

A MECHANISM FOR UPSCALE GROWTH OF CONVECTION IN THE COMPLEX
TERRAIN OF THE NORTHEAST U.S.

by

Brennan Stutsrim

A Thesis

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

College of Arts & Sciences

Department of Atmospheric and Environmental Sciences

Summer 2021

ABSTRACT

Upstate New York has a variety of complex terrain that can interact with the background flow to create mesoscale heterogeneities in the lower troposphere. The major valleys of Upstate New York, the Hudson and Mohawk Valleys, often have increased moisture content and stronger surface winds than the higher terrain surrounding them. These features can have a profound effect on the evolution of convective storms, especially in cases characterized by low-to-moderate shear, which tends to favor pulse-like or multicellular convection. Analysis of composite radar imagery has indicated that convective storms often change mode while descending from the Catskills Mountains into the Hudson Valley, coinciding with an increase severe weather reports. Some storms exhibited back building once reaching the Hudson Valley. Back building is when a convective line has new cells initiating adjacent to the mature cells such that the line propagates upstream with regards to the low-level flow. Back-building mesoscale convective systems (MCS) have been connected to an increased threat of heavy rainfall and flash flooding, especially in the Northeast.

A back-building MCS from 21 August 2019 was simulated using WRF-ARW to study the mesoscale interactions between the background flow, complex terrain features in and around the Hudson Valley, and the MCS's convective cold pool. During the three hours preceding the MCS, southerly terrain-channeled flow created a favorable pre-convective environment in the Hudson Valley through a low-level maximum in water vapor flux. Discrete convection from the Catskill Mountains intensified once reaching the Hudson Valley, creating a cold pool with an outflow boundary oriented across the valley. The channeled flow increased the low-level convergence along the southern portion of the outflow boundary causing high equivalent

potential temperature (θ_e) air from the lower Hudson Valley to be lifted, initiating new convective cells. The MCS propagated down the valley until the channeled flow was cut off by another convective line entering the lower Hudson Valley.

Decision trees were created to identify characteristics of the pre-convective environment that are conducive to back building in the Hudson Valley. Composite radar imagery was analyzed to identify cases with (n=15) and without (n=55) back building from June, July and August 2015-2020. HRRR 0-hour analyses, valid at 1800 UTC, were used to calculate area averaged variables in the Hudson Valley for each case and analyzed by the decision tree classifier. Variables related to surface-based instability, such as surface-based CAPE and lifted index (LI), and low-level moisture content, such as 2-m AGL dew point depression (DD), were chosen most often by the decision tree classifier. A high value of surface-based instability makes it more likely that a new updraft along the outflow boundary will grow into deep convection. A low 2-m DD makes it more likely that a small vertical displacement of the surface parcel by the outflow boundary will result in saturation and positive buoyancy. Comparison of composite wind profiles from cases with and without back building revealed a difference in the wind speeds from 900 hPa to the tropopause. A mean tropospheric wind of 30 kt made it more likely that cells would move out of the Hudson Valley before forming a strong cold pool. A mean tropospheric wind of 20 kt resulted in slower cell motion, allowing stronger cold pools to form in the Hudson Valley and increasing the likelihood of back building.

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iv
List of Figures.....	viii
1. Introduction.....	1
1.1 Motivation.....	1
1.2 Literature review.....	3
1.2.1 Terrain Channeling.....	3
1.2.2 Back-Building Convection.....	5
2. Data and Methodology.....	19
2.1 Case Selection.....	19
2.2 WRF Simulations.....	20
2.3 HRRR Analysis Data.....	21
2.4 Decision Trees.....	24
3. Physical Mechanisms.....	34
3.1 Case Study 21 August 2019.....	34
3.1.1 Case Overview.....	34
3.1.2 Synoptic Environment.....	35
3.1.3 Local Environment.....	36
3.2 WRF Simulation.....	37
3.2.1 Comparison with Observed Case.....	37
3.2.2 Convective Cold Pool.....	39
3.2.3 Terrain Channeling.....	41

4. Prediction	58
4.1 Decision Tree Results	58
4.2 Synoptic Environment	61
4.3 False Positives	63
4.4 Failure Modes	65
5. Illustrative Cases	83
5.1 Overview	83
5.2 29 June 2021	83
5.3 30 June 2021	84
5.2 Summary	86
6. Conclusions and Future Work	91
6.1 Conclusions	91
6.1.1 Physical mechanisms	91
6.1.2 Prediction	93
6.2 Future Work	95
References	97

LIST OF FIGURES

Table 1. Variables populating the sets of features used in decision tree analyses. Variables marked as “taken from analysis” are fields found in 1800 UTC HRRR 0-hour analyses surface files, while variables marked as “calculated from analysis” were calculated using fields found in the surface files.

