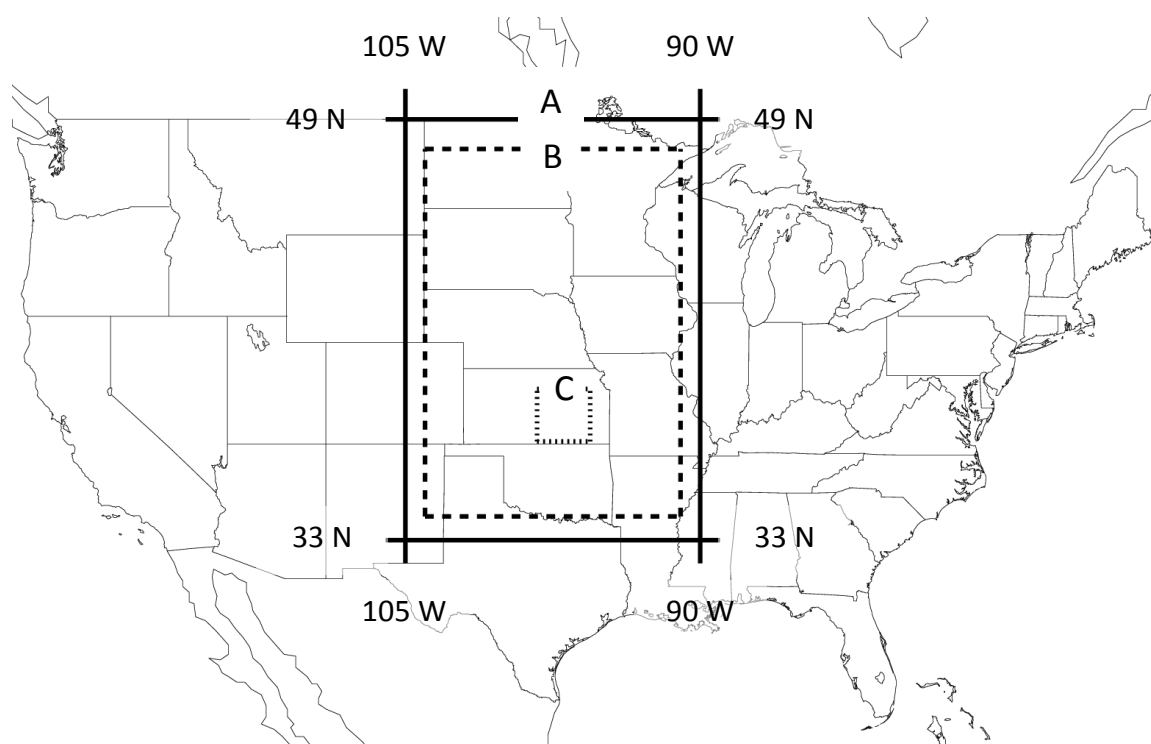


Figure 3.1. A) Total area of study (solid outline). B) Parallax and track-corrected area of study (dashed outline ranging from 91 – 104 W, 34 – 48 N). C) CI truth cluster validation domain (dotted outline).

a. National Mosaic Quantitative Precipitation Estimation

NMQ is used to improve precipitation estimation through multiple radars and sensors. NMQ uses level two radar data from more than 140 WSR-88D radars and 31 Canadian C-band radars to create a CAPPI radar dataset (Zhang et al. 2011). NMQ is created by first quality controlling raw radar data to remove effects such as ground clutter, sun spikes, wind farms and biological targets. Products are also filtered where significant beam blockage occurs, as is common near mountains. In addition, the NMQ dataset filters out bright banding features caused by melting ice particles by deriving vertical profiles of reflectivity and analyzing their uniformity (Zhang and Qi 2010).

After the individual data are quality controlled, they are merged into a final mosaic product. Merging is accomplished using a weighted function described in Zhang et al. (2011). The weighted function includes distance the scanned volume is from the radar, height above sea level and time at which the scan was taken. A weighted approach was chosen over a nearest neighbor mapping approach to mitigate discontinuities between radar datasets. The final product is a merged radar reflectivity dataset that covers a large section of North America (Zhang et al. 2011). At the time of this writing, the grid reaches from 60° to 130°W and 20° to 50°N. NMQ is produced on a cylindrical equidistant map with a 0.01°x0.01° resolution (Zhang and Qi 2010; Zhang et al. 2011). NMQ data are obtained from the National Severe Storms Laboratory (NSSL) in Norman, Oklahoma.

b. Rapid Refresh

The RAP NWP model recently replaced the rapid update cycle model (RUC). RAP has been operational as of 1 May 2012 (see rapidrefresh.noaa.gov). RAP is an hourly-updating model with 50 vertical levels and a Lambert conformal ~13 km resolution grid covering North America. The standard pressure level data produced and archived span from 1000 – 100 hPa, with a vertical resolution of 25 hPa. Several improvements have been provided with the implementation of RAP. RAP has a larger domain than RUC and uses the more advanced nonhydrostatic grid point statistical interpolation (GSI) over the older three-dimensional variational (3DVAR) system (Kleist et al. 2009). Previous studies on the 3DVAR system have shown that several parameters such as surface-based CAPE have a tendency to be too unstable. Overestimation of instability is consistent with the RUC's bias towards large moisture in the low levels (Coniglio 2012). Unstable trends are mitigated with the GSI system, as well as integration of real time radar and satellite data into the RAP (Devenyi, et al. 2007). While RUC CAPE relied upon a hydrostatic assumption in its formulation from the first law of thermodynamics, RAP CAPE does not require this formulation. RAP CAPE also uses a virtual temperature correction. The overall effect is smaller CAPE in very unstable regions, however larger CAPE is found in regions with modest instability with rapid decreases in mixing ratio with height (see http://ruc.noaa.gov/rr/RAP_var_diagnosis.html#CAPE-def). Despite these weaknesses, RAP should offer a sufficient look at pre-convective environments. RAP data are obtained from the National Climatic Data Center in Asheville, NC.

c. Defining Convective Initiation

Validation first requires a definition for convection. A radar-based definition appropriately characterizes CI over the Great Plains and has been used in past studies (Roberts and Rutledge 2003; Mecikalski and Bedka 2006; Walker et al. 2012). For this study, CI is defined as the first occurrence of a radar return of 35 dBZ. Since our goal is to determine points of CI at any height, NMQ composite data are used. Composite values are calculated by finding the maximum reflectivity in a column at a constant grid point. Bright band issues seen in Mecikalski et al. (2008) are removed by NMQ's initial filtering processes, so all heights in the NMQ dataset can be used. The composite dataset requires radar return segmentation and tracking for objective validation.

The Warning Decision Support Services-Integrated Information (WDSS-II) package offers a suite of algorithms for severe weather analysis and forecasting. WDSS-II was developed by the NSSL and the Cooperative Institute for Mesoscale Meteorological Studies at the University of Oklahoma (Lakshmanan et al. 2007). Among the suite of products lies the w2segmotion algorithm, which is utilized for this study. W2segmotion clusters radar data based on a combined k-means and enhanced watershed technique (Lakshmanan and Smith 2009). While k-means alone can serve to segment NMQ data, the w2segmotion algorithm k-means is used solely for the purpose of quantization.

The k-means technique is a process which assigns a grid into a discrete hierarchy of quanta. K-means clustering is achieved by iteratively minimizing a cost function for

every pixel above a specified threshold (e.g. 35 dBZ) that accounts for Euclidian distance in textural space and contiguity (similarity to neighboring pixels) with possible candidate clusters (Lakshmanan et al. 2002; Lakshmanan et al. 2003). Textural Euclidian distance is determined using the vectors of mean, variance and coefficient of variance for the pixel and cluster. Pixels are reassigned to new candidate clusters iteratively until the process is stable. Stable refers to the state at which no pixels are reassigned to new clusters. The enhanced watershed technique (Lakshmanan et al. 2009) is then used on the quantized clusters in k-quanta space to grow segmented clusters to a set minimum scale threshold (Lakshmanan and Smith 2009). The k-quanta dataset is considered as terrain where water begins at the global maximum and is removed by discrete increments. Clusters are then grown on these discrete increments from their maxima until a minimum scale threshold (saliency) is reached or surpassed. When the scale threshold is reached, growth by discrete k-increments is ceased and the cluster is identified. If a cluster is grown that does not reach the minimum scale before it drops below a minimum threshold, w2segmotion has the ability to combine nearby clusters to meet the saliency threshold (Lakshmanan and Smith 2009). If the clusters cannot be combined or grown further and do not reach the minimum scale threshold, they are not considered as identified peaks. The segmentation method has the advantage of not being dependent on a threshold value only (e.g. ≥ 35 dBZ). The end result is an efficient cluster segmentation that allows for storm identification on a grid (Lakshmanan and Smith 2009).

The thunderstorm observations by radar tracking algorithm (ThOR) was developed by the University of Nebraska to identify and track convective cells (Houston

et al. 2013). ThOR uses clusters identified in the w2segmotion algorithm and storm motion estimates to develop tracks for identified cells through time. Tracking is achieved by considering all possible tracks within an acceptable error radius from a first guess location (Figure 3.2). First guess locations are based solely on North American Regional Reanalysis (NARR) storm motion for the first identified clusters in tracks. The first guess is weighted for observed track motions as tracks grow in length (Houston et al. 2013, updraft.unl.edu/thor/wiki). Explained another way, when a cluster is initially identified, future motion estimates are created using only resolved storm motion. When a track is built, ThOR can use previously known locations to improve the guess on where a cell is moving. Adaptive first guess locations are advantageous compared to a storm motion only based approach particularly when tracking objects not constrained to the 0-6 km mean wind (e.g. supercells). Pixel clusters are tracked until they drop below a specific spatial threshold or move outside regions of interest (see updraft.unl.edu/thor/wiki). ThOR then considers all possible tracks found and assigns the track containing the least amount of mean error (average distance from the first guess locations to the actual cluster locations). The initial resolved cluster of a track is considered the CI cluster.

The combination of WDSS-II and ThOR will objectively identify CI using NMQ data. WDSS-II requires a minimum spatial threshold to define a convective system. WDSS-II also performs smoothing on a data grid to remove spurious peaks caused by data noise. The spatial and smoothing scales are determined using a hand validation technique.

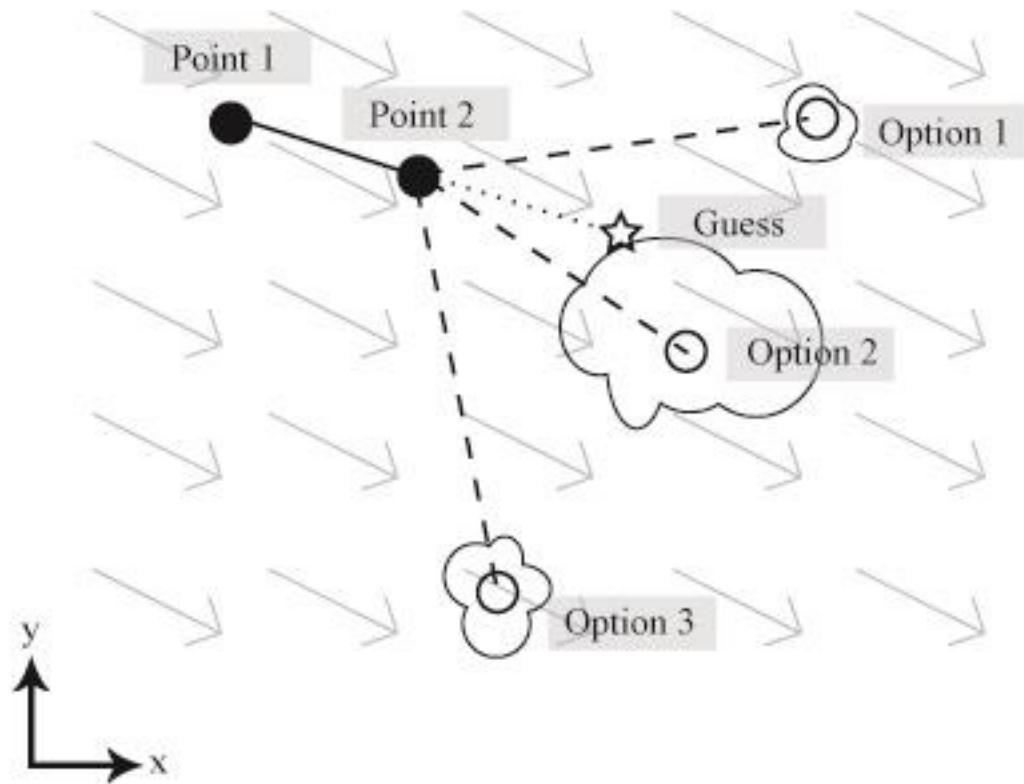


Figure 3.2. Schematic of ThOR tracking. Light gray arrows represent NARR storm motion data. Solid shapes represent current cluster locations. The first two points of track are represented with black dots, in addition to three possible tracks shown as dashed lines. The ThOR track guess shown is based on NARR storm motion and previous direction travelled from point 1 to point 2. ThOR chooses between the optional tracks based on the distances (error) from the guess centroid to the optional observed centroids (options 1, 2 and 3). In this figure, ThOR would test all three possible tracks, and choose the track with the least overall mean error (adapted from updraft.unl.edu/thor/wiki).

d. Hand Validation Technique

A small dataset on 25 July 2012 (Figure 3.1c) is examined to determine which WDSS-II settings would work best for an objective scheme. The domain is used from 2000 UTC 25 July 2012 to 0000 UTC 26 July 2012 due to the abundance of CI events. The validation process begins with forecasters at the University of Nebraska identifying events they believed as CI in this time frame, which are referred to as “truth” clusters. Forecasters identified “truth” clusters using a spatial scale they believed accurately represented a CI event. The objective clusters are then validated using the given “truth” clusters.

The first variable tested is the smoothing filter WDSS-II uses to remove spurious peaks and noise. The default WDSS-II setting uses a 50th percentile filter with a box half-size of 5. Half-size refers to the spatial number of pixels around which the smoothing filter is applied. For example, a half size of 5 refers to an 11x11 pixel box which is considered in the smoothing process (see www.wdssii.org). Larger half sizes can remove information that may be important to identifying CI events. Lowering half-sizes comes with the consequence of increased spurious peaks and noise values, so half-sizes of 2, 3, 4 and 5 are tested. A minimum scale is also tested with different half-sizes. Minimum scale saliency values of 20, 30, and 40 pixels are examined with all three half-sizes.

Four different half-sizes with three different scales are input into ThOR to determine the location of a CI event, which is then compared to the “truth” dataset. A

ThOR CI cluster with (without) a corresponding “truth” cluster within 10 min is considered a positive (false) indication. A “truth” cluster without a corresponding ThOR cluster is considered a missed detection. The probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI) are determined for the WDSS-II/ThOR objective validation techniques are calculated as follows:

$$POD = \frac{a}{a + b}$$

$$FAR = \frac{c}{c + d}$$

$$CSI = \frac{d}{d + b + c}$$

where a is the number of detected CI “truth” clusters, b is number of missed CI “truth” clusters, c is number of false WDSS-II/ThOR CI detections and d is number of positive WDSS-II/ThOR CI detections. Optimizing WDSS-II settings will allow for robust identification of CI events which are used in objective validation.

e. Algorithm Validation

SATCAST, UWCTC, NMQ and RAP are produced with ~1 km, ~4km, ~1 km and ~13 km resolutions respectively on differing projections. Since SATCAST is produced on a ~1 km resolution flat plan projection, this grid was chosen as the base grid to use. All other datasets are remapped to the SATCAST grid using a nearest neighbor approach. Remapping is done by assigning each new grid pixel the value that is closest to the grid point to be remapped (Figure 3.3). Any SATCAST-based grid pixel that is

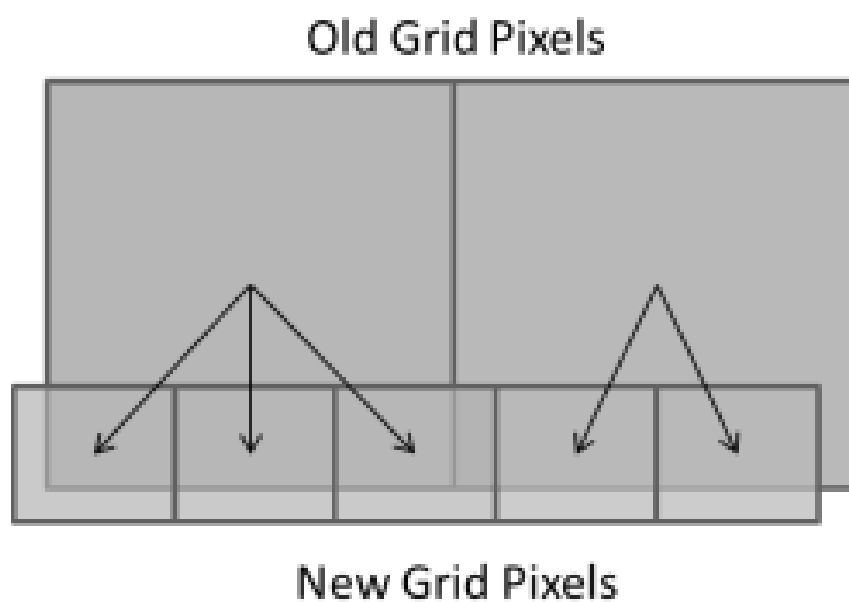


Figure 3.3. Schematic of nearest neighbor grid remapping. Values of the old grid (large squares) are reassigned to the closest new grid pixels (small squares).

equidistant to two or more pixels is reassigned the maximum value of the set (only valid for NMQ remapping). Once all data are transformed to the same grid, the validation process can begin.

SATCAST and UWCTC are produced on a grid of individual pixels, however indications are considered as clusters. Clustering is done by grouping any neighboring pixels and considering them as a single indication. When either SATCAST or UWCTC produces a forecast indication, the indication is compared to the radar data to see if the forecast is valid. An indication that corresponds with the resolved CI cluster location within one hour is considered a positive detection. If an indication occurs with no corresponding CI event, then that indication is considered false. A UWCTC product is considered a CI indication if the cluster contains a value $< -4 \text{ K } 15 \text{ min}^{-1}$.

Since indications can be created up to one hour before they are validated, vertical development that created the indication is expected to advect out of its original location. Thus CTC may be witnessed well upstream of the first 35 dBZ reflectivity cluster. If left unchecked, indications with large storm motions would be considered false with an objective approach because their respective CI points are well outside of the considered stationary validation area. To correct for storm motion, the validation process includes use of RAP derived 0-6 km SM vectors to advect indication validation areas through time. When an indication is first detected, the storm motion vector for that indication is used to advect the indication's validation area forward appropriately for each radar time step (every 5 min). The RAP 0-6 km storm motion vectors are then considered for the

new position of the indication's validation area at each time step to appropriately advect the area with the flow.

Three environments are presented to illustrate the need for advected validation (Figure 3.4). In environment A, given that CI occurred within the validation areas, the indication would be considered positive for both advected and non-advected validation. Environment B illustrates a situation in which an indication would be considered false in non-advected validation given that the cloud object moved away from its originally detected region. Situation C is an environment where non-advected validation would produce a positive indication for the incorrect cloud object. Modifying for advection allows for products to be correctly validated in high storm motions and also prevents false positives that occur when CI from another system advects into the area of the original indication.

While the nearest neighbor approach does not change the values of the data, it does change the intended position of the forecast (with errors $< \sim 1$ km). Spatial correlation problems also exist when comparing satellite brightness temperature values to reflectivity. A spatial displacement of satellite-based indications can also exist due to parallax. Parallax is resolved using a similar method to Sieglaff et al. (2011) where indications are assumed to be at a height of 7 km. When correcting for parallax, $\sim 0.5^\circ$ are lost on all sides of the domain. In order to correctly identify CI points near the boundary of the domain using ThOR, another $\sim 0.5^\circ$ is used only to track objects and not for validation. The resulting domain ranges from $91^\circ - 104^\circ$ W, $34^\circ - 48^\circ$ N (Figure 3.1). All other spatial correlation issues are resolved as part of the hand analysis. Positive

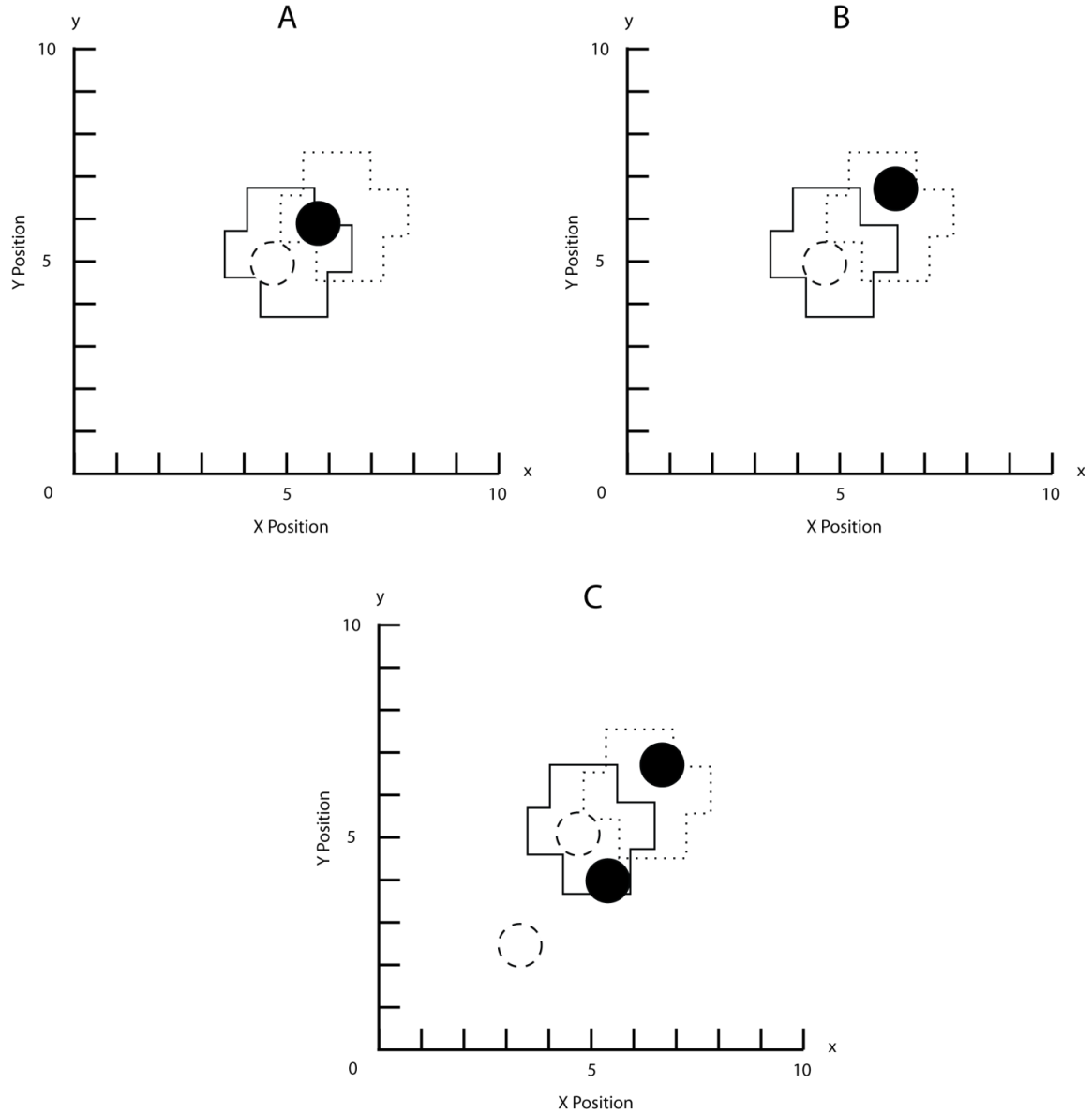


Figure 3.4. Three different environments (notated A, B and C) of validation areas (solid polygons) and a validation area that is advected forward (dotted polygons). Solid circles indicate locations of CI, and dashed circles indicate original cloud object location.

indications were found over a period of four hours on 21 July 2012 across the entire domain. The positive centroids are advected and compared to the centroids of their respective CI detection. The centroid distance was compared to the average spatial coverage of both SATCAST and UWCTC indications. It was found that SATCAST pixels were on average within 2.96 km in spatial coverage, however CI detections were as far away as 4.58 km. UWCTC on average covered a larger area at ~5.68 km, with CI detections 5.86 km away. It is possible that this spatial error is caused by parallax issues, or issues in the remapping data process. To compensate for the maximum possible spatial error, indications within ~3 pixels (~3 km) of a CI detection are considered positive forecasts.

A separate error that needs to be addressed in validation is the problem of multiple indications on a single storm changing the number of positive and false indications. If one system produces multiple indications, a system with several indications either positive or false would be favored in an environmental analysis over systems that may only produce few indications. Data from GOES rapid scans would be favored in study period collections unless indications are tracked and coupled. ThOR is used to track indications from both UWCTC and SATCAST. An indication track is considered as either one positive or one false indication. The track is positive if it contains an indication deemed to successfully forecast CI. Thus from this point on, a positive or false indication refers to a group of tracked indications. The averages of environmental variables for the entire track are recorded for the single positive or false indication.