Fig. 1.1. 3-D rendering of the topography in New York State. The major terrain features in eastern New York are labeled. [Adapted from <https://pixels.com/featured/new-york-state-usa-3d-render-topographic-map-border-frank-ramspott.html>]

Fig. 1.2. Mean MLCAPE from HRRR model on convective days between 1 Feb and 30 Sep 2013-15 for (a) all days, (b) SW-flow days and (c) SE-flow days in the Northeast. [Figure 2 and caption adapted from Katona et al. (2016, section3).]

Fig. 1.3. (a) Storm frequency Pr_{storm} per 15-min interval (shaded every 0.03) for a composite reflectivity ≥ 45 dBZ from April through September during the warm seasons of 1996–2007. Values have been multiplied by 100 and therefore are equivalent to percentages. (b) The monthly cloud-to-ground lightning frequency per square kilometer (shaded every 0.10) during the warm season. Terrain height is contoured every 200 m starting at 100 m. [Figure 2 and caption recreated from Murray et al. (2011, section 3).]

Fig. 1.4. Cross section from an idealized terrain channeling simulation of the velocity component v (parallel to the valley) for a geostrophic wind from the west (perpendicular to the valley) showing a pronounced channeling; intervals, 1 m/s. [Figure 11 and caption adapted from Gross et al. (1987, section 4).]

Fig. 1.5. Relationships between above-valley (geostrophic) and valley wind directions for four possible forcing mechanisms: thermal forcing, downward momentum transport, forced channeling, and pressure-driven channeling. The valley is assumed to run from northeast to southwest. [Figure 1 and caption reproduced from Whiteman et al. (1993, section 1).]

Fig. 1.6. Isolines for the relative frequency of the surface wind directions associated with a given direction of the geostrophic wind at the station Mannheim in the upper Rhine Valley. The heavy solid line connects the most frequent directions; data for stable stratification. [Figure 4 and caption reproduced from Gross et al. (1987, section 2).]

Fig. 1.7. Wind rose showing the frequency distribution of wind speed (U) and direction at HOB0 station H1 (located on eastern edge of central Hudson Valley) for Days 150-270 (30 May – 27 September) during years 2004 and 2005. [Figure 3 and caption adapted from Freedman et al. (2017, section 3).]

Fig. 1.8. Schematic reflectivity drawing of idealized life cycles for three linear MCS archetypes: (a) TS, (b) LS, and (c) PS. Approximate time intervals between phases: for TS 3–4 h; for LS 2–3 h; for PS 2–3 h. Levels of shading roughly correspond to 20, 40, and 50 dBZ. [Figure 4 and caption reproduced from Parker et al. (2000, section 4).]

Fig. 1.9. Schematic diagram of the radar-observed features of the (a) TL/AS and (b) BB patterns of extreme-rain-producing MCSs. Contours (and shading) represent approximate radar reflectivity values of 20, 40, and 50 dBZ. In (a), the low-level and midlevel shear arrows refer to the shear in the surface-to-925-hPa and 925–500-hPa layers, respectively. The dash-dot line in (b) represents an outflow boundary; such boundaries were observed in many of the BB MCS cases. The length scale at the bottom is approximate and can vary substantially, especially for BB systems, depending on the number of mature convective cells present at a given time. [Figure 3 and caption reproduced from Schumacher et al. (2005, section 3).]

Fig. 1.10. Schematic diagram of the four back-building subclasses of events. Subclasses are determined by the relative motion of back-building cells and MCSs and their relative transits across the flood area. Gray areas indicate the areas where flash flooding is most likely. [Figure 2 and caption reproduced from Jessup et al. (2012, section 3).]

Figure 1.11. Comparison schematics of (top) original (upwind) and (bottom) downwind versions of the vector technique to forecast short-term motion of upwind-developing systems (MCS motion given by thick red arrows) and downwind-developing systems (MCS motion given by thick green arrow at bottom of figure), respectively. Vector lengths are proportional to wind speed; MCS centroids are denoted by the cross symbol. [Figure 6 and caption reproduced from Corfidi (2003, section 3).]

Fig. 1.12. Composite wind profile at (a) the grid center for TL/AS and BB MCS extreme rain events at the peak rainfall time. Wind barbs are plotted conventionally in kt. [Figure 13 and caption adapted from Schumacher et al. (2005, section 4).]

Fig. 1.13. Frequency of northeast U.S. flash flood events, 2003–07, by event type, with radar-estimated rainfall meeting the 24-h precipitation threshold of [Schumacher and Johnson \(2006\)](#). Both the total frequency and the relative frequency (%) are given for these 23 events. [Figure 7 and caption recreated from Jessup et al. (2012, section 4).]