Tracking satellite-based CI indications requires some changes from the radar tracking approach. Radar-based detections that are tracked are above the set spatial and reflectivity thresholds and thus normally propagate with the mean 0-6 km wind. UWCTC indications are only given if substantial vertical cooling is identified, and usually propagate similar to radar indications along the mean 0-6 km wind. SATCAST variables are produced on all clouds determined ‘pre-convective’ (Walker et al. 2012), and thus misidentified objects (edges of cirrus) may not be influenced by the lower tropospheric wind at all. While ThOR tracking with reasonable temporal scales (7-15 min) was able to handle this issue through track weighted first guess values, large temporal gaps on full disc scans created tracks deemed unreasonable with closer examination. Thus SATCAST tracks were not allowed to make temporal jumps ≥ 30 min, and new tracks were created after each full disc scan. SATCAST tracking with ThOR also had issues with large numbers of candidate tracks being generated due to the large number of indications. Since ThOR considers all possible tracks for each indication, data rich areas such as cumulus fields generated large numbers of candidate tracks which became unreasonable for computation. Excessive candidate tracks are mitigated by reducing the normal search radius of consideration and discarding tracks that would cause unreasonable mean error. Since the search radius around a first guess is a function of time (see updraft.unl.edu), the issue of too many candidate tracks will be mitigated with datasets containing larger temporal resolution. Both corrections with SATCAST tracking should be noted by the reader, however both should not be necessary when this study is repeated on higher temporal resolution data with the launch of GOES-R.

f. Environmental Analysis

The remapped RAP data can now be used with validated SATCAST and UWCTC indications to help explain why some indications are positive while others are false. Since products are evaluated in clusters, environments can vary for different regions of SATCAST and UWCTC indications. Environmental variables for an indication are considered as the average of the indication clusters' spatial and temporal domain. This study focuses on variables that would affect CI and its detection by satellite (Table 3.1).

All algorithms use brightness temperature differences to derive areas of upward vertical motion (Mecikalski and Bedka 2006; Sieglaff et al. 2011; Mecikalski et al. 2013). When diagnosing areas prone to convective motion, parcel theory based convective variables are useful forecasting tools. CAPE is a common variable when attempting to quantify the amount of instability to be released vertically in the atmosphere (Williams and Renno 1993). An environment containing larger CAPE values is more conducive of faster transitions from immature cumulus to mature cumulonimbus (Mecikalski et al. 2013). CIN is another parcel theory based term for the amount of negative buoyant energy found in the lower troposphere. A deep layer of CIN can prevent convection from reaching set thresholds for objective validation. Definitions for parcel-based quantities such as CAPE, CIN, level of free convection (LFC), equilibrium level (EL) and lifted condensation level (LCL) are dependent on the lifted parcel level (LPL). A moist unstable approximation is used to account for both surface-based and elevated convection. The maximum pseudo-equivalent potential temperature (θ_e) found in the lowest 300 hPa is assumed to be the LPL.

Table 3.1. List of variables used with respective abbreviations and descriptions.

Variable	Abbr.	Description
Convective available potential energy	CAPE	Total integrated positive potential energy calculated from the LFC to the EL
Convective Condensation Level	CCL	Level at which the surface mixing ratio meets the temperature profile
Convective inhibition	CIN	Total integrated negative (downward) potential energy from the lifted parcel level to the LFC
Equilibrium level	EL	Level above LFC where a lifted parcel becomes cooler than the surrounding environment
Lifted condensation level	LCL	Level at which a lifted parcel saturates
Level of free convection	LFC	Level where an adiabatically lifted parcel becomes warmer than the surrounding environment
Lifted Parcel Level	LPL	Level of maximum θ_e from the surface pressure to the surface pressure-300 hPa
LPL Divergence	LPLD	$\nabla \cdot \vec{V}$ composite value calculated at the LPL
Lapse rates	LR	Change in temperature with respect to height calculated at several levels (0-3 km, 700-500 hPa)
Mean Wind Differential	MWD	Change in wind speed and direction with respect to height calculated from several levels (0-6 km)
Normalized CAPE	NCAPE	CAPE divided by the depth of the unstable layer
Normalized CIN	NCIN	CIN divided by the depth of the stable layer
Layer Relative Humidity	RH	Average relative humidity calculated from the LCL to the LFC
Storm motion	SM	Averaged layer of mean wind from 0-6 km assumed to be the characteristic storm motion
Convective Environmental Difference	$T_c - T_e$	Approximation of cumulus field formation potential, the difference of the convective and environmental temperatures
EL-LFC	Z_{EL-LFC}	Distance in m from the LFC to the EL
LFC-LCL	$Z_{LFC-LCL}$	Distance in m from the LCL to the LFC

Positive and false indications are compared to the layers which a convective parcel will have to traverse including the LCL to LFC ($Z_{LFC-LCL}$) and LFC to EL (Z_{EL-LFC}). This writing is focused on $Z_{LFC-LCL}$ which is a characteristic of the depth of the stable layer visible by satellite. The focus on $Z_{LFC-LCL}$ is twofold. Walker et al. (2012) suggested that rapid growth seen beneath a midlevel capping inversion can fool the algorithm into thinking pre-convective clouds will mature. If this is the case over the Central Plains, false indications should be witnessed with large $Z_{LFC-LCL}$ values. Z_{EL-LFC} is examined to monitor the convective cooling depth throughout the unstable layer. Areas with larger Z_{EL-LFC} have greater adiabatic temperature change throughout parcel excursion and should be easier for brightness temperature differencing algorithms to correctly resolve.

To supplement the depths of convective and stable layers, it is helpful to include normalized values of instability and stability into our analysis. CAPE normalized throughout the depth of the unstable layer (NCAPE) can parameterize a parcel's susceptibility to entrainment during ascent. Large NCAPE values suggest the environment is conducive of more explosive development (less entrainment) with larger buoyant accelerations in a layer. More explosive updrafts yield larger brightness temperature changes due to faster adiabatic cooling through ascent. For the same reasons explained for $Z_{LFC-LCL}$, normalized CIN (NCIN) may be important in reference to satellite CTC detection below the capping inversion. Lower NCIN values (more negative) are a characteristic of deeper stable layers, which if a parcel is allowed to cool through a deeper layer may be more prone to false identification (Walker et al. 2012).

The effect of vertical shear, or mean wind differential (MWD), on CI indications is not yet understood. While surface to midlevel shear is commonly used to discriminate between ordinary cell convection and supercell complexes (Weisman and Klemp 1984), the effects of highly sheared environments have not yet been explored on satellite-based CI forecasts. The 0-6 km MWD magnitude is examined in this study.

Lapse rate (LR) values are measured two ways: 0-3 km (LOWLR) and from the 700-500 hPa pressure levels (MIDLR). The LOWLR and MIDLR are analyzed for any specific patterns that may exist in a layer with positive and false indications. Layer-based lapse rates can break CAPE into its individual components to determine if a specific layer is important to performance of satellite-based CI detection.

Layer relative humidity (RH) is the mean relative humidity from the LCL to the LFC. RH has not yet been examined with respect to CTC-based algorithms. Values of RH near 1 are indicative of environments where the LCL is near the LFC large low level moisture limiting the impact of dry air entrainment through the stable layer. Two additional variables analyzed include the convective condensation level (CCL) and the difference of environmental temperature from the convective temperature ($T_c - T_e$). Larger values of $T_c - T_e$ can be indicative of dry, stable areas while smaller values indicate areas more prone to create satellite-detectable cumulus fields with relatively unstable low-level layers. SM is also included in this study.

Divergence is the final variable examined in this study. Resolved divergence at the LPL (LPLD) can offer some indication of the presence of boundaries where there is

mass conservation related low level convective forcing. As mentioned in Mueller et al. (1993) and discussed by Banacos and Schultz (2005), use of a surface-based divergence would not be sufficient to nowcast deep convection alone. Therefore, this study uses convergence along the most unstable parcel height.

Temporal resolution of the RAP is an important limitation in convective time scale studies. Convection can occur on 0-1 hour time scales, while the RAP has a temporal resolution of only one hour. Thus, convective events that occur on sub-hour time scales change the environmental parameters in such a way that cannot be detected by RAP analysis. Since convection-contaminated environments cannot yet be resolved by operational NWP models, areas near ongoing convection are removed from datasets used for statistical analysis. This study uses a 50 km radius mask ($\sim 7850 \text{ km}^2$) around all WDSS-II resolved convective clusters to be considered contaminated (and therefore unused) areas. The 50 km radius ensures that convective events cannot advect into areas that are being considered for pre-convective environmental analysis on sub-radar ($< 5 \text{ min}$) time scales. The 50 km radius mask does not affect calculation of validation statistics, as RAP variables are not required to determine algorithm performance.

g. Data Fusion

The final step of this study is to take the environmental variable analysis of positive and false indications and use it on future products to improve satellite-only CI forecasting algorithms. Once the original CTC algorithms are validated and the environmental variables are found for each indication, three different statistical

approaches are used. The first is a PCA. PCA is a simple statistical way to reduce the dimensionality of data (Hotelling 1933). PCA is accomplished by finding the eigenvalues and eigenvectors of the covariance matrix for the environmental variables. The first vectors give us the axes of greatest variation in our data, and allow us to organize an 18-dimensional dataset into two or three common factors (principal components). This method was previously used to analyze the variation in several satellite-based interest fields in Mecikalski et al. (2008). The top two principal components are resolved and plotted with respect to positive and false indications.

The second statistical approach is a calculation of actual statistical differences between environmental variables for positive and false indications using an analysis of variance (ANOVA) and analysis of covariance (ANCOVA) approach (Cochran 1957). ANOVA tests evaluate statistically significant differences of means assuming normal distributions. Use of ANCOVA tests for statistically significant differences can account for variation present within treatments as a result of linear correlation with other environmental variables. While ANCOVA is a sharper test than ANOVA, reducing statistical significance as a result of between-treatment correlations can be misleading when comparing two very similar datasets such as $Z_{LFC-LCL}$ and CIN. Since the $Z_{LFC-LCL}$ is related to CIN, it is to be expected that correlation exists between the two, so ANCOVA tests would only show one or the other as significantly different. The ANCOVA approach is flawed when testing treatments that are directly involved in the calculation of other treatment variables. For example, a correlation is expected between $Z_{LFC-LCL}$ and LFC because LFC is directly involved in the calculation of $Z_{LFC-LCL}$. Thus

variables such as LFC and LCL are removed from the ANCOVA analysis. Statistically significant differences are applied to a probability-based data fusion methodology. The data fusion methodology chosen for this work is a quadratic discriminant analysis (QDA).

It should be taken into consideration that ANOVA, ANCOVA and QDA all rely on the assumption of multivariate normality. In the atmosphere, most data are not always normally distributed. To account for potential non-normal distributions, transformations on particular variables are performed (Table 3.2). Results from means and distributions are reported in an untransformed form. When ANOVA, ANCOVA and QDA are performed, the data input are transformed appropriately such that the assumptions for multivariate normality do not fail. Data are also transformed and tested through two sample rank testing in the form of Mann-Whitney (MW) based Z statistics (Mann and Whitney 1947). MW is not sensitive to non-normality in datasets given that it is an examination of ranked data. If a statistically significant difference is not found for an environmental variable, it is removed from QDA. ANOVA, ANCOVA and MW results are compared to determine which set of variables produce the most helpful QDA.

Discriminant analysis was initially proposed as a means of classifying data into two separate categories based on another variable (Fisher 1936). QDA is a statistical method of classifying data into groups based on outside variables that does not require the assumption of equal covariance matrices between positive and false indications. In our particular case, QDA is used to classify indications into positive and false groups based on environmental data found with RAP analysis. QDA works by directly

Table 3.2. Transformations done to normalize convective variable distributions. These transformations are used in the statistical tests and results are reported in the original forms.

Variable	Transformation
Convective available potential energy	\sqrt{x}, \sqrt{x}
Normalized CAPE	$x^{\frac{1}{8}}, x^{\frac{1}{6}}$
Convective inhibition	$-1 * (x)^{\frac{1}{10}}, -1 * (x)^{\frac{1}{5}}$
Normalized CIN	$-1 * (x)^{\frac{1}{3}}, -1 * (x)^{\frac{1}{9}}$
Lifted Parcel Level	\sqrt{x}
Lifted condensation level	\sqrt{x}, x
Level of free convection	\sqrt{x}, \sqrt{x}
Equilibrium level	$x^{1.5}, x^{1.5}$
Convective Condensation Level	\sqrt{x}
Convective Environmental Difference	x
Storm motion	x
$Z_{LFC-LCL}$	x, x
Z_{EL-LFC}	x, x
Mean Wind Differential	x
Lapse rates	x
Layer Relative Humidity	x, x
LPL Divergence	x

comparing the data distributions of environmental variables for two groups and assigning a new indication to a group with the highest probable outcome (Figure 3.5). If the two environmental variable distribution variances are equal, then a QDA is the same approach as a linear discriminant analysis (LDA) where only one group probability shift exists. Mathematically, a QDA looks like this:

$$\begin{aligned} \text{Allocate } x_0 \text{ to group 1 if: } & -\frac{1}{2}x_0'(S_1^{-1} - S_2^{-1})x_0 + (\bar{x}_1'S_1^{-1} - \bar{x}_2'S_2^{-1})x_0 - k \\ & \geq \ln \left[\left(\frac{c(1|2)}{c(2|1)} \right) \left(\frac{p_1}{p_2} \right) \right] \\ \text{where } k = & \frac{1}{2} \ln \left(\frac{|S_1|}{|S_2|} \right) + \frac{1}{2} (\bar{x}_1'S_1^{-1}\bar{x}_1 - \bar{x}_2'S_2^{-1}\bar{x}_2) \end{aligned}$$

Where S_n refers to the $m \times m$ covariance matrix of group n , x_0 is a matrix of m environmental variables to be classified, and \bar{x}_n is an $m \times 1$ matrix of the average environmental variables for group n . The covariance matrix and mean matrix are derived from the data collected, and new data stored in x_0 is used from different case studies. The right hand side of the equation deals with prior probabilities p_1 and p_2 and costs of misclassification ($c(1|2)$ and $c(2|1)$). Costs of misclassification are assumed to be equal in this study. Costs can be adjusted by future users to weight the discriminant function towards assigning a specific group if necessary. Prior probabilities are used when discriminating SATCAST data, where p_1 is the probability that a detection will be false ($1 - \text{SoS value}$) and p_2 is just SoS. It is important to mention that current SoS values take into account several environmental variables through logarithmic regression. The training dataset for calculation of the covariance and mean matrices was created without considering SoS values, so statistical differences found between indication groups are not

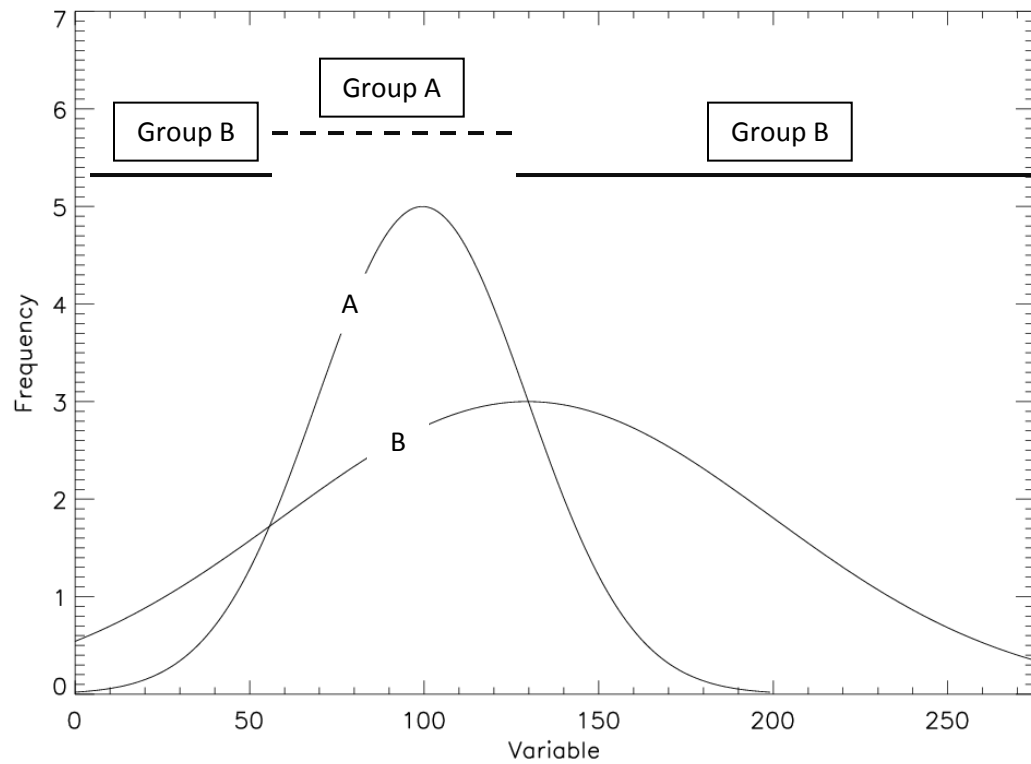


Figure 3.5. Sample variable distribution comparison performed by QDA. Two group distributions are presented, Group A and Group B. QDA will use these distributions on and classify new data into groups based on probability (area under the curve). For example, if these distributions are used to classify a new point with a variable value of 100, that variable would be assigned to group A given a higher probability. In this paper this technique is performed on a multivariate basis.

affected. Use of SoS in prior probabilities will act to merge the two statistical approaches for incorporation of environmental variables, and should not harm the end result of the QDA. However, if a variable is deemed not significantly different, a QDA can be further improved if the SoS value does not consider the environmental variable in question.

While the calculation of SoS is beyond the scope of this paper, future studies can use the statistically significant differences found in environmental variables of positive and false indications to improve new product output.

Since UWCTC is not produced in probability format, prior probabilities are estimated based on the given CTC value. p_1 is assumed to be 0.25 for weak UWCTC signals ($-10 < \text{CTC} < -4 \text{ K } 15 \text{ min}^{-1}$), 0.5 for moderate signals ($-20 < \text{CTC} < -10 \text{ K } 15 \text{ min}^{-1}$) and 0.75 for strong signals ($\text{CTC} < -20 \text{ K } 15 \text{ min}^{-1}$) (Hartung et al. 2013). Prior probabilities can be adjusted by future users to account for observed CTC with different parameterizations.

The QDA function is initially developed using the two collected study periods, where values are determined for the mean and covariance matrices. The means and covariances of variables with statistically significant differences between positive and false indication groups are the only variables considered. QDA performance is then evaluated on separate case studies collected that are not part of the original dataset. Evaluation is done using a dichotomous confusion matrix, which quantifies performance of a QDA by determining the matches between actual and classified groups (Table 3.3). QDA-resolved false variables are removed to create a filtered satellite-based forecasting product. It is common in previous literature to evaluate algorithm skill using validation

Table 3.3. Example confusion matrix. Columns represent QDA classification, rows represent actual classification. Percentage of improvement is found using

$$j = 100\% * ((\frac{a}{e}) - (\frac{c}{d}))$$

Example

Actual	Predicted		
	False	Positive	Total
False	a	b	e
Positive	c	d	f
Total	g	h	i

Percentage of Improvement: j

statistics such as POD, FAR and CSI for UWCTC and Brier score values for SATCAST. The confusion matrix approach is used to determine if overall skill scores (CSI and Brier score) have changed using the relative difference between positive and false indications. The percentage of improvement (POI) is the relative number of false indications removed compared to the relative number of positive indications removed. Positive POI values suggest that more false indications than positive indications have been removed, thus the resulting product will have improved skill scores. Five case study days in 2013 are selected to examine the effectiveness of the QDA using the confusion matrix approach.

h. Case Studies

The first case study occurred from 1800 UTC 9 April 2013 to 0300 UTC 10 April 2013, where post-frontal elevated thunderstorms developed with observed reflectivity values > 70 dBZ. This case presented an example where traditional surface boundary analysis would not help in a nowcasting sense, and yet several thunderstorms produced > 1 " diameter hail. A majority of severe convection developed where surface temperatures were near freezing (Figure 3.6). 9 April 2013 was chosen to analyze the effectiveness of an NWP data fusion method in an otherwise 'unusual' convective situation behind a shallow cold front.

The second case study chosen was 1800 UTC 20 May 2013 to 0000 UTC 21 May 2013. A surface low located over the Northern Plains stalled with a cold front/stationary boundary draped from South Dakota to the northern Texas panhandle (Figure 3.7). A dry line present over west-central Oklahoma produced discrete supercells in a "classic"

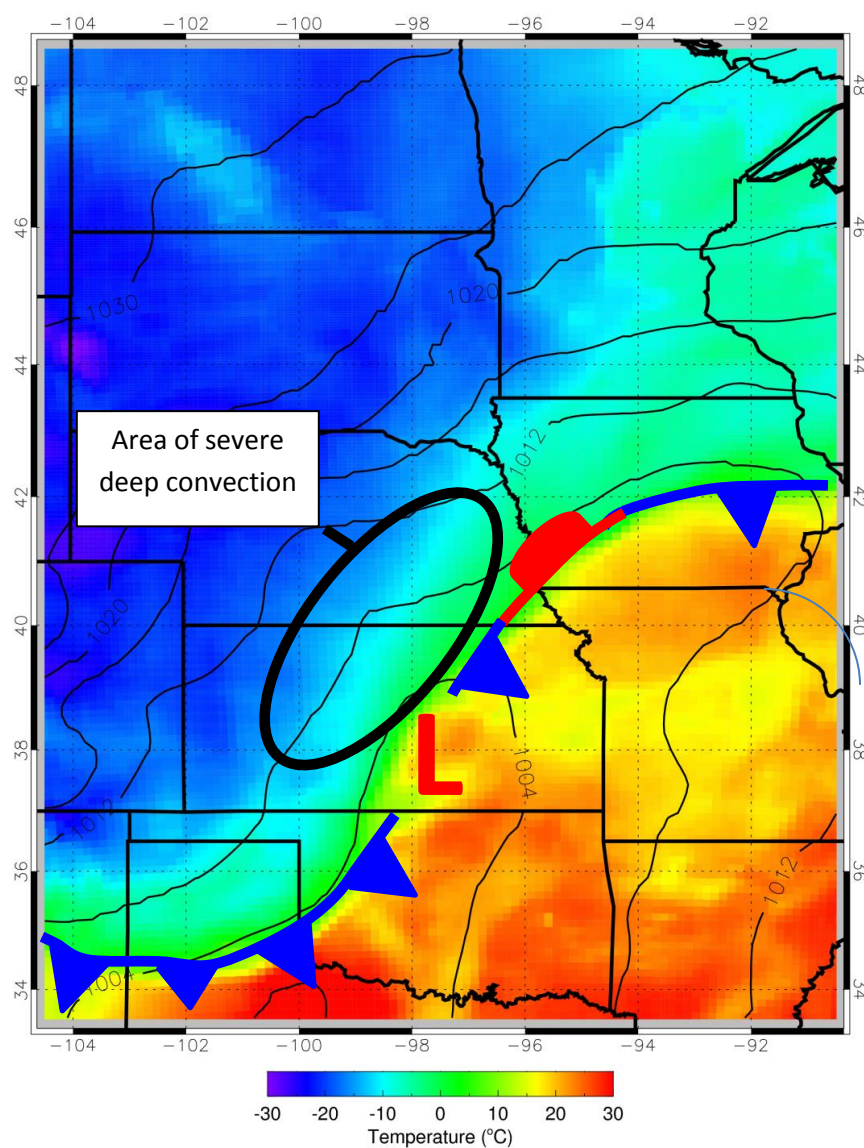


Figure 3.6. 1800 UTC 9 April 2013 RAP MSLP and surface temperature with surface boundaries.

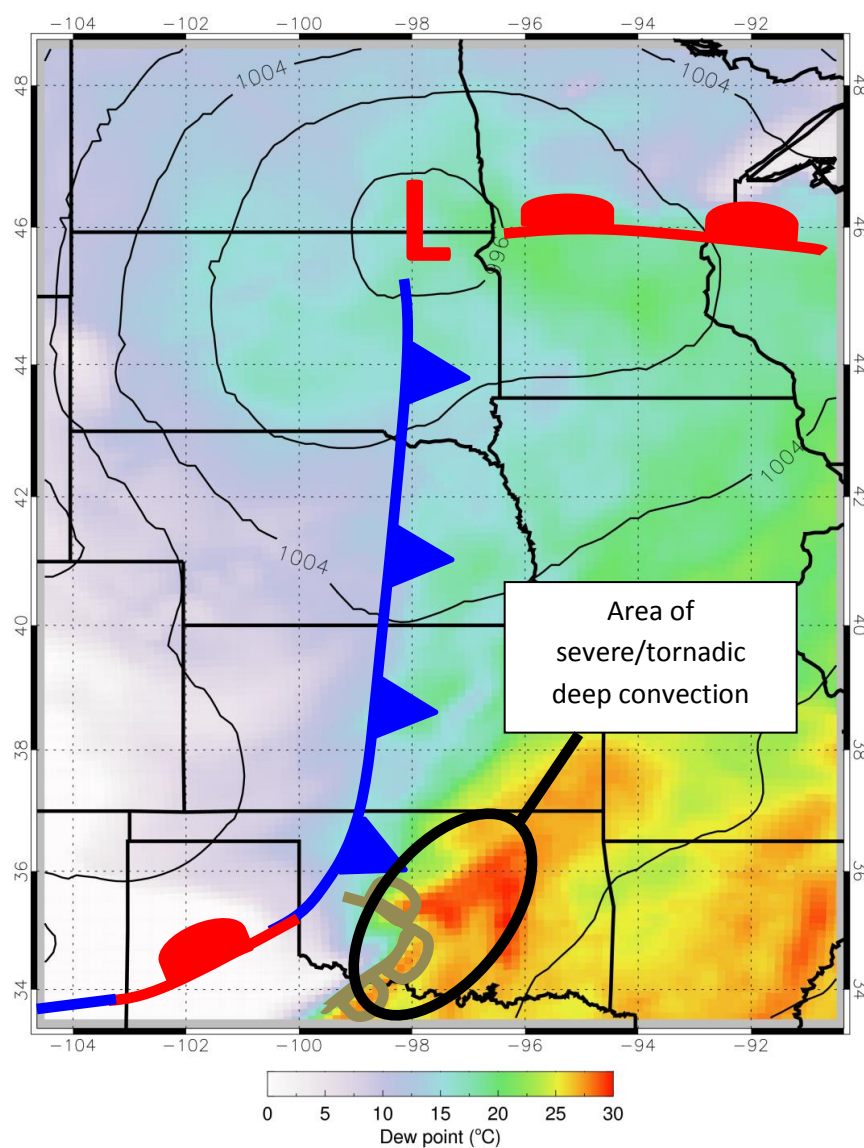


Figure 3.7. 1800 UTC 20 May 2013 RAP MSLP and surface dew point with surface boundaries.

severe convective scenario. With CAPE values reaching $> 5000 \text{ J kg}^{-1}$ and strong 0-6 km shear, a supercell developed and spawned a deadly tornado in Moore, Oklahoma. This day was chosen to analyze NWP-fused CTC-based algorithms in a high CAPE/high shear environment, so it is expected that data fusion works well.

The third and fourth cases are analyzed over a period of two days (1200 UTC 20 June 2013 through 1200 UTC 22 June 2013). 20 June 2013 had modest CAPE $> 2000 \text{ J/kg}$ with very little CIN. Given the lack of a robust capping inversion, CI along the quasi-stationary boundary that occurred quickly grew upscale into a large MCS that propagated through east/southeast Minnesota. This particular case is useful given that CIN would not contribute to removal of indications, so it may be helpful to observe the characteristics of false discriminant classifications in the NWP-fused CI algorithms. The fourth case study occurred on 21 June 2013 and was chosen for the same reasons as 20 June 2013. CI developed with enhanced convergence on the nose of a low level jet along the quasi-stationary boundary (Figure 3.8).

The fifth and final case study occurred on 24 June 2013. CI events occur at 0400 UTC 25 June 2013 along an enhanced convergent boundary in eastern Nebraska (Figure 3.9). A zone of less severe nocturnal convection developed later in central Iowa. The nocturnal storms, while not numerous, produce a unique look at substantial nocturnal events with reflectivity values $> 60 \text{ dBZ}$ in the Great Plains. All case studies are used to evaluate QDA performance. An advantage of using case studies will be to understand how well probability-based data filtering will work in an operational setting.

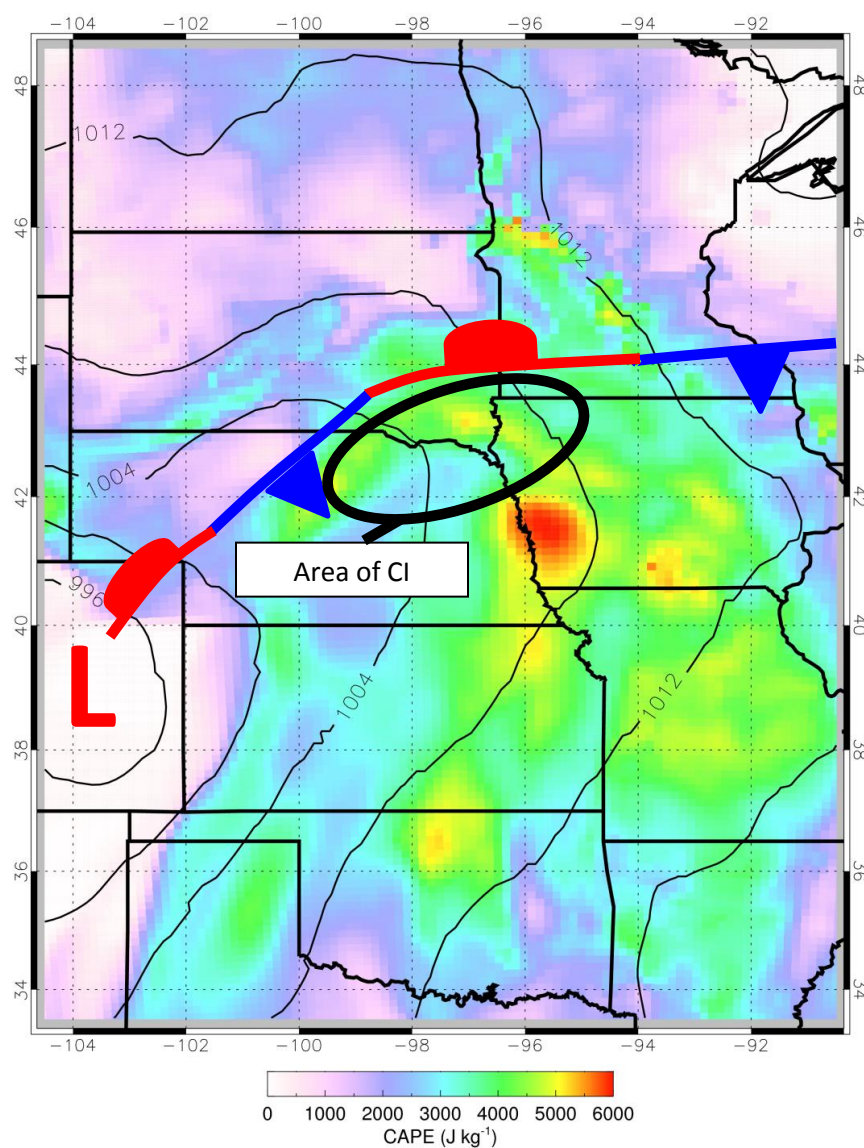


Figure 3.8. 0000 UTC 22 June 2013 RAP MSLP and MUCAPE with surface boundaries.

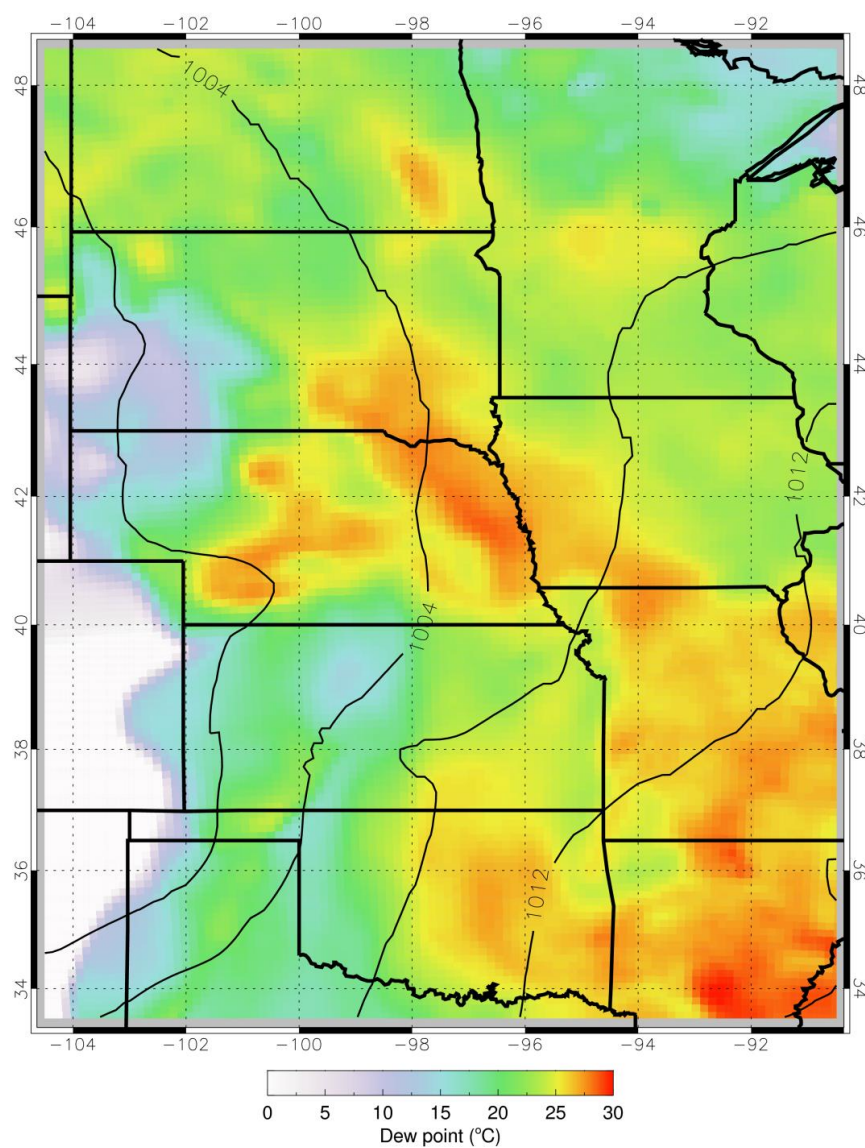


Figure 3.9. 0200 UTC 25 June 2013 RAP MSLP and surface dew point.

4. Hand Validation

An objective validation is designed to mimic the interpretation of CI by a forecaster. Thus the design of the objective scheme is optimized through comparison to a subjective validation. The objective scheme that best resembles the small-scale subjective validation is used. The subjective validation revealed optimal settings for the WDSS-II clustering algorithm (Table 4.1). The default half-size of 5 is too aggressive and removed results that could be important in an objective validation scheme. As lower half-sizes are used, several important features became visible to the clustering process (Figure 4.1). It is clear that increasing the half-sizes of the smoothing parameters allows for smaller saliency scales to be more skillful. Less smoothing results in too many clusters being identified and thus presents a need for larger saliency scales to improve the skill scores. For this study the medium (and most skillful) value half-size of 3 with a 30 pixel saliency scale was chosen.

A simple test of how well the validation works is to compare the skill scores found here to previous findings. Since the current UWCTC algorithm was extensively analyzed for skill in previous studies, this algorithm was chosen for subjective interpretation. The objective validation approach produced a 36.1% POD and 45.5% FAR for the UWCTC algorithm. The validation statistics for UWCTC presented here are less skillful than the statistics found in Sieglaff et al. (2011) and Hartung et al. (2013) (FAR values around 40%, POD around 50%). However, UWCTC presented a 45.2% POD and 27% FAR with subjective validation. In a closer examination, many subjectively identified positive indications occurred after CI was detected on a complex.

Table 4.1. Validation of WDSS-II/ThOR CI clusters compared to subjectively identified truth CI clusters. WDSS-II settings presented as low smooth (50th percentile smoothing over a 5x5 box), medium smooth (50th percentile smoothing over a 7x7 box) and high smooth (50th percentile smoothing over a 9x9 box).

WDSS-II Settings	Min. Scale	POD	FAR	CSI
High Smooth	20	73.1	20.4	63.5
Med Smooth	30	69.7	18.6	66.67
Low Smooth	40	62.7	18.6	61.2

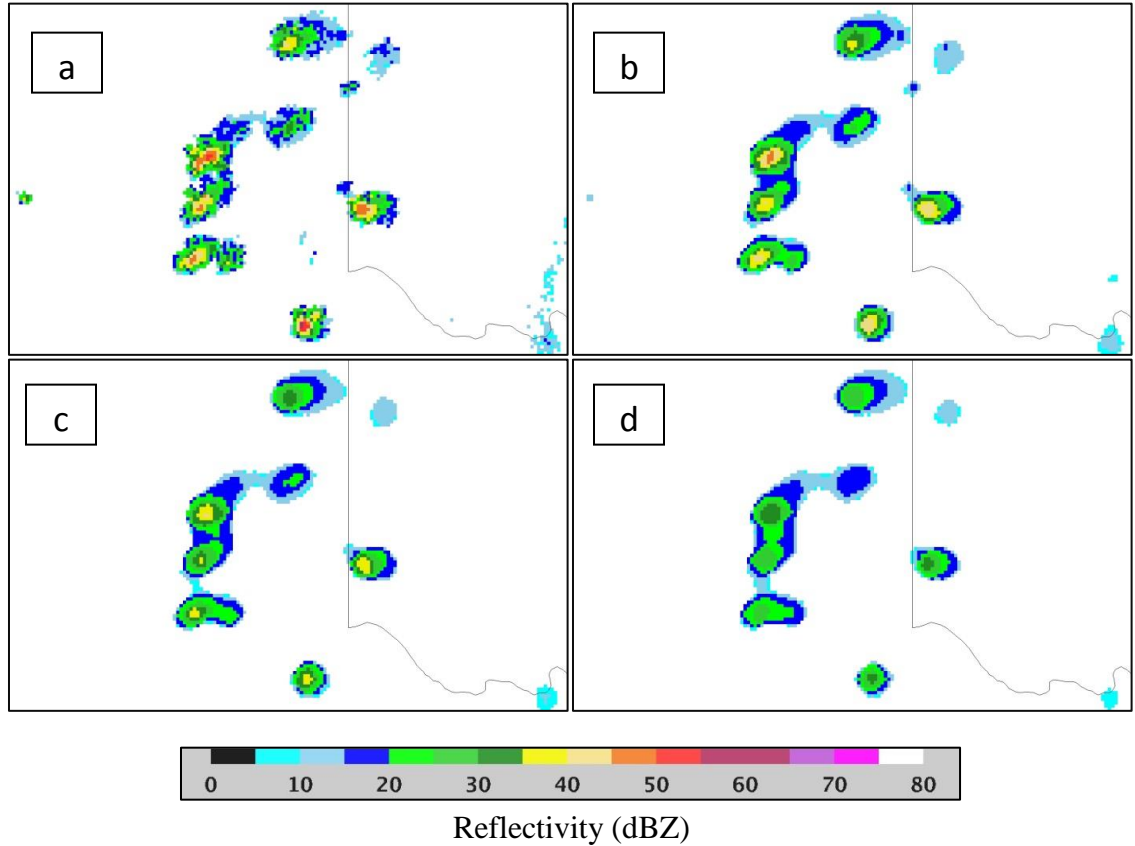


Figure 4.1. a) Example of unsmoothed composite NMQ data over the Texas panhandle and southwest Oklahoma. b) Light smoothing applied, 50th percentile filter over a 5x5 box. c) Medium smoothing applied, 50th percentile filter over a 7x7 box. d) Heavy smoothing, 50th percentile filter over a 9x9 box. Colors hotter than yellow are above 35 dBZ and considered convective.

A post-CI CTC indication is a satellite-based indication that occurs on a convective complex after it has reached reflectivity spatial and quantitative thresholds set to identify CI. Most post-CI CTC indications were indicators of a strengthening system in which reflectivity would increase after the forecast was made. Since subjective validation considers post-CI CTC indications as reasonable forecasts, similar skill scores were found for CTC indications reported in Sieglaff et al. (2011) and Hartung et al. (2013). Upon further inspection it was found that the increased FAR (decreased POD) directly resulted from the exclusion of indications that occur on convective systems after CI was detected.

If CTC indications are considered positive for a 17 min period after CI detection (following Hartung et al. 2013), the objective validation used here produces similar validation statistics to previous studies. However, allowing for successful CTC indications after CI (Hartung et al. 2013) provides no benefit to environmental analysis and risks contaminating environmental data collection with respect to performance (Figure 4.2). All positive post-CI CTC indications occur within close proximity to ongoing convection, and therefore are excluded from an environmental analysis regardless of objective positive or false classification. Post-CI CTC false indications are not always in close proximity to ongoing convection, and therefore could be included in an environmental analysis. It is crucial to classify post-CI CTC indications as false for a correct assessment of environmental effect on CTC algorithm performance with a model that cannot resolve convective time scales. Thus all validations in this study are not a finalized evaluation of algorithm ‘skill’. Rather this validation technique is good for

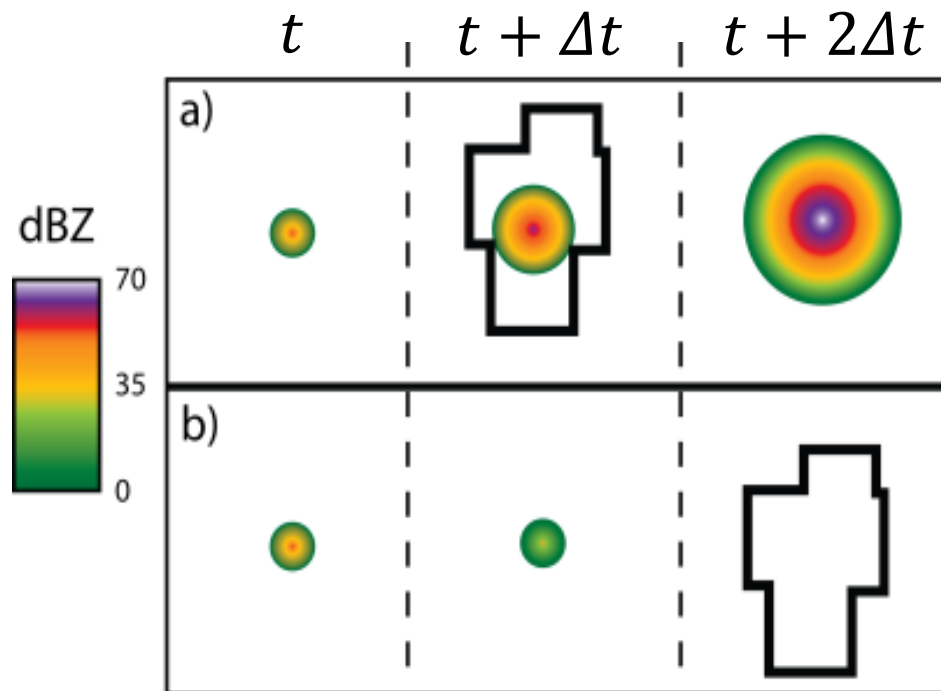


Figure 4.2. Radar shown in dBZ at three time steps, where initiation occurs at time t . Black outlined polygons represent CTC indications. a) Example of a positive post-CI CTC indication, where reflectivity values intensify after cooling. b) Example of a false CTC detection occurring after CI, with no nearby ongoing convection.

objectively identifying regions where positive or false indications occur for environmental analysis. Therefore validation statistics (POD, FAR, CSI and Brier score) should not be calculated using this methodology. Objective validation methods for environmental analysis that do include post-CI indications will need to use models that sufficiently resolve convective-scale motions.

Objective validation determines positive and false groups within the July 2012 and April 2013 study periods. Objective validation also determined positive and false groups for selected case studies (Figure 4.3). For UWCTC environmental analysis there are 1211 (284) false (positive) indications during the July study period, and 797 (391) false (positive) indications during the April study period. For SATCAST environmental analysis there are 357,496 (4,209) false (positive) indications for the July study period and 282,965 (1,160) false (positive) indications for the April study period. These numbers may seem daunting to prospective users, however the reader is reminded that these data are all indications that are at least 50 km away from ongoing convection. It is common for positive indications to be clustered along a boundary. If several positive indications occur near or along a boundary, the first indications will be collected in the dataset and all subsequent indications cannot be considered. Also, SATCAST indications are produced on all cloud types determined pre-convective (Walker et al. 2012). SoS is only changed based on witnessed interest fields. For the July study period, of the 357,496 false indications found, 187,355 indications are below 30% SoS (~50% of all false indications). Approximately 85% of the 4209 positive indications have SoS values greater than 30%. For the April study period, of the 282,965 false indications, 254,025

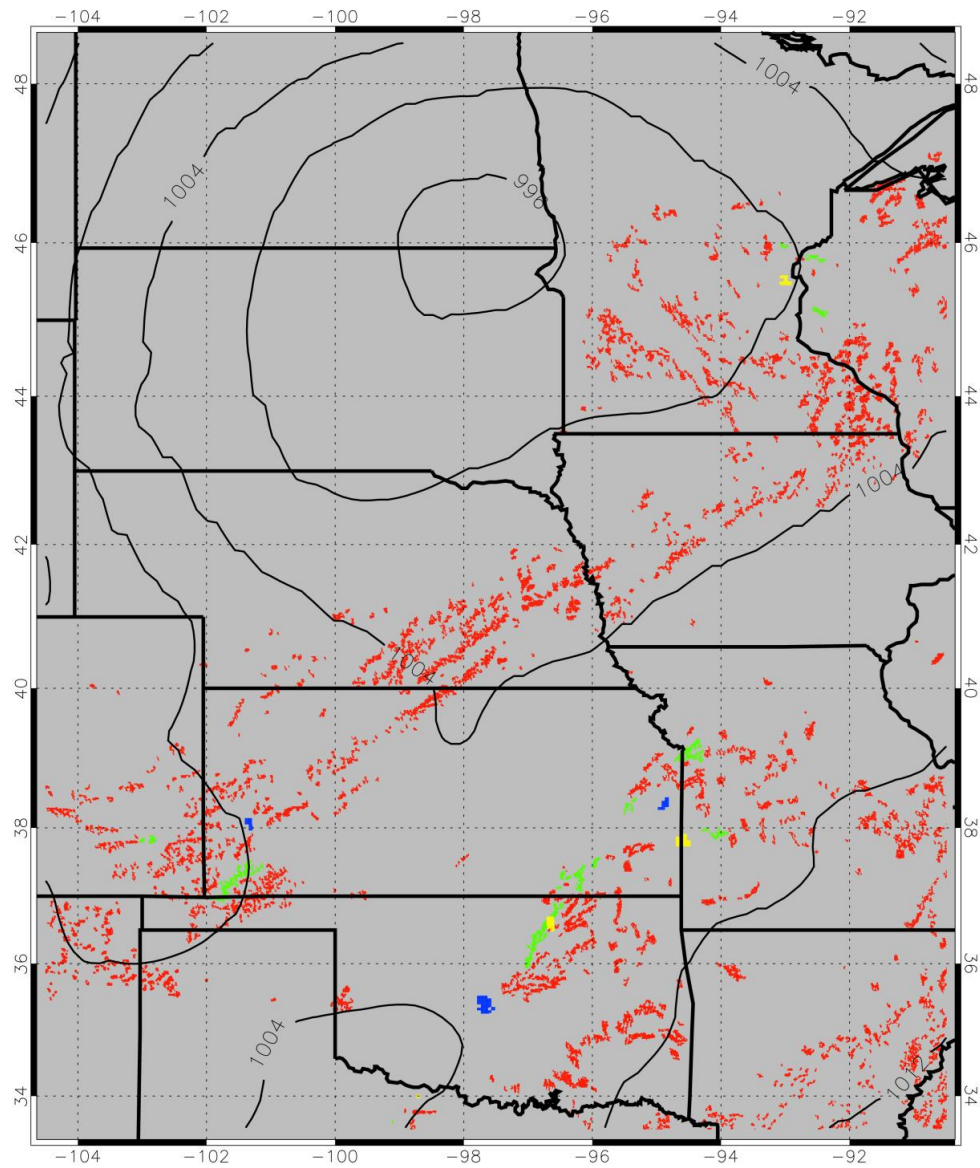


Figure 4.3. 1910 UTC 20 May 2013 example of validated UWCTC and SATCAST data one hour before the Moore, Oklahoma, tornado with contoured MSLP (hPa). Positive indications are shown in green (blue) for SATCAST (UWCTC). False indications are red (yellow) for SATCAST (UWCTC).

contain a SoS value less than 30% (~90% of all false indications). 457 positive indications, or 40% of all positive indications, have a SoS value greater than 30%. The substantial improvement in SoS value for the false indications in the April SATCAST is due to the incorporation of convective variables into the SoS value calculation.

SATCAST output post-April 2013 now includes several convective variables in the logarithmic regression equation, including CAPE, CIN, LCL, LFC, CCL and many others which will be discussed in upcoming publications (Mecikalski, personal communication: 2013). There are still a large number of false indications with SoS values greater than 30%, which suggests QDA incorporation may be helpful.

5. Results

a. Principal Components Analysis

A statistical examination is performed to determine if there is any possible pattern between pre-convective environments and indication performance. PCA is the first step in the statistical examination, and used to determine if relationships can be seen along the main sources of variation with respect to satellite-based CI forecasting algorithm performance. PCA for the UWCTC indications determined that the top two sources of variance in the July 2012 dataset explained 45% of overall variation (Table 5.1). The eigenvector for the first principal component (PC1) suggests that the first source of variation is related to LFC height. CIN also has a notable effect on the first source of variation, despite not being among the top five components. However, variables related to LFC ($Z_{LFC-LCL}$ and Z_{EL-LFC}) had larger components in PC1. It is likely that changes in CIN are related to changes in $Z_{LFC-LCL}$. The second principal component (PC2) eigenvector was largely related to instability present with large eigenvector components in CAPE, LCL and NCAPE. The signs of PC1 and PC2 suggest that LFC height increases as PC1 increases and instability decreases as PC2 increases. Dependency upon values of PC1 and PC2 become apparent when the principal component scores are plotted (Figure 5.1). Positive indications favor areas with lower than average LFC heights, with ~80% of positive indications occurring in areas with $PC1 < 0$. The dependence on CAPE is not as clear with the July PCA. It is notable that there are a larger number of positive indications in more unstable regions. However, the relative difference between quadrants is too small to make a definitive conclusion about CAPE using only the PCA. The next

Table 5.1. July UWCTC PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.