Fig. 2.1. Composite radar imagery showing an example of a back-building MCS (circled in red) in the Hudson Valley from 23 July 2020 at (a) 1930 UTC, (b) 2030 UTC, (c) 2155 UTC and (d) 2325 UTC.

Fig. 2.2. Composite radar imagery showing an example of convection without back building (circled in red) in the Hudson Valley from 13 August 2016 at (a) 1830 UTC, (b) 1930 UTC, (c) 2020 UTC and (d) 2120 UTC.

Fig. 2.3. Location of model domains used in WRF simulations of back-building convection. The inner most domain is centered over the Catskill Mountains, a common convective initiation location, and the Hudson Valley, a common back-building location. The horizontal resolution used in each domain is 9, 3 and 1 km respectively.

Fig. 2.4. Map of terrain height from the inner-most domain of the WRF simulations. Terrain height is filled at an interval of 25 m and county borders appear in grey.

Fig. 2.5 Monthly frequencies of high-impact (blue), missed (red), and good (green) events from 1980 to 2013. [Figure 6 and caption reproduced from Vaughan et al. (2016, section 3).]

Fig. 2.6. Map of terrain height from inner-most domain of WRF simulations. Terrain height is filled at an interval of 25 m and county borders appear in grey. Red box denotes the area considered to be the Hudson Valley for area averages.

Fig. 2.7. Example decision tree that can provide guidance on the likelihood of back-building convection based on pre-convective parameters. Cases that have values lower than the threshold at each decision node enter the next lowest node on the left, while values higher than the threshold go to the right.

Fig. 2.8. Parameter phase space of 2-m dew point depression and 10-m wind direction. The thick black line shows an example of a threshold used in sequential feature selection to separate BB and no-BB cases. Cases with a DD below 17.5 would be kept and all other cases removed from the dataset.

Fig. 3.1. Composite radar imagery showing a back-building MCS in the Hudson Valley from 21 August 2019 at (a) 1855 UTC, (b) 2000 UTC, (c) 2030 UTC and (d) 2100 UTC.

Fig. 3.2. Composite radar imagery depicting the back-building MCS at (a) 1845 UTC, (b) 1900 UTC, (c) 1928 UTC and (d) 2032 UTC on 21 August 2019. Yellow lines denote severe thunderstorm warning areas issued by the Albany WFO at each time.

Fig. 3.3. Storm total quantitative precipitation estimate map valid at 2300 UTC 21 August 2019. Data is taken from NCEP Stage IV QPE with a horizontal resolution of 4 km. Thin black lines denote county borders and thick black lines denote state borders.

Fig. 3.4. Map of severe weather and flash flood reports from 21 August 2019. Text showing the approximate time of each cluster of reports and arrows denote the chronological evolution of the responsible convection. Adapted from NOAA Interactive Local Storm Reports.

Fig. 3.5. Maps of (a) 300-hPa geopotential height (dam), isotachs (kt) and wind (kt); (b) 500-hPa geopotential height (dam), relative vorticity (10^{-5}s^{-1}) and wind (kt); (c) 700-hPa geopotential heights (dam), precipitable water (mm) and wind (kt); (d) 850-hPa geopotential heights (dam), temperature (F) and wind (kt). All maps valid at 1800 UTC 21 August 2019 and data taken from HRRR 0-hr analysis.

Fig. 3.6. Weather Prediction Center (WPC) surface analysis valid at 1800 UTC 21 August 2019.

Fig. 3.7. Meteograms showing (a) temperature (F) and solar insolation (W/m^2), (b) wind (mph) and pressure (hPa), (c) relative humidity (%) and observed precipitation (in), (d) wind direction and magnitude (mph). Data taken from New York State Mesonet (NYSM) Station in Schodack, NY.

Fig. 3.8. Area-averaged Hudson Valley model soundings valid at 1800 UTC 21 August 2019 from (left) WRF simulation and (right) HRRR 0-hr analysis.

Fig. 3.9. Left panels showing composite radar imagery from 1855 UTC (top) and 2100 UTC (bottom) 21 August 2019. Right panels showing simulated reflectivity from WRF simulation of 21 August 2019, valid at 1855 UTC (top) and 2100 UTC (bottom). The 125-m terrain height contour is included to roughly outline the Mohawk and Hudson Valleys.

Fig. 3.10. Left panel showing simulated storm total precipitation from WRF simulation of 21 August 2019, valid at 2300 UTC. Right panel showing observed storm total precipitation from 21 August 2019, valid at 2300 UTC, taken from NCEP Stage IV QPE.

Fig. 3.11. Maps of (left) simulated reflectivity, (center) filled reflectivity with contoured θ_e every 1 K, and (right) filled reflectivity with upward (green) and downward (purple) vertical motion contoured at 0.5, 1, 1.5 and 3 m/s. The 125-m terrain height contour is included to roughly outline the central and upper Hudson Valley. All maps valid at (top) 1855 UTC and (bottom) 1910 UTC.