	PC1	PC2	PC3	PC4	PC5
CAPE	-0.279	-0.321	-0.073	0.238	0.193
CIN	-0.287	0.068	-0.294	-0.324	0.136
LCL	-0.007	0.464	-0.305	0.072	0.198
LFC	0.365	-0.069	-0.210	0.249	-0.019
EL	-0.281	-0.181	-0.095	0.423	-0.007
RH	-0.071	-0.211	-0.515	-0.253	-0.038
LPL	0.180	0.015	-0.621	-0.025	0.132
LPLD	0.069	-0.073	-0.009	-0.052	-0.165
$T_c - T_e$	0.376	-0.131	-0.042	0.170	0.072
CCL	0.143	0.423	-0.099	0.289	0.088
MID LR	0.067	0.242	0.142	0.323	0.474
LOW LR	-0.320	0.312	-0.005	0.036	-0.048
MWD	0.137	-0.103	0.227	-0.236	0.444
NCAPE	-0.163	-0.296	-0.142	0.300	0.266
NCIN	-0.070	0.042	0.029	-0.004	0.085
Z_{EL-LFC}	-0.398	-0.110	0.026	0.215	0.004
$Z_{LFC-LCL}$	0.326	-0.303	-0.026	0.174	-0.118
SM	0.045	-0.185	0.081	-0.292	0.575
Variation Explained	27.10%	18.03%	10.45%	9.91%	6.46%

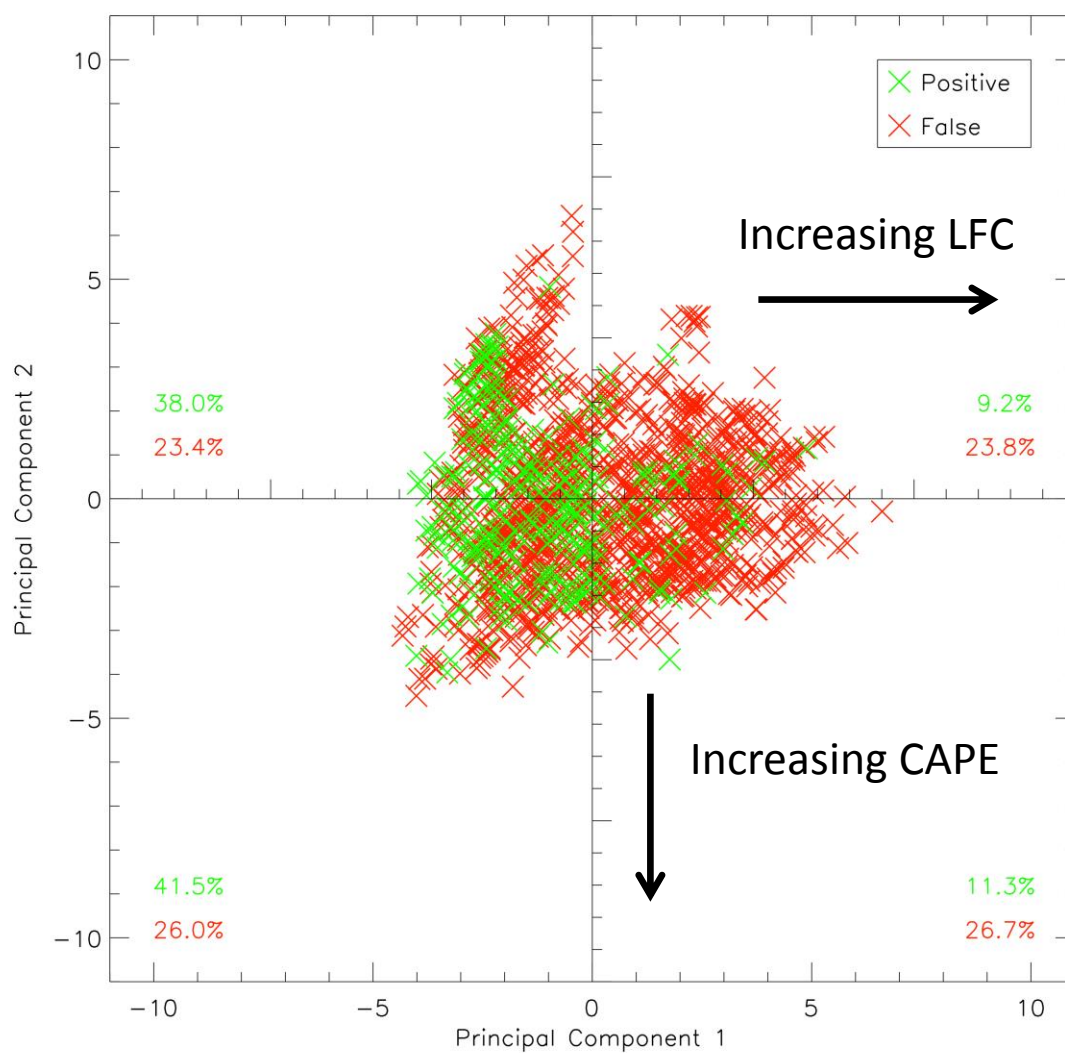


Figure 5.1. July UWCTC study period plot of the first (x axis) and second (y axis) principal components. Positive (false) UWCTC indications shown in green (red).

three components (PC3, PC4 and PC5) for the July UWCTC are more difficult to interpret, though still should be considered given that PC3 through PC5 account for ~26% of variation. PC3 has a notably large dependence upon LPL, suggesting that PC3 is related to the elevation of the largest instability. PC3 also has large eigenvector components for CIN, LCL and RH. As PC3 increases, LPL height is reduced, RH is reduced, there is more CIN (CIN becomes more negative) and LCL height is reduced. PC4 has a large dependence on CIN, suggesting that CIN is in fact one source of variation. However, the signs and magnitude of EL and MIDLR components are not consistent with an increase in stability alone. Thus PC4 is interpreted as a combination of stability and stable layer depth. PC5 has large magnitudes of SM, MWD and MIDLR, so it is interpreted as the general motion and tractability of an indication. Together, the five principal components explain ~72% of variation.

The April 2013 UWCTC data PCA displayed similar sources of variation as the July data (Table 5.2). PC1 and PC2 again explain the majority of the variation (~52%). However, the April PC1 did not lend itself to the same interpretation as the July PC1. The April PC1 dependence did not appear to be linked with LFC height, given the relatively low magnitude of the Z_{EL-LFC} eigenvector component. The April PC1 did display a dependence on $Z_{LFC-LCL}$, with another notably large magnitude on LOWLR. This lends itself to the interpretation that higher April PC1 scores suggest environments conducive of deeper stable layers visible by satellite. April PC2 again is related to instability, as higher PC2 scores are almost directly linked to less statically unstable environments. PC3, PC4 and PC5 again explain ~25% of the variation. PC3 appears to

Table 5.2. April UWCTC PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.

	PC1	PC2	PC3	PC4	PC5
CAPE	0.225	-0.383	0.100	-0.112	0.090
CIN	-0.076	-0.151	-0.351	-0.603	-0.081
LCL	-0.205	0.273	0.177	-0.372	-0.324
LFC	0.285	0.296	0.151	0.136	-0.220
EL	0.146	-0.367	0.234	0.008	-0.241
RH	0.288	-0.104	-0.250	-0.133	-0.294
LPL	0.302	0.252	-0.032	-0.234	-0.273
LPLD	0.059	0.002	-0.153	-0.058	0.669
T _c -T _e	0.254	0.164	0.287	-0.223	0.268
CCL	-0.151	0.200	0.459	-0.321	0.178
MID LR	-0.084	-0.105	0.547	0.002	-0.049
LOW LR	-0.395	-0.064	0.050	0.054	0.000
MWD	0.265	0.215	-0.100	-0.110	0.110
NCAPE	0.284	-0.322	0.125	-0.077	0.062
NCIN	-0.109	0.025	0.050	0.307	-0.134
Z _{EL-LFC}	-0.014	-0.459	0.127	-0.059	-0.098
Z _{LFC-LCL}	0.372	0.124	0.039	0.326	-0.023
SM	0.266	-0.024	0.173	-0.121	0.123
Variation Explained	29.46%	22.46%	11.29%	7.16%	6.05%

be a source of variation caused by MIDLR and CCL. Again the signs of the eigenvector components do not lend PC3 to an overall interpretation, and should rather be considered as a source of variation caused by increasing mid-level stability and decreasing low-level moisture. PC4 is interpreted as amount of stability with large eigenvector components in all CIN-related variables. LPLD has the largest component in PC5, which suggests that the April PC5 may be related to available forcing factors (i.e. convergent boundaries). The July PCA did not have an LPLD source of variation, possibly due to forced convection being uncommon during this time period. Since all 5 principal components explain ~77% of variation, all components except SM are considered important in the variation of the April UWCTC study period. The differences between July and April variation suggest that the choice of initial distributions for probability-based discrimination should vary for different time periods. April UWCTC also suggests dependence on both the amount of instability and stability in the atmosphere. Almost 80% of all positive indications occur in more unstable regions (Figure 5.2). Since almost 50% of positive indications occur in areas with both larger instability and smaller $Z_{LFC-LCL}$ layers, PCA suggests a relationship between instability and stability with respect to UWCTC performance.

The PCA for July SATCAST dataset (Figure 5.3) had similar variations explained with the top five eigenvectors (Table 5.3). PCA for SATCAST has a much larger sample size of points in the July dataset than UWCTC (Figure 5.1 and 5.2). The SATCAST July PC1 suggests similar dependence on LFC height as UWCTC. Larger PC1 scores are related to lower LFC heights, lower $Z_{LFC-LCL}$ and higher Z_{EL-LFC} . PC2 for the July

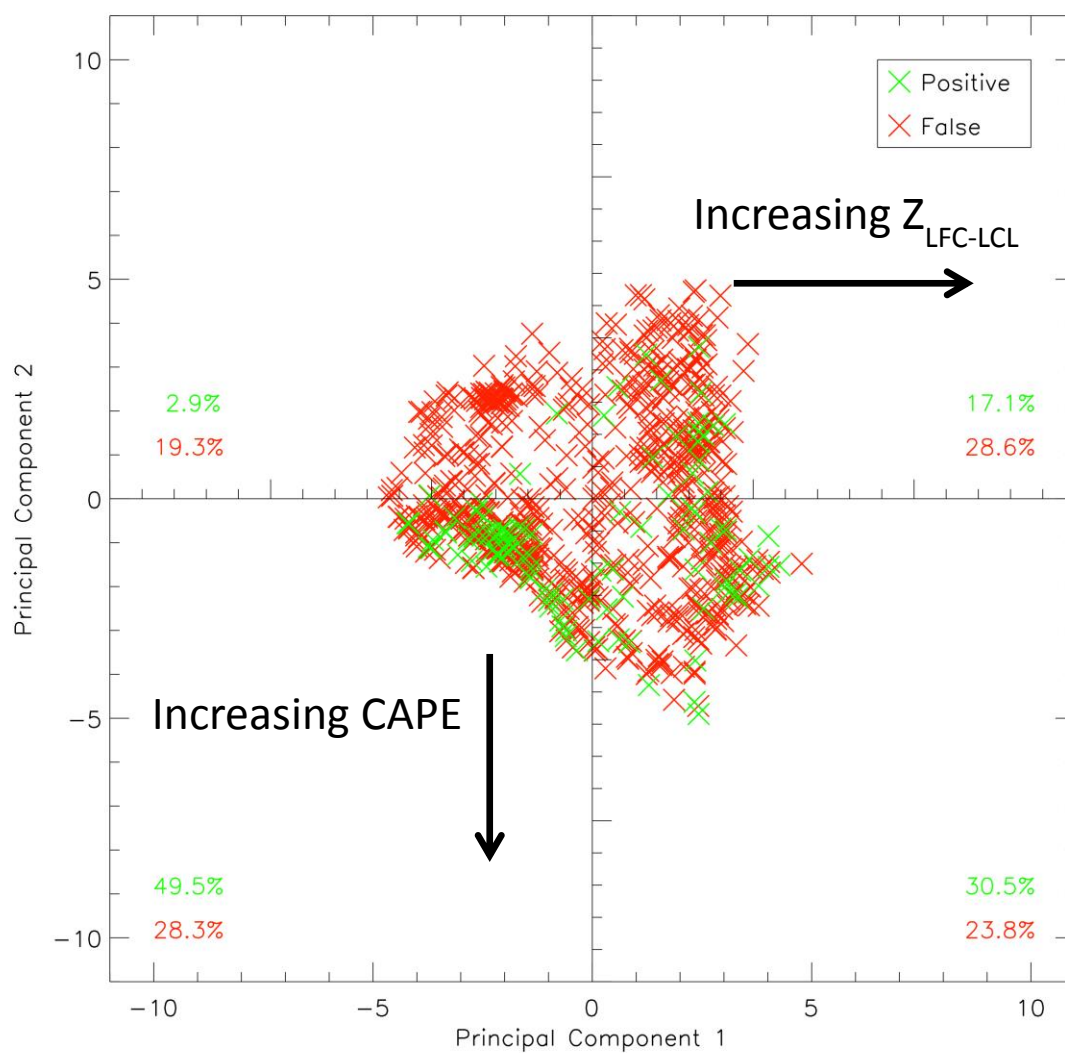


Figure 5.2. April UWCTC study period plot of the first (x axis) and second (y axis) principal components. Positive (false) UWCTC indications shown in green (red).

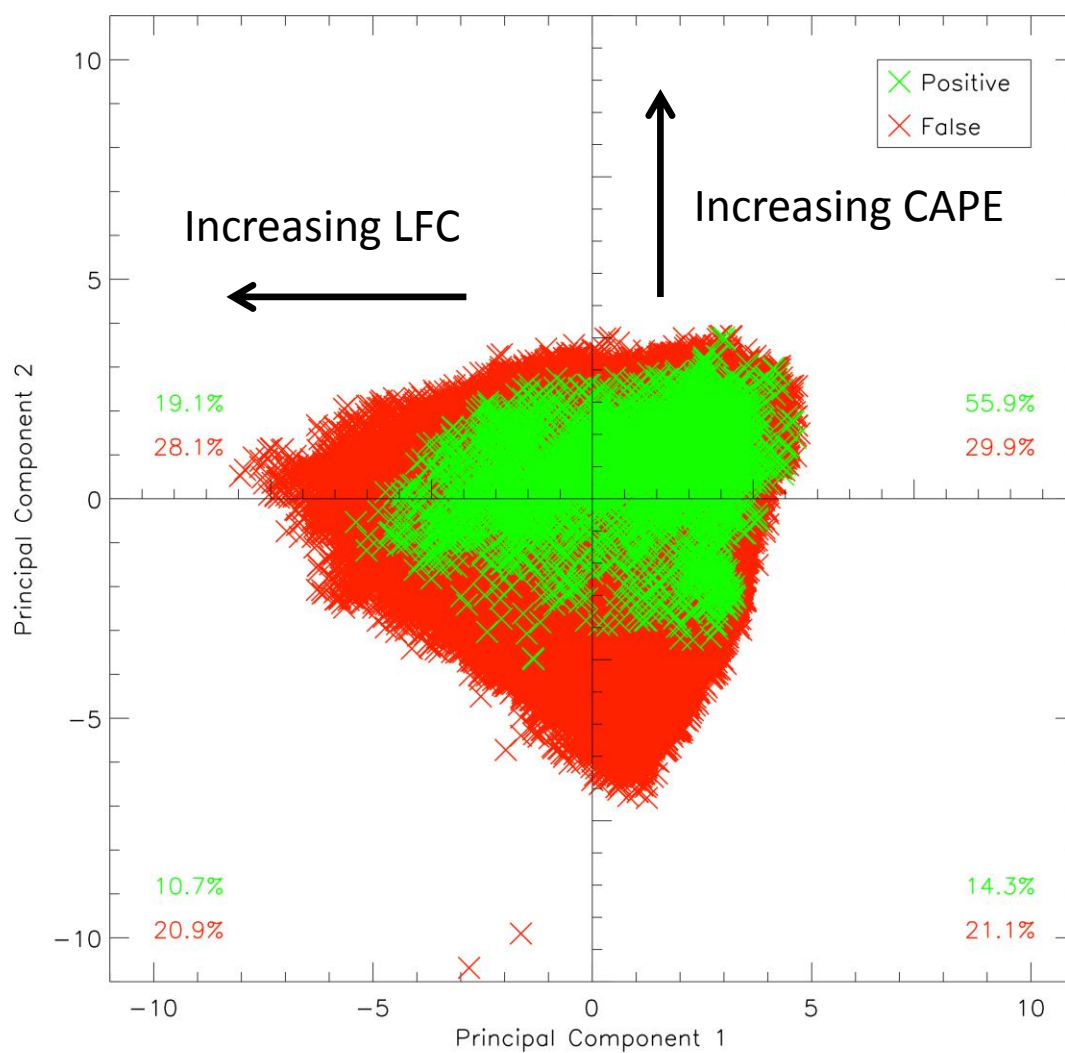


Figure 5.3. July SATCAST study period PCA with the first principal component on the x axis and the second principal component on the y axis. Positive (false) SATCAST indications are shown in green (red).

Table 5.3. July SATCAST PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.

	PC1	PC2	PC3	PC4	PC5
CAPE	0.199	0.310	0.298	0.278	0.207
CIN	0.297	-0.067	-0.299	0.240	0.084
LCL	-0.146	0.225	-0.534	0.049	0.121
LFC	-0.417	0.107	0.015	0.143	-0.045
EL	0.152	0.448	0.240	0.257	0.097
RH	0.132	-0.242	-0.169	0.528	-0.020
LPL	-0.212	-0.013	-0.322	0.536	0.082
LPLD	-0.015	-0.062	0.062	0.031	-0.389
T _c -T _e	-0.397	0.050	0.158	0.178	0.067
CCL	-0.259	0.367	-0.253	-0.144	0.049
MID LR	-0.229	0.259	-0.042	-0.195	0.377
LOW LR	0.289	0.231	-0.284	-0.268	-0.050
MWD	-0.046	-0.236	0.095	-0.086	0.673
NCAPE	0.001	0.003	0.009	0.008	0.010
NCIN	0.001	0.019	0.002	-0.018	-0.016
Z _{EL-LFC}	0.326	0.337	0.200	0.155	0.105
Z _{LFC-LCL}	-0.340	-0.034	0.346	0.117	-0.122
SM	0.075	-0.382	0.037	-0.034	0.368
Variation Explained	25.28%	15.37%	13.17%	8.89%	6.25%

SATCAST also displays a dependence upon instability, with a notably large component for storm motion. Thus higher PC2 scores suggest slow moving storms in more unstable environments. The PCA plot for July SATCAST (Figure 5.3) suggests a much larger performance on PC2 than the UWCTC. Positive indications are clustered in environments with lower LFCs and higher instability (with ~56% of all positive indications in the upper right quadrant of Figure 5.3). Few positive indications are observed below a PC2 score of -3, suggesting that there exists a minimum necessary instability factor for a SATCAST indication to be positive. As with previous analyses, PC1 and PC2 explain only parts of the variation. It is important to consider PC3, PC4 and PC5 for the July SATCAST PCA. PC3 is largely coupled with the amount of moisture present with high eigenvector components in LCL and related variables. PC4 is interpreted as a height of the LPL. PC5 is interpreted the same as UWCTC, where high component values of SM, MWD and MIDLR suggest a dependence on overall storm motion and tractability. Together, all components explain ~69% percent of overall variation, which is notably similar to UWCTC PCA.

The April SATCAST PC1 has large dependence on values related only to $Z_{LFC-LCL}$ and excluded values related to the actual value of CIN (Table 5.4). This result displays a consistently large source of variation explained by $Z_{LFC-LCL}$ throughout all CTC-based products. It is notable that SATCAST in April is also sensitive to LFC height. However, the lack of a robust Z_{EL-LFC} component prevents an interpretation based only on LFC. The April SATCAST PC2 is again instability, with the largest components of eigenvectors in CAPE and related variables. PC3, PC4 and PC5 do not appear to have

Table 5.4. April SATCAST PCA eigenvectors with amount of variation explained by the principal component. The top five variables are shown in bold for each principal component.

	PC1	PC2	PC3	PC4	PC5
CAPE	-0.053	-0.458	0.195	-0.185	-0.127
CIN	-0.226	0.085	-0.207	-0.423	-0.398
LCL	0.117	0.403	0.271	-0.208	-0.301
LFC	0.441	-0.033	0.045	0.133	-0.071
EL	-0.097	-0.428	0.315	0.031	-0.224
RH	0.052	-0.207	-0.419	-0.177	-0.375
LPL	0.372	0.049	-0.099	-0.219	-0.397
LPLD	0.049	-0.017	-0.130	0.039	-0.061
T _c -T _e	0.397	-0.038	0.109	0.037	-0.099
CCL	0.153	0.314	0.413	-0.039	-0.125
MID LR	0.114	-0.084	0.495	0.040	-0.099
LOW LR	-0.361	0.260	0.158	0.009	0.024
MWD	0.162	-0.038	0.048	-0.556	0.363
NCAPE	0.005	-0.018	0.001	-0.009	-0.026
NCIN	-0.001	0.005	-0.002	0.004	-0.001
Z _{EL-LFC}	-0.289	-0.359	0.255	-0.034	-0.163
Z _{LFC-LCL}	0.367	-0.259	-0.119	0.241	0.104
SM	0.126	-0.138	0.091	-0.520	0.414
Variation Explained	24.11%	16.76%	13.31%	8.80%	6.77%

a single clear meaning, despite that they explain ~29% of variation. PC3, PC4 and PC5 could be related to surface moisture, system motion and CIN respectively.

Further analysis of the first two principal components of April SATCAST yields two distinct clusters of positive indications (Figure 5.4). Since the first component is again related to the depth of the stable layer, the first positive detection cluster appears to occur when instability is low or nonexistent. The apparent PC1 minimum is caused by data censorship when $CIN = 0$, thus the PC1 score near ~ 3 should be interpreted as near-zero stability with very low or negative $Z_{LFC-LCL}$. However, this begs the question why a second cluster would be seen with relatively large (compared to the first) values of $Z_{LFC-LCL}$. The clustering could be caused by the abundance of forced convection in April, yielding two different stability environments where convection is witnessed. Despite this clustering, again a majority of positive indications are located in areas with small $Z_{LFC-LCL}$ values and large instability.

Throughout all four PCA analyses, it became clear that the largest source of variation in the overall data is related to $Z_{LFC-LCL}$ and not CIN. This large variation does not mean that $Z_{LFC-LCL}$ is a better discrimination variable to use than CIN. Determining which variable is a better discriminator is done using ANOVA, ANCOVA and MW testing. While it is possible to create plots along the PC3, PC4 and PC5 values for all study periods, the plots for PC1 and PC2 were sufficient to conclude that a general relationship is apparent between environmental variables and CI algorithm performance. For a more detailed analysis of just what that relationship is, tests for statistically significant differences are analyzed. Finding a general relationship with the largest

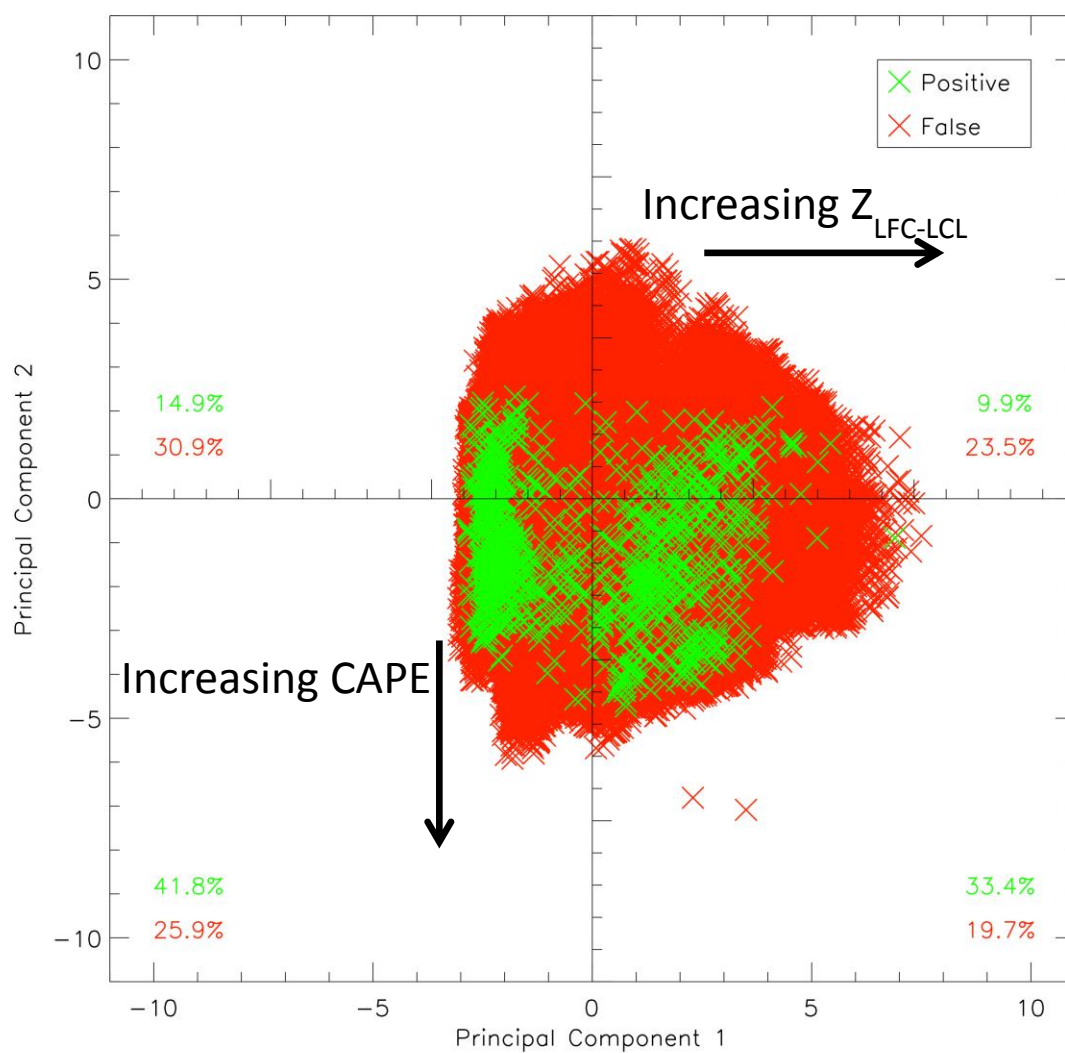


Figure 5.4. April SATCAST study period PCA with the first principal component on the x axis and the second principal component on the y axis. Positive (false) SATCAST indications are shown in green (red).

sources of variation ($Z_{LFC-LCL}$ and CAPE) does suggest that one or the other should be important.

b. Variable Contribution

The ultimate goal of a discriminant analysis is to only use RAP observations that contribute to CTC-based CI algorithm performance. Using observations that do not have significant differences in a QDA could result in bad classifications of new data. Several variables appear to have statistically significant differences upon initial examination (Figure 5.5). Variables such as CAPE and Z_{EL-LFC} vary as expected with notably larger values for positive indications than false indications in UWCTC and SATCAST (Figure 5.5 a, p). The importance of instability is also found in LFC and EL differences where positive indications have lower LFC values and higher EL values than false indications for both algorithms (Figure 5.5 d, e). CIN-related variables also have a consistent relationship, such that false indications are found in more stable regions for UWCTC (Figure 5.5 b, o, q). RH variables are consistently different across the time periods for UWCTC where larger RH is seen with positive indications (Figure 5.5 f). T_c - T_e also appears to have a relationship to performance, since lower values are common with positive indications (Figure 5.5 i). Several pre-convective environmental relationships yield differences that change depending on the time period. LPL is larger for positive UWCTC indications in July than false indications. Positive indication LPL is lower than false indications for April (Figure 5.5 g). The same characteristic is witnessed in UWCTC LCL, MIDLR and SM (Figure 5.5 c, k). The total dataset (combination of July and April study periods) with variables that exhibit differing relationships end up

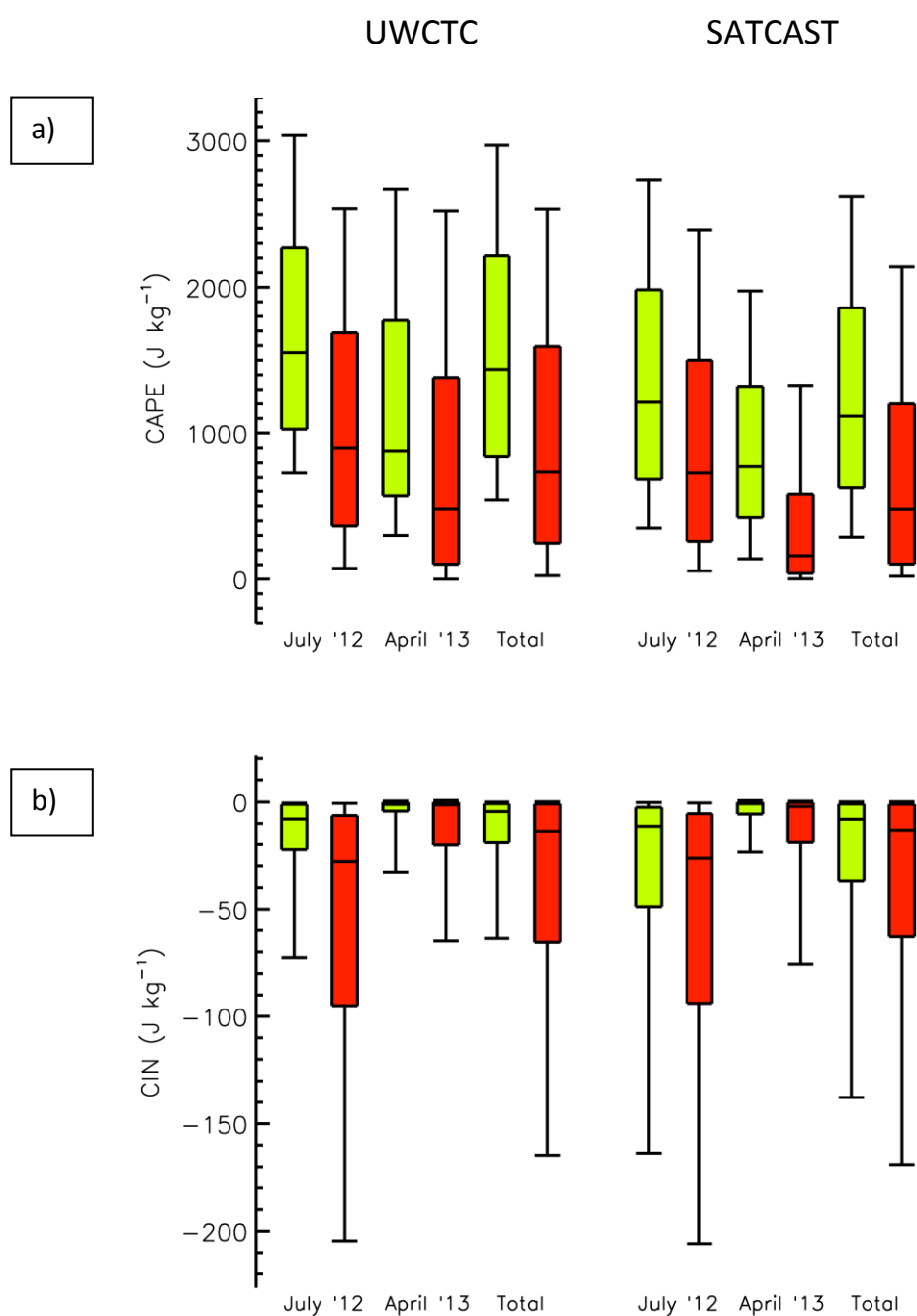
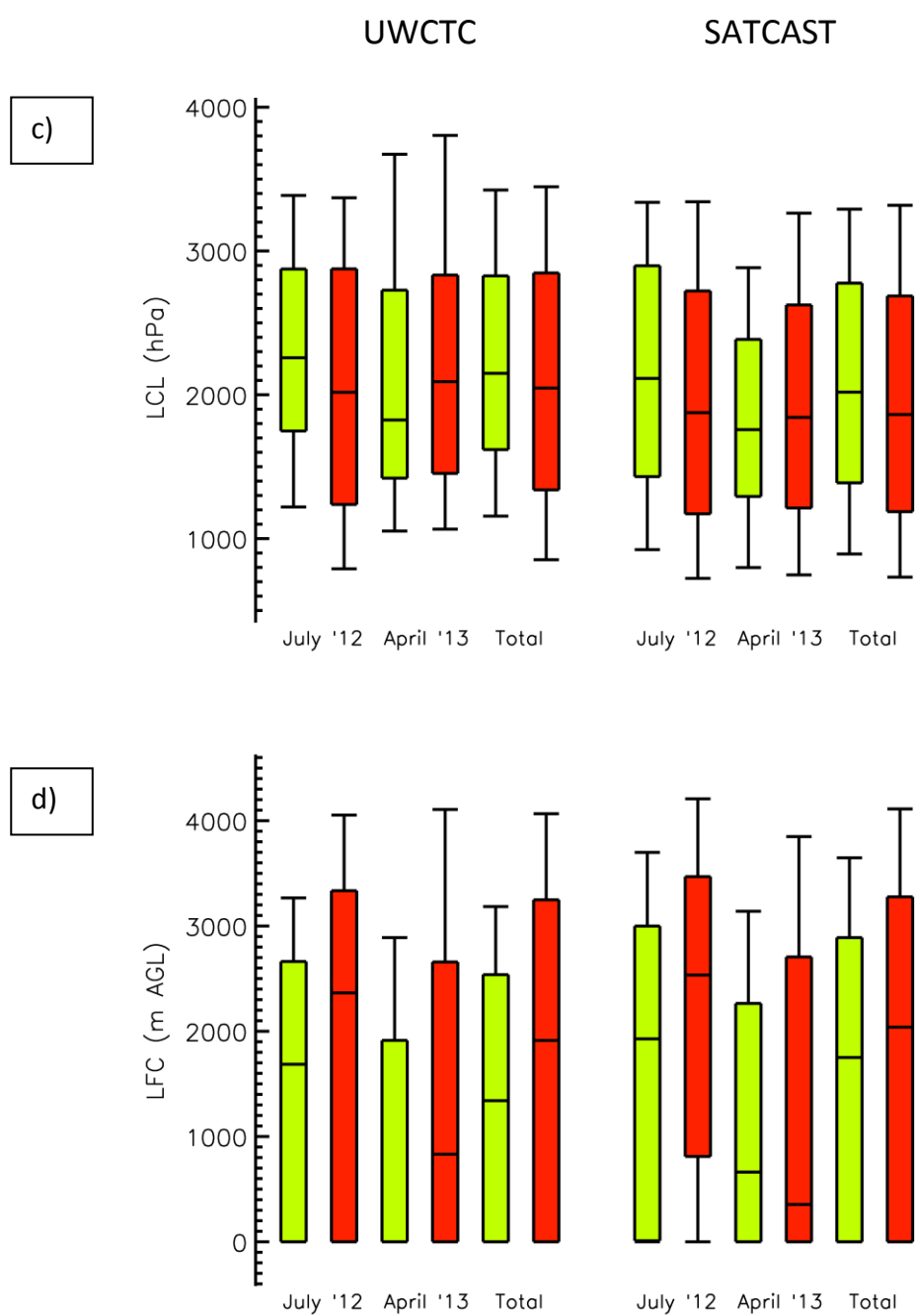
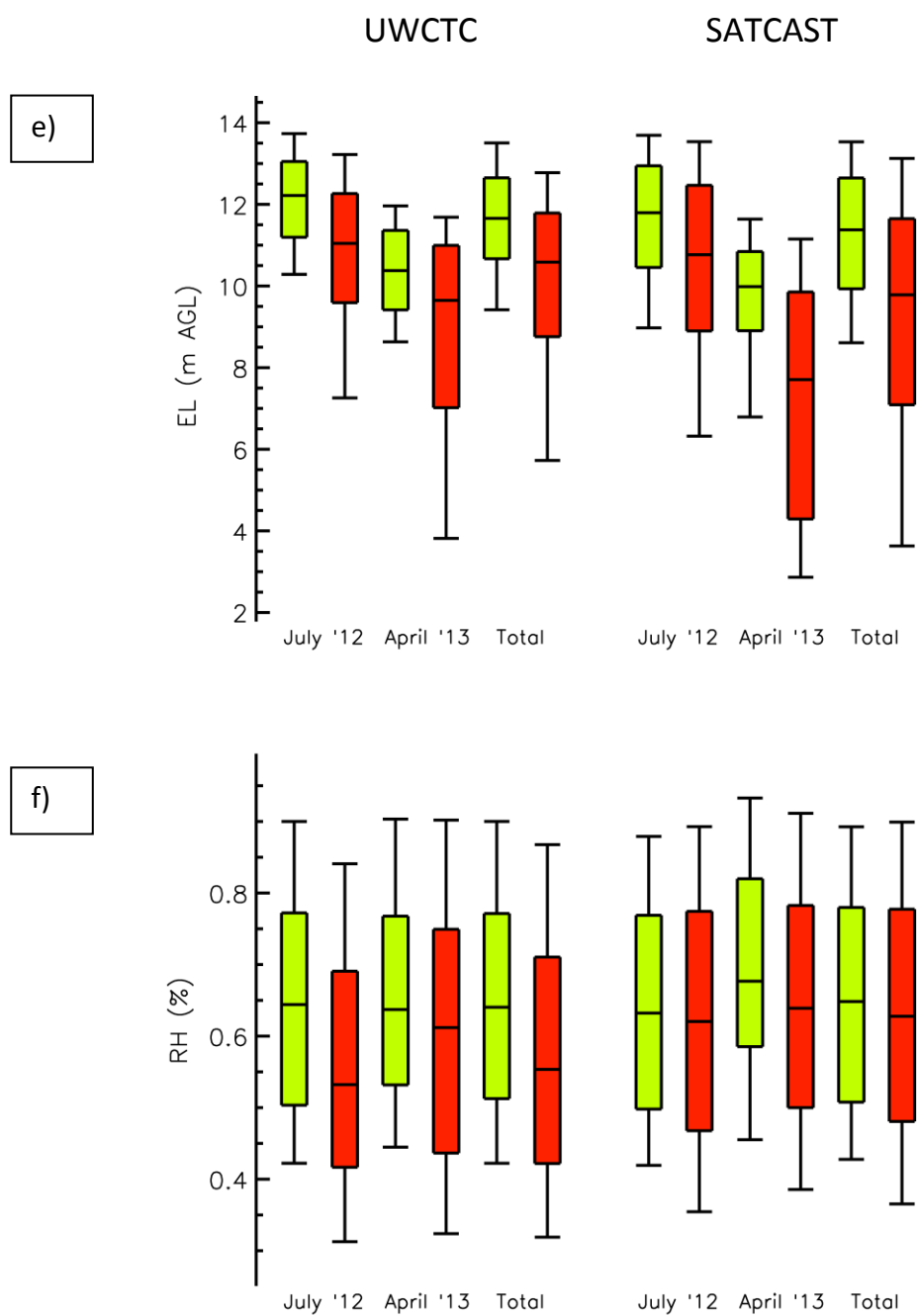
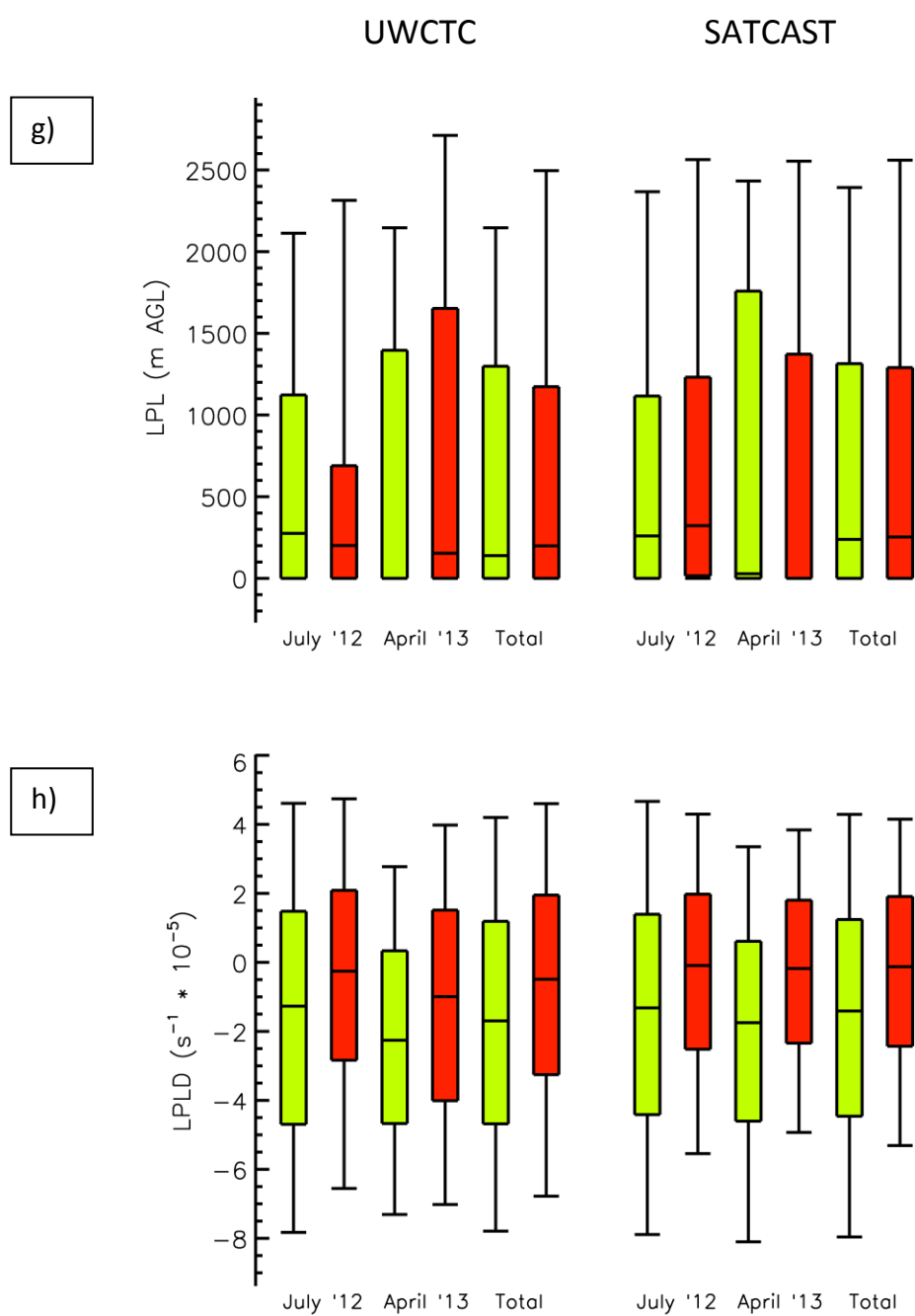
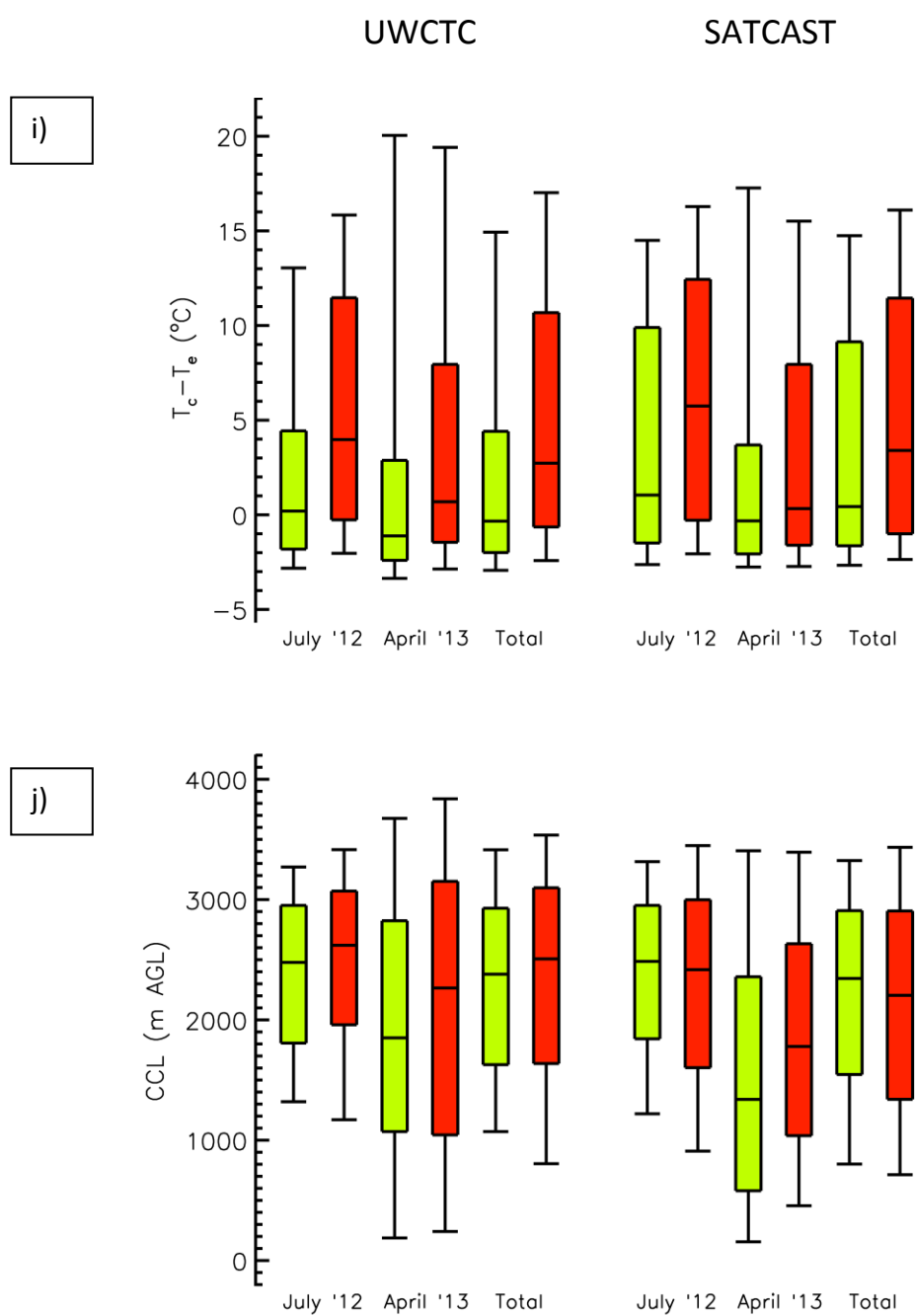


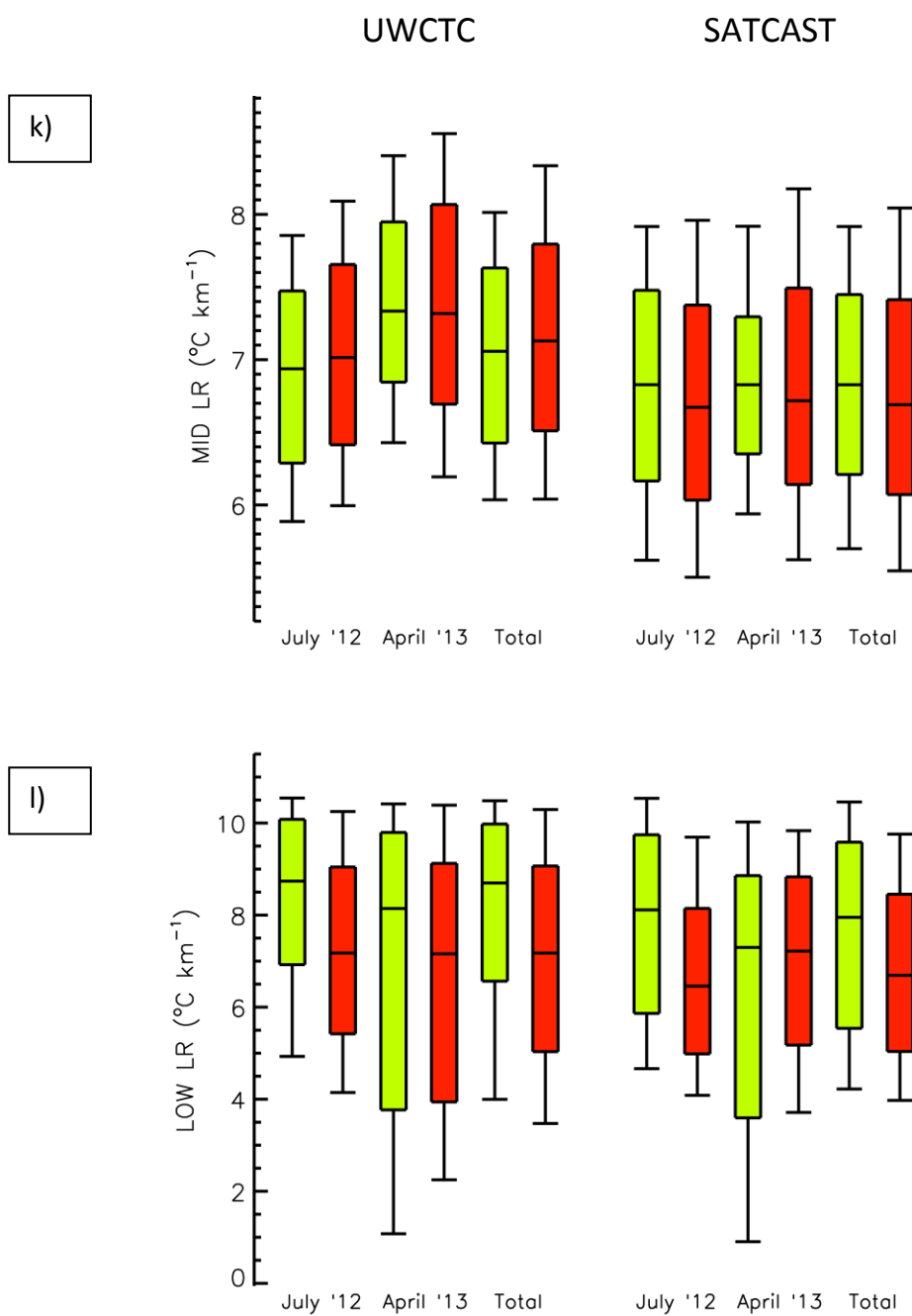
Figure 5.5. July, April and total UWCTC (left) and SATCAST (right) study period RAP values for positive (green) and negative (red) indications. Boxes represent the 25th to 75th percentiles and whiskers are from the 10th to 90th percentiles. Medians are shown as solid lines within the boxes.

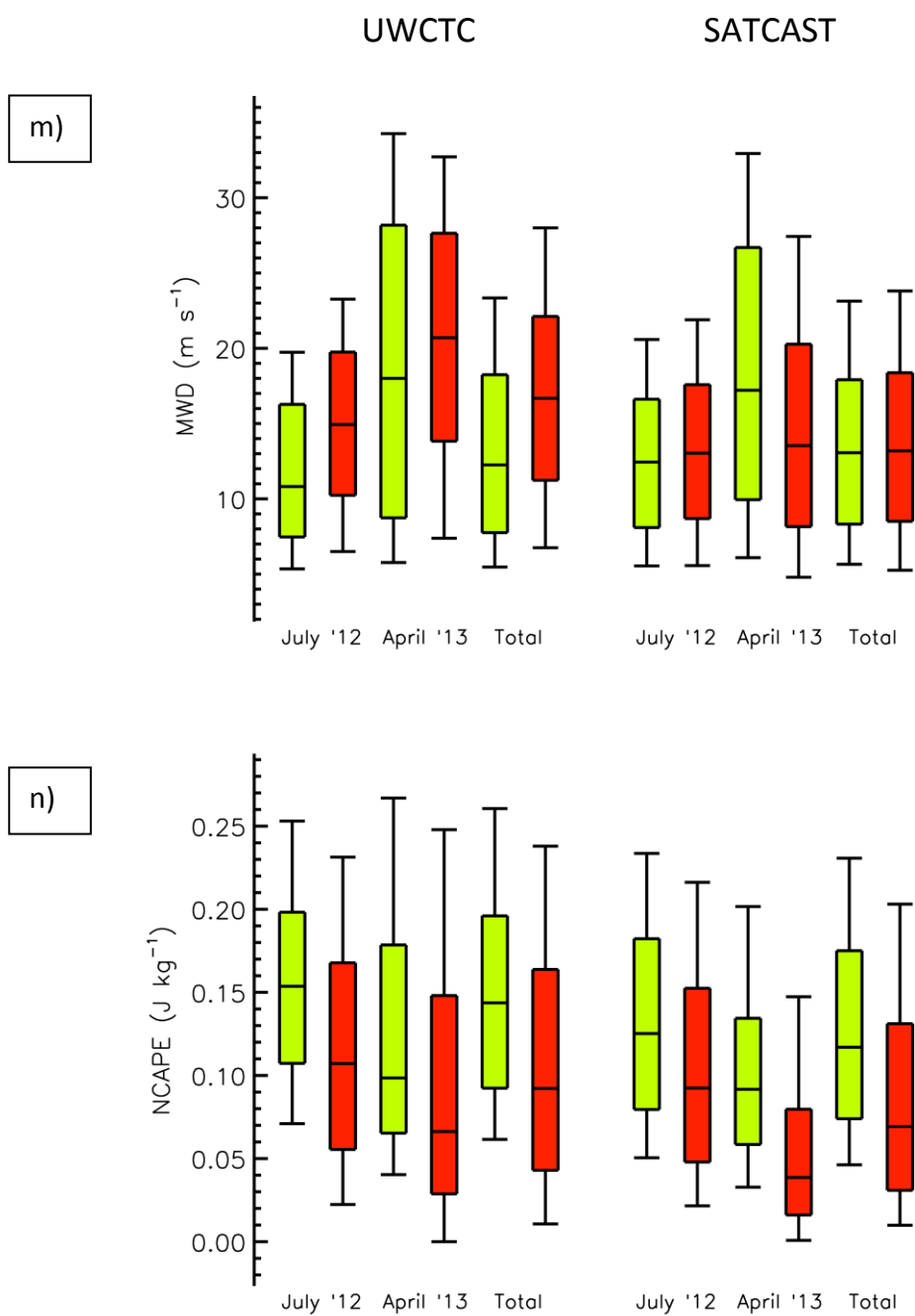
Figure 5.5. *continued.*

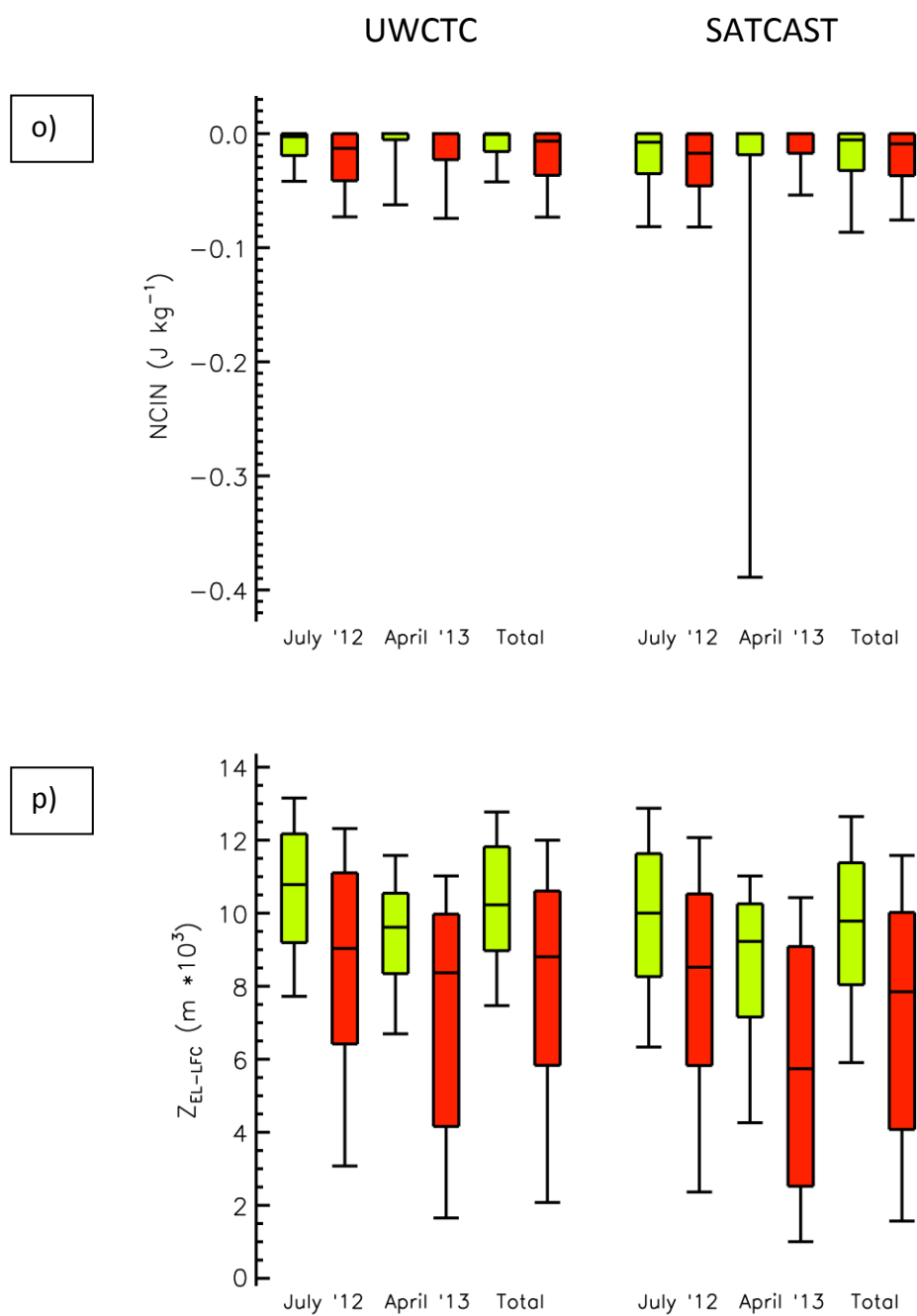
Figure 5.5. *continued.*

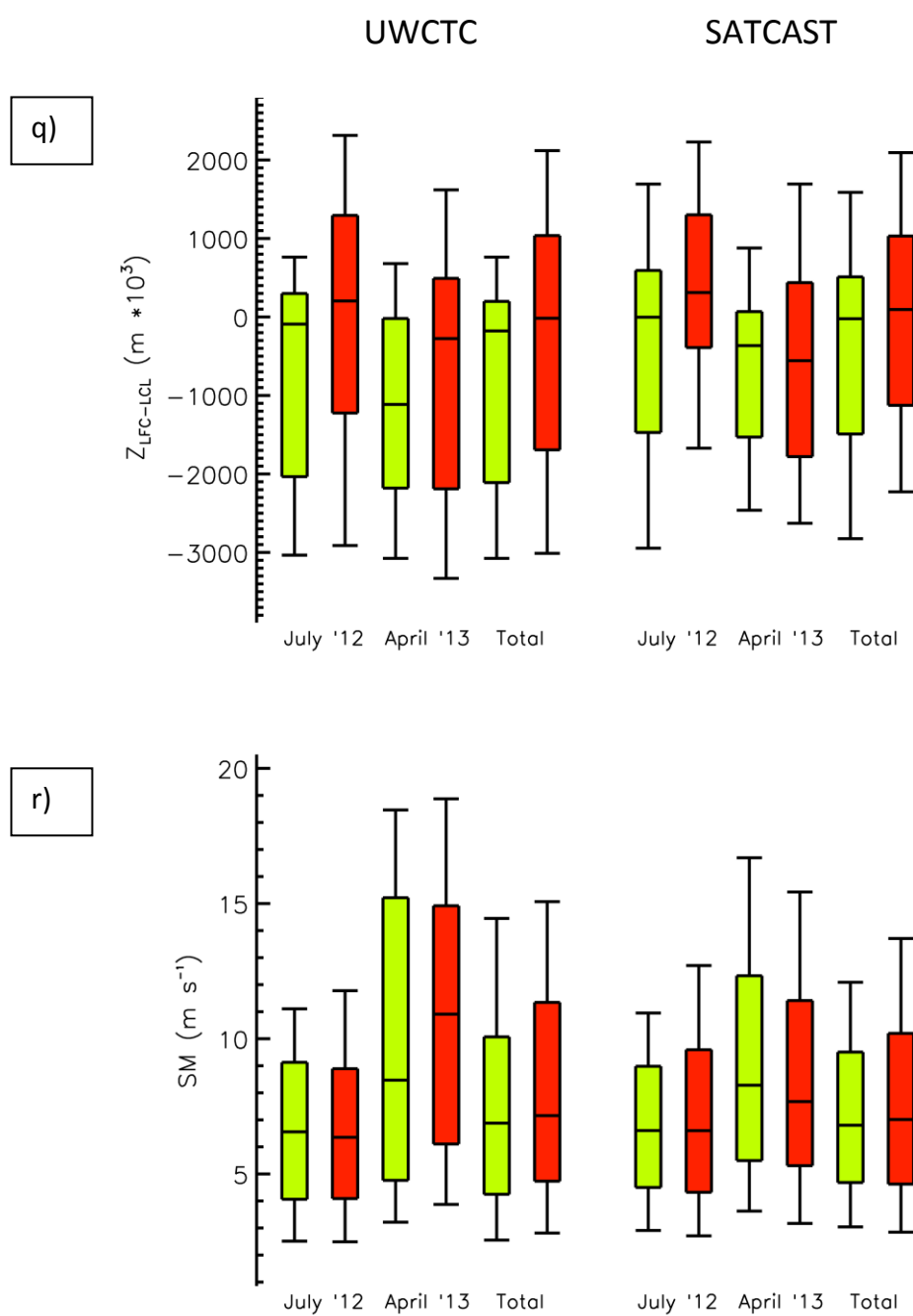
Figure 5.5. *continued.*

Figure 5.5. *continued.*

Figure 5.5. *continued.*

Figure 5.5. *continued.*

Figure 5.5. *continued.*

Figure 5.5. *continued.*

being relatively similar (no significant difference should be found). If no significant difference is found, then the variables are deemed not important to QDA, and potentially not important to overall algorithm performance.

SATCAST RAP environmental variable results have similar characteristics to UWCTC values (Figure 5.5). Instability-related factors (CAPE, EL, Z_{EL-LFC}), all show large differences (Figure 5.5 a, e, p) in a manner consistent with the PCA findings (i.e. more instability, higher likelihood of a positive detection). Other variables exhibit different tendencies depending on the time period, such as LFC, LCL, LPL, CCL, MWD and SM (Figure 5.5 c, d, g, j, m, r). Variables that have differing relationships between time periods for positive and false groups result in little change when considered as a total dataset. For example, CCL is larger for positive indications than false indications in the July study period for SATCAST (Figure 5.5 j). In the April study period for SATCAST, positive indications have a much lower CCL than false indications. The resulting total dataset exhibits a difference that is not statistically significant. Other data do not exhibit large differences between positive and false SATCAST indications for both time periods, such as RH and SM (Figure 5.5 f and r).

The total UWCTC dataset had several statistically significant differences between positive and false indications in both ANOVA and MW tests (Table 5.5). ANOVA and MW suggest a lack of significance in LCL, LPL and CCL. MW also suggests that MIDLR and SM differences are not statistically significant. ANCOVA is more aggressive by additionally removing CIN, LPLD and T_c-T_e . The aggressive reduction in

Table 5.5. P-values for UWCTC tests for significant difference between positive and false indications for the total dataset (combined July and April). Variables with 99.5% significance are shown in bold, stars indicate that the variable was omitted from the test.

	ANOVA	MW	ANCOVA
CAPE	0.000	0.000	0.000
CIN	0.000	0.000	0.016
LCL	0.020	0.008	*
LFC	0.000	0.000	*
EL	0.000	0.000	*
RH	0.000	0.000	0.002
LPL	0.985	0.455	0.281
LPLD	0.001	0.000	0.068
T_c-T_e	0.000	0.000	0.011
CCL	0.830	0.042	0.001
MID LR	0.000	0.009	0.000
LOW LR	0.000	0.000	0.445
MWD	0.000	0.000	0.000
NCAPE	0.000	0.000	0.237
NCIN	0.000	0.000	0.632
Z_{EL-LFC}	0.000	0.000	0.736
$Z_{LFC-LCL}$	0.000	0.000	0.128
SM	0.002	0.014	0.942

variable importance by ANCOVA can be attributed to the high correlation between certain variables in the dataset.

Fewer variables are found to not have statistically significant differences between SATCAST positive and false indications (Table 5.6). LPL, MWD and NCIN are all found not statistically significant in ANOVA, and MW keeps NCIN. ANCOVA again is aggressive in removing variables from the QDA with CIN, T_c-T_e , CCL, MIDLR and Z_{EL-LFC} . All three tests show several more significant variable differences for the SATCAST dataset than the UWCTC dataset. It is possible that statistical differences are caused by the object identification schemes used by each algorithm. UWCTC is a fairly conservative algorithm which only produces indications if a cloud object is deemed vertically cooling. The SATCAST identification scheme identifies all cloud objects rather than only vertically cooling areas, which produces more indications in non-convective areas. These indications, while normally low in SoS, cause many significant differences between the convective variables selected for this study. A QDA approach will thus be helpful to remove such points from the overall product to assist the operational forecaster.

Splitting UWCTC into individual study periods implies that different variables are statistically significant at different times of the year (Table 5.7). ANOVA and MW agree that MIDLR and SM are not significant for the July study period. A considerably larger amount of variables are deemed not significant during the April study period, most notably in $Z_{LFC-LCL}$ (Table 5.8). The lack of $Z_{LFC-LCL}$ importance can be explained through

Table 5.6. P-values for SATCAST tests for significant difference between positive and false indications for the total dataset (combined July and April). Variables with 99.5% significance are shown in bold.

	ANOVA	MW	ANCOVA
CAPE	0.000	0.000	0.000
CIN	0.000	0.000	0.840
LCL	0.000	0.000	*
LFC	0.000	0.000	*
EL	0.000	0.000	*
RH	0.000	0.000	0.000
LPL	0.064	0.108	0.000
LPLD	0.000	0.000	0.000
T_c-T_e	0.000	0.000	0.330
CCL	0.000	0.000	0.139
MID LR	0.000	0.000	0.003
LOW LR	0.000	0.000	0.000
MWD	0.213	0.057	0.000
NCAPE	0.000	0.000	0.000
NCIN	0.075	0.000	0.000
Z_{EL-LFC}	0.000	0.000	0.047
$Z_{LFC-LCL}$	0.000	0.000	0.000
SM	0.000	0.000	0.000

Table 5.7. P-values for UWCTC tests for significant difference between positive and false indications for the July study period. Variables with 99.5% significance are shown in bold.

	ANOVA	MW	ANCOVA
CAPE	0.000	0.000	0.000
CIN	0.000	0.000	0.000
LCL	0.011	0.000	*
LFC	0.000	0.000	*
EL	0.000	0.000	*
RH	0.000	0.000	0.000
LPL	0.000	0.118	0.989
LPLD	0.000	0.000	1.000
T _c -T _e	0.000	0.000	1.000
CCL	0.004	0.041	0.044
MID LR	0.875	0.011	0.135
LOW LR	0.974	0.000	1.000
MWD	0.000	0.000	0.000
NCAPE	0.000	0.000	0.998
NCIN	0.000	0.000	0.008
Z _{EL-LFC}	0.000	0.000	0.993
Z _{LFC-LCL}	0.000	0.000	0.785
SM	0.065	0.353	0.000

Table 5.8. P-values for UWCTC tests for significant difference between positive and false indications for the April study period. Variables with 99.5% significance are shown in bold.

	ANOVA	MW	ANCOVA
CAPE	0.000	0.000	0.020
CIN	0.000	0.021	0.567
LCL	0.250	0.126	*
LFC	0.008	0.021	*
EL	0.000	0.000	*
RH	0.007	0.022	0.199
LPL	0.052	0.075	0.905
LPLD	0.002	0.005	0.003
T_c-T_e	0.108	0.000	0.250
CCL	0.406	0.085	0.110
MID LR	0.920	0.455	0.384
LOW LR	0.449	0.114	0.733
MWD	0.011	0.030	0.474
NCAPE	0.000	0.000	0.338
NCIN	0.009	0.020	0.559
Z_{EL-LFC}	0.000	0.000	0.966
$Z_{LFC-LCL}$	0.029	0.031	0.987
SM	0.057	0.041	0.002

examination of the PCA (Figure 5.2). April data exhibit evidence of more forced convection, which is less dependent on instability than convection is during July.

SATCAST has a number of significant variables for July (Table 5.9). Again, this significance is attributed to less conservative convective cloud object identification. ANOVA found all variables significant in the dataset, and MW only removes SM. Even ANCOVA, the most conservative of the tests, only removes LPL, MWD, NCIN and SM. Thus during July, most of the variables selected for this study have some importance to the performance to SATCAST. Changing statistical differences between both algorithms for two different time periods implies that different relationships could be more important at different times of the year. Examining seasonality changes is beyond the scope of this writing; however future studies with larger datasets can explore seasonal relationships further.

SATCAST shows a similar lack of dependence to UWCTC on $Z_{LFC-LCL}$ in April (Table 5.10). Since both algorithms found that $Z_{LFC-LCL}$ is not as crucial to the success of an indication, despite the fact that $Z_{LFC-LCL}$ is a large source of variation, use of CIN is a better choice to understand the effects of a capping inversion on CTC-based algorithm performance. When examining $Z_{LFC-LCL}$ which only pertains to depth of the stable layer, it is found that the value is comparable to CIN (Figure 5.6). Statistically, the two datasets both exhibit significant differences. CIN exhibits slightly larger differences. Since the two are considered virtually the same measure, only CIN is used in QDA. Now that statistically significant variables have been identified, the total study periods can be applied to future case studies with QDA.

Table 5.9. P-values for SATCAST tests for significant difference between positive and false indications for the July study period. Variables with 99.5% significance are shown in bold.

	ANOVA	MW	ANCOVA
CAPE	0.000	0.000	0.000
CIN	0.000	0.000	0.000
LCL	0.000	0.000	*
LFC	0.000	0.000	*
EL	0.000	0.000	*
RH	0.000	0.000	0.000
LPL	0.000	0.000	0.294
LPLD	0.000	0.000	0.000
T_c-T_e	0.000	0.000	0.000
CCL	0.000	0.000	0.000
MID LR	0.000	0.000	0.000
LOW LR	0.000	0.000	0.000
MWD	0.000	0.000	0.991
NCAPE	0.000	0.000	0.000
NCIN	0.000	0.000	0.592
Z_{EL-LFC}	0.000	0.000	0.000
$Z_{LFC-LCL}$	0.000	0.000	0.000
SM	0.000	0.005	0.006

Table 5.10. P-values for SATCAST tests for significant difference between positive and false indications for the April study period. Variables with 99.5% significance are shown in bold.

	ANOVA	MW	ANCOVA
CAPE	0.000	0.000	0.000
CIN	0.000	0.000	0.000
LCL	0.000	0.003	*
LFC	0.055	0.024	*
EL	0.000	0.000	*
RH	0.000	0.000	0.029
LPL	0.000	0.002	0.000
LPLD	0.000	0.000	0.000
T_c-T_e	0.000	0.000	0.000
CCL	0.000	0.000	0.000
MID LR	0.007	0.001	0.000
LOW LR	0.000	0.000	0.000
MWD	0.000	0.000	0.000
NCAPE	0.000	0.000	0.002
NCIN	0.004	0.107	0.000
Z_{EL-LFC}	0.000	0.000	0.000
$Z_{LFC-LCL}$	0.066	0.246	0.000
SM	0.000	0.000	0.000

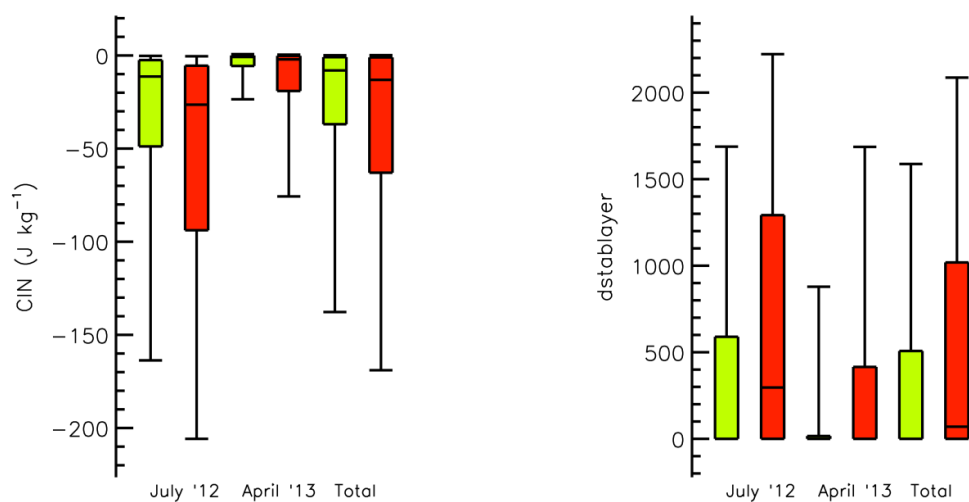


Figure 5.6. July, April and total SATCAST study period RAP values for positive (green) and negative (red) indications. Boxes represent the 25th to 75th percentiles and whiskers are from the 10th to 90th percentiles. Medians are shown as solid lines within the boxes.

c. Case Studies

Examination of the performance of ANOVA, ANCOVA and MW tests are analyzed on the second case study, 20 May 2013. Application of the QDA to UWCTC on 20 May 2013 using significant ANOVA variables for the total dataset was found to make the ratio of positive to false indications worse (Table 5.11). That is, the combination of the statistically significant differences found by the ANOVA approach did not result in a better CI forecasting product. MW tests, however, provide consistent improvement to the algorithm and conform to large differences seen in the box plots. ANCOVA issues with QDA stem from linear relationships between variables deemed significant (such as CAPE to NCAPE), which normally result in one being deemed significant and the other removed. Linear dependence may be useful for other types of analyses. However removing environmental variables removes additional information that can be gained from RAP which improves QDA on case studies (such as removing Z_{EL-LFC} in favor of CAPE). MW significant variables applied through QDA improve the 20 May 2013 case (Table 5.11).

ANOVA and MW on the 20 May 2013 case improved the SATCAST product (Table 5.12). ANCOVA again seemed to remove too much information and did less to improve the overall SATCAST algorithm. In all types of analyses, QDA will not only remove false alarms. That is, the QDA would consider some positive indications as false depending on their environment. QDA issues can stem from data collection methodology problems to issues with the RAP model in fully resolving indication environments. Despite the loss of positive indications, more false indications are correctly discarded and

Table 5.11. 20 May 2013 UWCTC performance of QDA removing insignificant variables diagnosed by ANOVA, ANCOVA and MW.

		Predicted		
		False	Positive	Total
Actual	False	22	72	94
	Positive	26	51	77
	Total	48	123	171

Percentage of Improvement: -10.36%

		Predicted		
		False	Positive	Total
Actual	False	56	38	94
	Positive	42	35	77
	Total	98	73	171

Percentage of Improvement: 5.03%

		Predicted		
		False	Positive	Total
Actual	False	15	79	94
	Positive	7	70	77
	Total	22	149	171

Percentage of Improvement: 6.87%

Table 5.12. 20 May 2013 SATCAST performance of QDA removing insignificant variables diagnosed by ANOVA, ANCOVA and MW.

Actual		SATCAST ANOVA		
		Predicted		
		False	Positive	Total
False	False	9758	21118	30876
Positive	Positive	296	804	1100
Total	Total	10054	21922	31976

Percentage of Improvement: 4.69%

Actual		SATCAST TOTAL ANCOVA		
		Predicted		
		False	Positive	Total
False	False	18633	12243	30876
Positive	Positive	645	455	1100
Total	Total	19278	12698	31976

Percentage of Improvement: 1.71%

Actual		SATCAST MW		
		Predicted		
		False	Positive	Total
False	False	9845	21031	30876
Positive	Positive	302	798	1100
Total	Total	10147	21829	31976

Percentage of Improvement: 4.43%

thus types of QDA applied to UWCTC and SATCAST are useful (shown through percentage of improvement). Examining all three approaches with respect to percentage of improvement reveals that the choice for an optimal test is relatively ambiguous (Table 5.13). Through all four cases, MW was consistently helpful for both algorithm products. Therefore, MW was the approach selected to analyze QDA performance for the case studies.

The QDA approach improved both UWCTC and SATCAST for the first case study, 9 April 2013 (Table 5.14). Important indications are kept in central Kansas which result in CI ~1 hour later (Figure 5.7). Several false indications with relatively high SoS values are removed from areas such as eastern Kansas. The resulting POI value near 14% improves the Brier score values for the dataset. UWCTC saw a notably smaller improvement with POI near 1%. Several positive indications are lost in the 9 April 2013 case for the UWCTC dataset behind the cold front (Figure 5.8). Indications are in regions containing little instability, and thus are deemed environments not conducive of positive indications.

An examination using MW on the 20 May 2013 case study finds several false indications correctly removed from the SATCAST and UWCTC datasets (Figures 5.9 and 5.10). Both SATCAST and UWCTC show impressive lead time in forecasting the strong updraft on the storm that produced the Moore, Oklahoma tornado. These values are correctly kept in all analyses, and false indications are correctly removed around central NE. The QDA successfully keeps what is probably the strongest of updrafts analyzed in

Table 5.13. UWCTC and SATCAST percentage of improvement values for all case studies using all three statistical approach solutions for QDA.

	9 April 2013	20 May 2013	20-22 June 2013	25 June 2013
UWCTC				
ANOVA	1.55%	-10.36%	16.60%	37.07%
ANCOVA	-10.61%	5.03%	22.51%	16.34%
MW	1.08%	6.87%	14.44%	35.85%
SATCAST				
ANOVA	16.51%	4.69%	19.06%	30.67%
ANCOVA	13.16%	1.71%	15.69%	28.11%
MW	13.78%	4.43%	18.36%	30.96%

Table 5.14. 9 April 2013 confusion matrices for UWCTC and SATCAST performance of QDA removing insignificant variables diagnosed by MW.

		Predicted		
		False	Positive	Total
Actual	False	88	34	122
	Positive	27	11	38
	Total	115	45	160

Percentage of Improvement: 1.08%

		Predicted		
		False	Positive	Total
Actual	False	9154	9546	18700
	Positive	153	282	435
	Total	9307	9828	19135

Percentage of Improvement: 13.78%

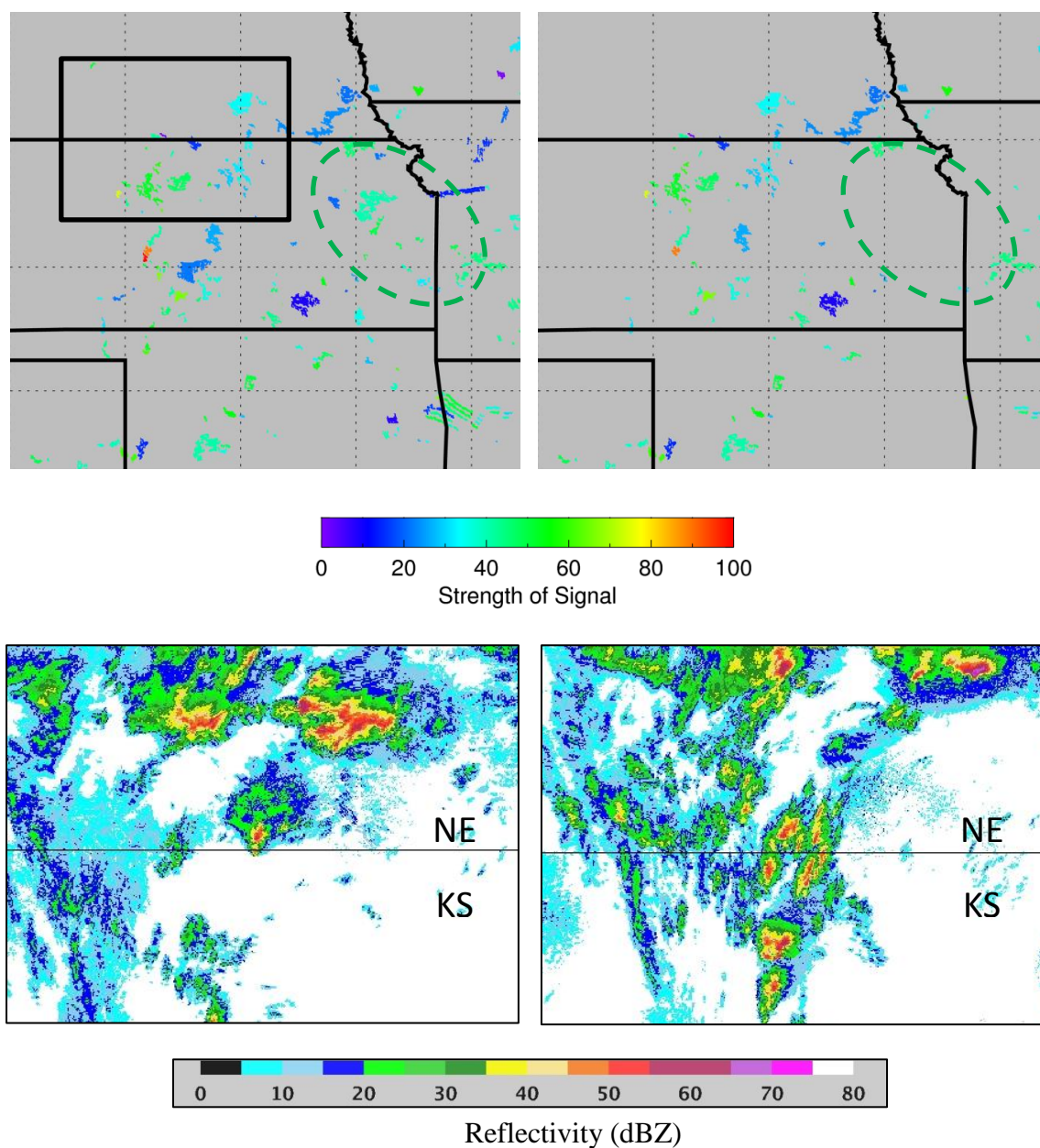


Figure 5.7. 1832 UTC 9 April 2013 unfiltered SATCAST (top left) and QDA-filtered SATCAST (top right). The green circle displays an area where QDA successfully removes non convective SATCAST indications. 1835 UTC 9 April 2013 composite NMQ radar reflectivity data (dBZ) inside the rectangle (bottom left) and 1935 UTC 9 April composite radar reflectivity (dBZ) displaying several CI events (>35 dBZ) successfully forecast by SATCAST (bottom right).

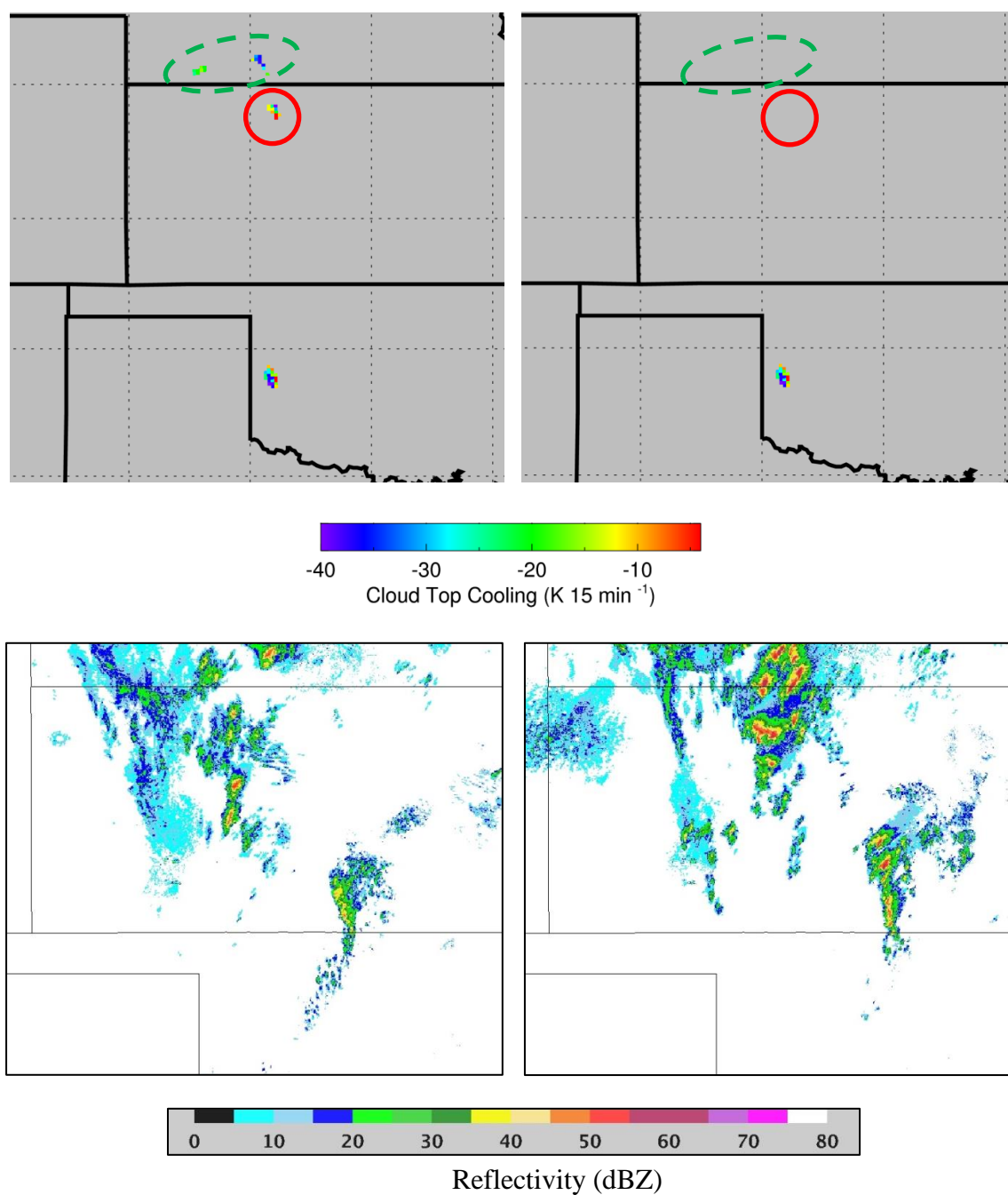


Figure 5.8. 1855 UTC 9 April 2013 UWCTC product with one positive detection in northern Kansas and 3 false detections (top left) with a QDA filter applied to the product (top right). 1855 UTC 9 April 2013 NMQ reflectivity data (dBZ) (bottom left) and one hour later (bottom right). Two false detections are correctly dropped (green circle) at the cost of one positive (red circle).

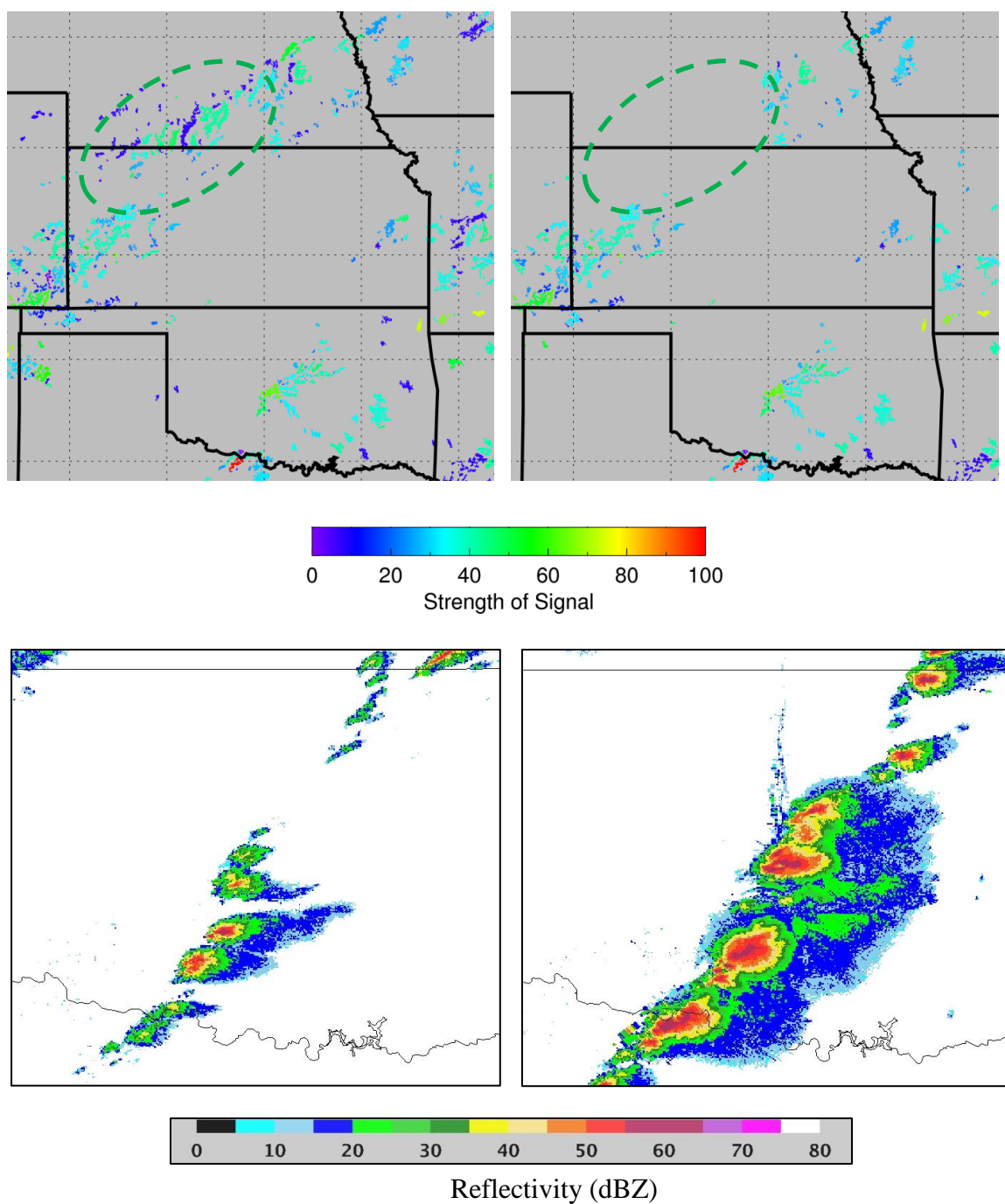


Figure 5.9. 1825 UTC 20 May 2013 unfiltered (top left) and QDA-filtered (top right) SATCAST data. The green circle highlights an area where false detections are correctly removed. 1910 UTC 20 May 2013 NMQ reflectivity data (dBZ) over central Oklahoma (bottom left) and one hour later (bottom right).

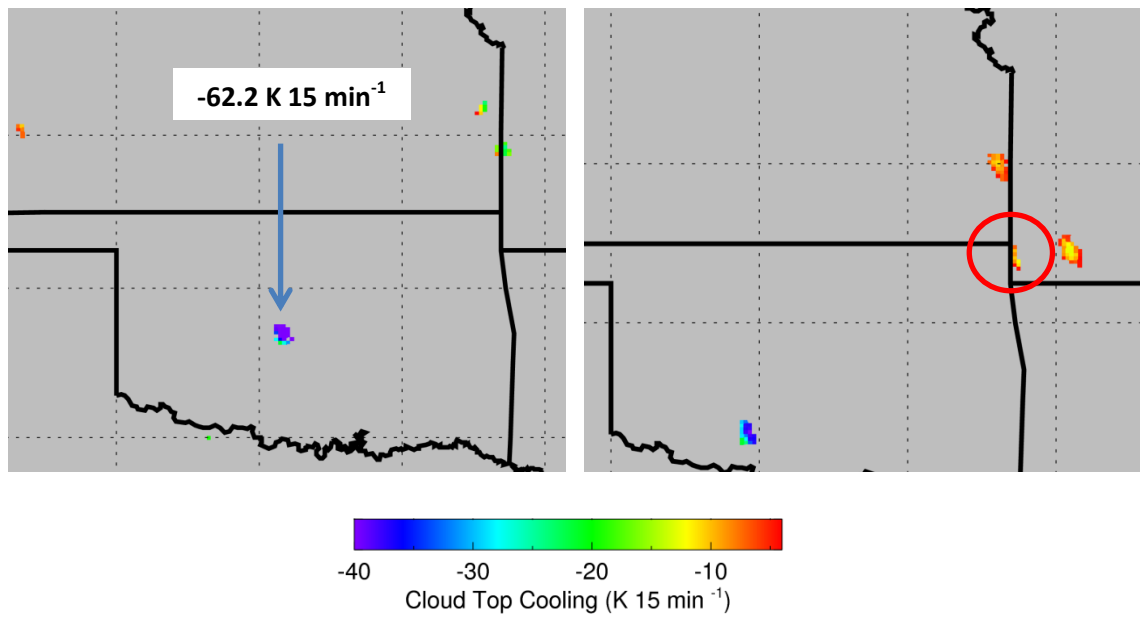


Figure 5.10. 1910 UTC 20 May 2013 UWCTC product with QDA filtering and the maximum value of measured cooling for the Moore, Oklahoma, tornadic supercell (left) and 1815 UTC 20 May 2013 example of correctly QDA-filtered UWCTC product circled in red (right). For radar refer to Figure 5.9.

this study with the Moore, Oklahoma cell, where cooling values $\sim 62 \text{ K } 15 \text{ min}^{-1}$ are witnessed with an apparent bounded weak echo region on the composite NMQ dataset.

A notably larger improvement in the UWCTC dataset using QDA is witnessed in the 20 June – 22 June 2013 case studies, where only $\sim 19\%$ of positive indications were considered false by the filtering process. The QDA removed $\sim 33\%$ of false indications, resulting in a $\sim 14.4\%$ improvement (Table 5.15). SATCAST displayed considerable improvement as well (Figure 5.10). However, it is found that several false indications are left unfiltered near the area of interest. These indications contain relatively large SoS values and if left unchecked may cause confusion to a nowcaster using such products. Combined use of the SATCAST and UWCTC datasets in this case is found to be helpful, with both containing CI indications in northern Nebraska resulting in several CI events (Figures 5.11 and 5.12). The lack of filtering on several large SATCAST indications suggests that while a QDA does arguably improve the satellite-based CTC forecasting dataset, the filtering method does not remove all false indications. A forecaster therefore should still use this product coupled with several other observational tools (surface data, soundings, high resolution models) to correctly nowcast CI in specific areas. A QDA-filtered dataset should therefore be considered only for decision support.

25 June 2013 showed some of the largest percentage of improvement values for both UWCTC and SATCAST (Table 5.16). Several false indications are correctly removed from SATCAST, which correctly detects the small nocturnal line of CI in central Nebraska (Figure 5.13). While UWCTC did not pick up this line, indications for several additional events are found in central IA (Figure 5.14). QDA successfully filters

Table 5.15. 20 June 2013 – 22 June 2013 confusion matrices for UWCTC and SATCAST performance of QDA removing insignificant variables diagnosed by MW.

		UWCTC		
		Predicted		
Actual		False	Positive	Total
	False	133	272	405
	Positive	23	102	125
	Total	156	374	530

Percentage of Improvement: 14.44%

		SATCAST		
		Predicted		
Actual		False	Positive	Total
	False	18845	35474	54319
	Positive	74	379	453
	Total	18919	35853	54772

Percentage of Improvement: 18.36%

Table 5.16. 25 June 2013 confusion matrices for UWCTC and SATCAST performance of QDA removing insignificant variables diagnosed by MW.

		UWCTC		
		Predicted		
Actual		False	Positive	Total
	False	54	28	82
	Positive	3	7	10
	Total	57	35	92

Percentage of Improvement: 35.85%

		SATCAST		
		Predicted		
Actual		False	Positive	Total
	False	10352	9239	19591
	Positive	28	100	128
	Total	10380	9339	19719

Percentage of Improvement: 30.96%

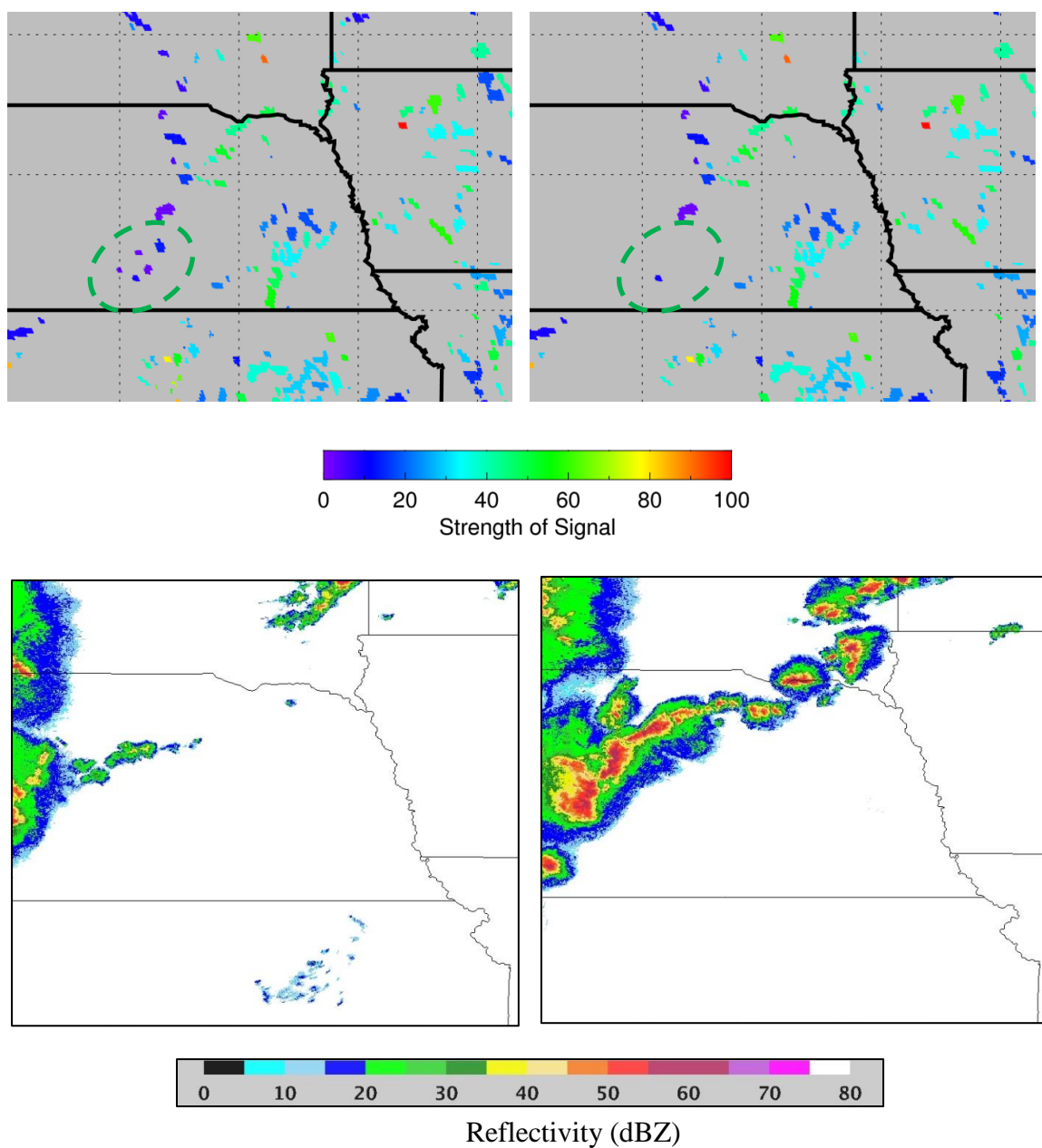


Figure 5.11. 0055 UTC 22 June 2013 unfiltered (top left) and QDA-filtered (top right) SATCAST data. The green circle highlights an area where false detections are correctly removed. 0125 UTC 22 June 2013 NMQ reflectivity data (dBZ) over central Nebraska (bottom left) and one hour later (bottom right).

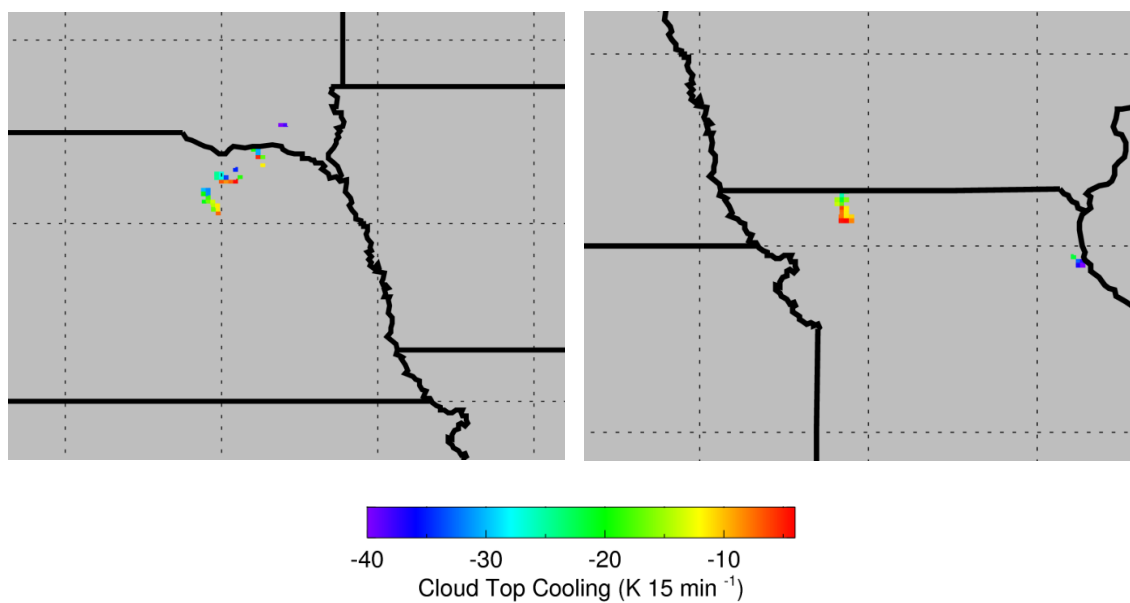


Figure 5.12. 0125 UTC 22 June 2013 QDA-filtered UWCTC product over northern Nebraska (left) and 0155 UTC 22 June 2013 example of two correctly removed products in the non-convective area of northern Missouri (right).

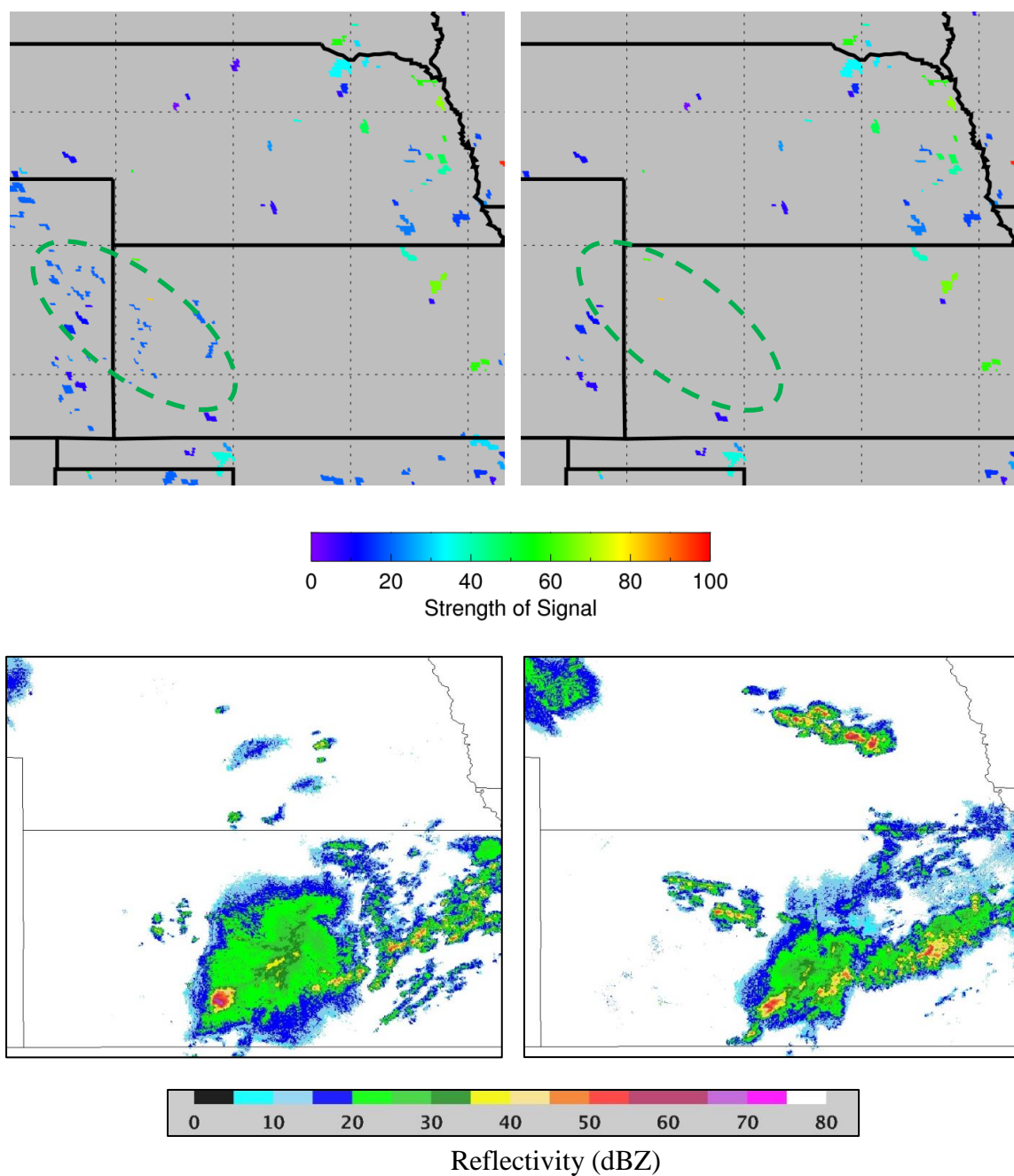


Figure 5.13. 0215 UTC 25 June 2013 unfiltered (top left) and QDA-filtered (top right) SATCAST data. The green circle highlights an area where false detections are correctly removed. 0215 UTC 22 June 2013 NMQ reflectivity data (dBZ) over southern Nebraska (bottom left) and one hour later (bottom right).

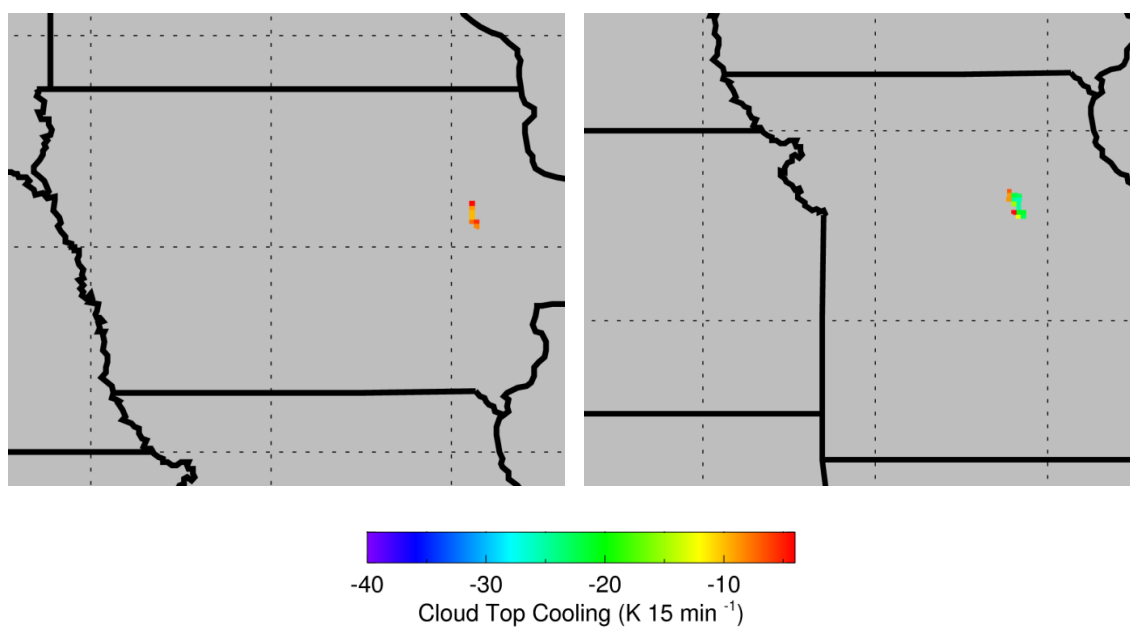


Figure 5.14. 0815 UTC 25 June 2013 QDA-filtered UWCTC product over Iowa (left) and 0945 UTC 25 June 2013 example of a UWCTC removed value.

two separate indications, keeping a positive yet relatively weak ($\sim 7 \text{ K } 15 \text{ min}^{-1}$) CI indication in central IA and discarding a moderate ($\sim 14 \text{ K } 15 \text{ min}^{-1}$) yet false indication in central MO.

Overall, simple application of QDA is found to improve all case studies examined herein. While it does not remove all false indications from the two algorithms, it does show that it is possible to successfully apply NWP variables to a CTC-based algorithm to improve the skill of the overall product. Therefore, probability-based filters should be used with unfiltered products and other datasets to improve the situational awareness of a CI nowcaster.

6. Summary and Conclusion

Convection initiation (CI) can be determined using satellite-based brightness temperature and interest field trends. UWCTC and SATCAST algorithms both make use of differing convection initiation forecasting techniques which will utilize the advanced baseline imager on board the GOES-R series. Both algorithms are being extensively tested before they become operational after the launch. While several studies have examined algorithm performance and relation to severe weather occurrence, only a few have speculated on the effects of unfavorable pre-convective environments. This study therefore set out to answer three questions: Do unfavorable pre-convective environments affect algorithm performance? Can these environmental variables be used to improve future algorithm products? Will the incorporation of these products be helpful to an operational forecaster? To answer all three questions, an objective validation study was conducted to create a large dataset of positive and false indications for both algorithms. Positive and false indications were examined with rapid refresh-resolved pre-convective environmental variables using a principal components analysis, and significant differences between both groups were determined. Environmental variables that were significantly different between groups were used to filter five case studies using a quadratic discriminant analysis approach.

Objective validation took place in several steps. Algorithm products are remapped to one grid with NMQ radar data to achieve the validation. NMQ radar data are segmented using the WDSS-II framework, and the centroids of the segmented clusters are tracked through space and time using ThOR. Tracking allows for the

identification of the first occurrence of a segmented WDSS-II cluster, also known as the point of convection initiation. WDSS-II requires both spatial and quantitative reflectivity thresholds to create clusters analyzed by ThOR. WDSS-II also implements a smoothing technique for faster and more accurate calculation of segmented centroids that are not affected by spurious peaks. In order to optimize the combination of WDSS-II and ThOR CI identification techniques, a hand validation is performed over a small area during a highly convective 4-hour period. The hand validation identified the location of actual CI events (truth clusters) and compared the events to various WDSS-II smoothing and spatial scale settings. The WDSS-II settings that best matched the clusters identified by a forecaster are used.

Objective validation is heavily dependent upon spatial overlap of an indication area with a CI truth cluster. Spatial decorrelation between algorithm data and radar data is reasonable and expected with errors caused by parallax and variations between two different measurement techniques. Since parallax error is approximately accounted for, all other issues are resolved using a comparison of hand-identified positive indications to ThOR-identified CI centroids. The average distance between centroids is compared to the average spatial coverage of SATCAST and UWCTC, and a generous ‘forgiveness region’ is applied for optimal objective validation.

The objective validation results in large positive and false samples of SATCAST and UWCTC indications. Study periods are taken from the late convective season in 2012 and early convective season in 2013 over the Great Plains. Study periods depended upon data availability from GOES, NMQ and RAP data sources. A principal components

analysis was completed on the data to examine the sources of variation in the entire dataset. Examining data plotted in terms of the top two sources of variation (related to $Z_{LFC-LCL}$ and CAPE) suggest that a relationship exists between how good an indication forecast convection initiation and the pre-convective environment. PCA found that more positive indications exist in environments conducive of larger-than-average instability and lower-than-average stability.

Significance testing was also performed to further the exploration of relationships between pre-convective environments and the satellite algorithm performances. Larger (smaller) CAPE is found with indications that are positive (false) CI forecasts. The relationship to CAPE could have something to do with the difficulty of identifying areas that are actually vertically cooling rather than falsely-resolved horizontal changes. CAPE is found in truly convective areas, where false areas (identified as objects in SATCAST or vertically cooling pixels in UWCTC) can reside in both unstable and stable regions. Two different interpretations of CIN are examined in this study, one of the negative integrated buoyancy below the LFC, and one as depth of the stable layer visible by satellite. Through analysis of two separate datasets, it was found that the depth of the stable layer visible by satellite had very little difference to measured CIN. Thus CIN alone was determined sufficient for discrimination.

In addition to CAPE, CIN is an important variable to consider for algorithm performance, which is consistent with speculations from Mecikalski et al. (2008) and Walker et al. (2012). It is also found that low-level divergence is a helpful discrimination variable when resolved at the LPL through RAP data. No definitive conclusion can be

reached regarding the impact of shear, lapse rates, convective condensation level, layer relative humidity and lifted condensation level due to varying relationships between time periods. It is possible that all four variables have little impact on the overall performance of either algorithm. Shear had a consistent impact on the UWCTC approach, which may suggest that CTC is more difficult to achieve in highly sheared environments. Shear does not have a consistent impact with SATCAST between time periods, which may simply be a result of the different object identification schemes between the two algorithms. Storm motion is not significant for the likelihood of a positive or false indication. The storm motion problem may stem from an issue regarding missed detections rather than likelihood of an indication performing well. Any issues regarding storm motion are heavily dependent upon the temporal resolution of the satellite imagery and should be examined in further detail after the launch of the GOES-R series.

All significant variables, including CAPE, CIN, divergence and related variables are then used in a probability-based QDA. Use of the QDA consistently improved the performance of both algorithms when the proper variables were removed from the analysis. Even though MW does not remove some of the inconsistent variables discussed previously, the overall algorithm is still improved with the application of the QDA. Thus, it is shown that differences exist in the pre-convective environments for positive and false indications in both algorithms. The application of the statistically significant environmental variables resulted in improvement in both products. Data fusion did not remove several important indications, including the Moore, Oklahoma, cell. Thus a data-

fused product can reduce product noise and give a forecaster more confidence in an indication seen on both unfiltered and QDA-filtered products.

Comparison of radar data to CTC indications can introduce issues stemming from unresolved parallax problems. Spatial decorrelation between radar and satellite data is also possible because the two datasets are measuring two different things, where radar measures the droplet distribution in an area and CTC or cloud-based interest fields measure the area of a cold cloud top temperature. While both issues are accounted for in this study, a better validation in future data periods can take advantage of the GOES-R geostationary lightning mapper (GLM, www.goes-r.gov) to determine locations of convection initiation. The GLM will not suffer from issues stemming from parallax, beam blockage or lack of radar coverage, and much larger domains can be used for validation.

Sieglaff et al. (2013) also suggest that a validation based on the tracking and identification of cloud objects may be a better approach to identifying convective events than radar-based event identification methodology. The advantage of cloud object identification stems from the ability to diagnose correct null forecasts, where a cumulus pixel exists without a CTC indication. The diagnosis of correct null forecasts will allow for the computation of Heidke skill score values. The reason a cloud object approach was not adopted for this work was that the current generation of GOES has a very low temporal resolution (normally 7-15 min, with 30 min data gaps every three hours for full disc scans). The cloud object approach could result in a better objective validation after the launch of GOES-R since temporal resolution will increase to five min.

Several problems occurred when tracking large volume datasets such as SATCAST. The ThOR methodology, to consider all tracks in an acceptable range, had to be tweaked to handle the low temporal resolution of GOES and the high object density of SATCAST. Without correction, ThOR would consider too many candidate tracks with SATCAST objects for reasonable computation. The reduction in first guess search radius allowed ThOR to produce usable tracks with reasonable computational time. However, some tracking error is expected with the reduction in search radius, particularly with fast-moving cloud objects that could be lost by ThOR's tracking scheme. Future studies should not have the same tracking issues with a ~5 min GOES-R temporal resolution.

An issue that may have occurred in the QDA is that convective contamination is not considered in the application of RAP-based variables to new cases. Ongoing convection processes contaminating RAP analysis were not included to new cases in order to simulate the performance of QDA under real time conditions. This way, all products are filtered with the approximate environment they are in. This limitation of RAP data may explain some of the performance variability of the QDA filter on differing case studies. As higher temporal resolution models become operational, such as the high resolution rapid refresh, QDA-filtered data performance should improve in the real time environment. Another issue that may have occurred in QDA is the fact that SATCAST SoS values take into account convection related variables that were not removed by this analysis. Removing variables that are not statistically different from the QDA increased the observed percentage of improvement. Variables that are not statistically different used in the calculation of SoS can mitigate the improvement of QDA. Future products

can take into account the observed statistical relationships in this work to improve output when fusing indications with NWP data. Finally, all derived relationships are found on the current GOES series proving ground products which have a lower temporal resolution than projected for GOES-R. However, it is not the goal of this study to make the perfect QDA. The goal of this study was to show that it is possible to apply specific environmental variables to improve current products, which was successfully done.

Future studies can take advantage of GOES-R data after the launch and can create larger study periods to analyze. While this study did show a difference in pre-convective environmental conditions between positive and false indications for both products, it has not yet been answered why these relationships exist. Hypotheses have been suggested as to good and poor cloud top cooling and interest field based performance with specific environmental conditions. However, tests in a controlled environment are necessary to fully explain the impact of differing variables such as instability and shear on a satellite-based CI forecast. Controlled tests in model runs may yield new pre-convective environmental variables that have not yet been considered to filter algorithm products. This study has thus created the initial framework for future examinations of pre-convective environmental effect on CTC and interest field forecasting. New data-fused products will improve the situational awareness of a CI nowcaster.

7. References

- Banancos, P. and D. Schultz, 2005: The use of moisture flux convergence in forecasting convection initiation: historical and operational perspectives. *Wea. Forecasting*, **20**, 351-366.
- Berendes, T. A., J. R. Mecikalski, W. M. Mackenzie, K. M. Bedka and U. S. Nair, 2008: Convective cloud identification and classification in daytime satellite imagery using standard deviation limited adaptive clustering. *J. Geophys. Res.*, **113**, D20207, doi:10.1029/2008JD010287.
- Browning, K. A. and Coauthors, 2007: The convective storm initiation project. *Bull. Amer. Meteor. Climatol.*, **48**, 1542-1561.
- Cochran, W. G., 1957: Analysis of covariance: Its nature and uses. *Biometrics*, **13**, 261-281.
- Coniglio, M. C., 2012: Verification of RUC 0-1-h forecasts and SPC mesoscale analyses using VORTEX2 soundings. *Wea. Forecasting*, **27**, 667-683.
- Devenyi, D., T. W. Schlatter, S. G. Benjamin and M. Hu, 2007: Hourly data assimilation with gridpoint statistical interpolation for rapid refresh. Preprints, 22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction, Park City, UT, Amer. Meteor. Soc., 4A.2. [Available online at <http://ams.confex.com/ams/pdfpapers/124535.pdf>]
- Fisher, R. A., 1936: The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, **7**, 179-188.

- Hartung, Daniel C., Justin M. Sieglaff, Lee M. Cronic and Wayne F. Feltz, 2013:
An intercomparison of UW cloud-top cooling rates with WSR-88D radar
data. *Wea. Forecasting*, **28**, 463–480.
- Heidinger, A. K., 2010: ABI cloud mask algorithm theoretical basis document.
NOAA/NESDIS/Center for Satellite Applications and Research, 67 pp.
- Hotelling, H., 1933: Analysis of a complex of statistical variables into principal
components. *Jour. of Educational Psychology*, **24**, 417-441, 498-520.
- Houston, A. L., J. Lahowetz, N. A. Lock, C. D. Oppermann, G. Limpert and B. L.
Barjenbruch, 2013: The thunderstorm observation by radar (ThOR)
algorithm. *Jour. of Appl. Meteo. And Climatology*, in preparation.
- Jedlovec, G. J., S. L. Haines and Frank J. LaFontaine, 2008: Spatial and temporal
varying thresholds for cloud detection in GOES imagery. *IEEE
Transactions on Geoscience and Remote Sensing*, **46**, NO. 6, 1705-1717.
- Johns, R. H. and C. A. Doswell, 1992: Severe local storms forecasting. *Wea.
Forecasting*, **7**, 588-612.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W. Wu and S. Lord, 2009:
Introduction of the GSI into the NCEP global data assimilation system.
Wea. Forecasting, **24**, 1691-1705.
- Lakshmanan, V., V. Debrunner and R. Rabin, 2002: Nested partitions using texture
segmentation. *Southwest Symposium on Image Analysis and Interpretation*,
IEEE, Santa Fe, New Mexico.
- Lakshmanan, V., R. Rabin and V. DeBrunner, 2003: Multiscale storm
identification and forecast. *J. Atmospheric Research*, 367-380.

- Lakshmanan, V., T. Smith, G. J. Stumpf and K. Hondl, 2007: The warning decision support system –integrated information, *Wea. Forecasting*, **22**, no. 3, pp. 596-612.
- Lakshmanan, V., K. Hondl and R. Rabin, 2009: An efficient, general-purpose technique to identify storm cells in geospatial images. *J. Ocean. Atmos. Tech.*, **26**, 523-537.
- Lakshmanan, V. and T. Smith, 2009: Data mining storm attributes from spatial grids. *J. Ocean. Atmos. Tech.*, **26**, 2353-2365.
- Mann, H. B. and D. R. Whitney, 1947: On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, **18**, 50-60.
- Markowski, P., C. Hannon and E. Rasmussen, 2006: Observations of convective initiation “failure” from the 12 June 2002 IHOP deployment. *Mon. Wea. Rev.*, **134**, 375-405.
- Mecikalski, J. R. and Kristopher M. Bedka, 2006: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.*, **134**, 49–78.
- Mecikalski, J. R., K. M. Bedka, S. J. Paech and L. A. Litten, 2008: A statistical Evaluation of GOES Cloud-Top Properties for Nowcasting Convective Initiation, *Mon. Wea. Rev.*, **136**, 4899-4914.
- Mecikalski, J. R., X. Li, L. D. Carey, E. W. McCaul Jr. and T. A. Coleman, 2013: Regional comparison of GOES cloud-top properties and radar characteristics in advance of first-flash lightning initiation. *Mon. Wea. Rev.*, **141**, 55-74.

- Mueller, C. K., J. W. Wilson and N. A. Crook, 1993: Utility of soundings and mesonets to forecast thunderstorm initiation. *Wea. Forecasting*, **8**, 132-146.
- Mueller, C., T. Saxen, R. Roberts, J. Wilson, T. Betancourt, S. Dettling, N. Oien and J. Yee, 2003: NCAR Auto-Nowcast System. *Wea. Forecasting*, **18**, 545-561.
- Pavlonis, M. J., 2010: Advances in extracting cloud composition information from spaceborne infrared radiances – a robust alternative to brightness temperatures. Part I: Theory. *Jour. of Appl. Meteo. And Climatology*, **49**, 1992-2012.
- Roberts, R. D. , T. Saxen, C. Mueller, J. Wilson, A. Crook, J. Sun and S. Henry, 1999: Operational application and use of NCAR's thunderstorm nowcasting system. Preprints, *15th Int. Conf. on Interactive Information and Processing Systems*, Dallas, TX, Amer. Meteor. Soc., 158–161.
- Roberts, R. D. and S. Rutledge, 2003: Nowcasting storm initiation using GOES-8 and WSR-88D data, *Wea. Forecasting*, **18**, 562-584.
- Roberts, R. D., A. R. S. Anderson, E. Nelson, B. G. Brown, J. W. Wilson, M. Pocerlich and T. Saxen, 2012: Impacts of forecaster involvement on convective storm initiation and evolution nowcasting. *Wea. Forecasting*, **27**, 1061-1089.
- Sieglaff, J. M., L. M. Counce, and W. F. Feltz, 2011: Nowcasting convective storm initiation using satellite-based box-averaged cloud-top cooling and cloud-type trends, *J. of Appl Meteor. and Climatol.*, **50**, 110-26.
- Sieglaff, J. M., D. C. Hartung, W. F. Feltz, L. M. Counce and V. Lakshmanan, 2013: A satellite-based convective cloud object tracking and multipurpose

- data fusion tool with application to developing convection. *J. Atmos. Oceanic Technol.*, **30**, 510–525.
- Walker, J. R., W. M. MacKenzie, J. R. Mecikalski and C. P. Jewett, 2012: An enhanced geostationary satellite-based convective initiation algorithm for 0-2 hour nowcasting with object tracking, *J. of Appl Meteor and Climatol.*, **51**, 1931-1949.
- Weckwerth, T. M. and D. B. Parsons, 2006: A review of convection initiation and motivation for IHOP_2002. *Mon. Wea. Rev.*, **134**, 5-22.
- Weisman, M. L. and J. B. Klemp, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, **112**, 2479-2498.
- Williams, E. and N. Renno, 1993: An analysis of the conditional instability of the tropical atmosphere. *Mon. Wea. Rev.*, **121**, 21-36.
- Zhang, J. and Y. Qi, 2010: A real-time algorithm for the correction of brightband effects in radar derived precipitation estimation. *J. Hydrometeor.*, **11**, 1157-1171.
- Zhang, J. and Coauthors, 2011: National mosaic and multi-sensor QPE (nmq) system: description, results and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321-1338.

8. Data Sources

National Severe Storms Laboratory, Norman, OK, October 2013: National Mosaic Quantitative Precipitation Estimation radar. Digital media. [Available online at nmq.ou.edu].

National Climatic Data Center, October 2013: Rapid Refresh. Digital media. [Available online at nomads.ncdc.noaa.gov].

The University of Alabama-Huntsville, Huntsville, AL, October 2013: Satellite Convection Analysis and Tracking, digital media. [Available online at www.nsstc.uah.edu].

Space Science and Engineering Center, Madison, WI, October 2013: University of Wisconsin Cloud Top Cooling algorithm, digital media. [Available online at <http://cimss.ssec.wisc.edu/snaap/convinit/quicklooks>].