Fig. 3.12. As in Figure 3.11, but maps valid at (top) 1940 UTC and (bottom) 2000 UTC.

Fig. 3.13. Cross section from WRF simulation across the gust front of the back-building MCS showing filled potential temperature and vertical motion contoured at 1, 3, 5, 8 and 12 m/s. Cross section valid at 2000 UTC. Inset shows simulated reflectivity, 125-m terrain height contour and location of the cross section (black line).

Fig. 3.14. As in Figure 3.11, but maps valid at (top) 1940 UTC and (bottom) 2000 UTC.

Fig. 3.15. As in Figure 3.13, but valid at 2050 UTC.

Fig. 3.16. As in Figure 3.11, but valid at (top) 2120 UTC and (bottom) 2140 UTC.

Fig. 3.17. Cross sections from WRF simulation across the Hudson Valley showing (left) wind speed (m/s) and (right) wind direction (deg), valid at 1800 UTC. Insets show simulated reflectivity, 125-m terrain height contour and location of cross section (black line).

Fig. 3.18. Cross section from WRF simulation across the Hudson Valley (left) and along the Hudson Valley (right) showing horizontal vapor flux (filled) and wind speed (contoured), valid at 1800 UTC. Inset shows simulated reflectivity, 125-m terrain height contour and location of cross section (black line).

Fig. 3.19. As in Figure 3.18, but valid at 2055 UTC, the time at which the back-building MCS was most intense.

Fig. 4.1. Histogram showing the number of decision trees in which each feature was chosen from the short list of variables. Colors indicate the weight that was assigned to back-building cases in the decision tree classifier. Features represent the area average in the Hudson Valley taken from the HRRR 1800UTC 0-hr analyses.

Fig. 4.2. Parameter space of back-building (red) and non-back-building (gray) cases. Parameters are the Hudson Valley area-averaged variables that were most commonly chosen from the short list by the decision trees. Back-building cases have been given a weight of (a) 1, (b) 2, (c) 3, and (d) 4 respectively. SBCAPE (J/kg) and 2-m DD (F) were chosen for all weights. Bold lines indicate the “best splits” from the decision tree, with the shaded area indicating the combination of parameters that yields the best identification of back-building versus non-back-building cases.

Fig. 4.3. As in Figure 4.1, but with the expanded list of variables.

Fig. 4.4. Parameter space of back-building (red) and non-back-building (gray) cases. Parameters are the Hudson Valley area-averaged variables that were most commonly chosen from the expanded list by the decision trees. Back-building cases have been given a weight of (a) 1, (b) 2, (c) 3, and (d) 4 respectively. Lifted index (F) was chosen for all weights, with the second selected features being (a) 850-hPa dew point (F), (b) 2-m temperature (F), (c) 2-m temperature (F), and (d) 10-m wind magnitude (m/s). Bold lines indicate the “best splits” from the decision tree, with the shaded area indicating the combination of parameters that yields the best identification of back-building versus non-back-building cases.

Fig. 4.5. As in Figure 4.1, but with factors identified through the PCA.

Fig. 4.6. Composite maps of back-building cases identified by the decision tree analysis with the expanded list of variables. Maps show (top) 300-hPa geopotential heights (dam), isotachs (kt) and wind (kt); and (bottom) 500-hPa geopotential heights (dam), relative vorticity (10^{-5}s^{-1}) and wind (kt) from 1800 UTC HRRR 0-hr analyses.

Fig. 4.7. Spaghetti plot of 500-hPa 582 dam geopotential height contours from back-building cases identified by the decision tree analysis with the expanded list of variables. The thick contour is the mean position of the 582 dam contour. Map taken from 1800 UTC HRRR 0-hr analyses.

Fig. 4.8. Composite maps of back-building cases identified by the decision tree analysis with the expanded list of variables. Maps show (top) 700-hPa geopotential heights (dam), precipitable water (mm) and wind (kt); and (bottom) 850-hPa geopotential heights (dam), temperature (F) and wind (kt) from 1800 UTC HRRR 0-hr analyses.

Fig. 4.9. Composite map of back-building cases identified by the decision tree analysis with the expanded list of variables. Map shows the MSLP (hPa), 2-m dew point (F) and 10-m AGL wind (kt) from 1800 UTC HRRR 0-hr analyses.

Fig. 4.10. Area averaged Hudson Valley sounding from back-building cases identified by the decision tree analysis with the expanded list of variables, taken from 1800 UTC HRRR 0-hr analyses.

Fig. 4.11. As in Figure 4.6, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

Fig. 4.12. As in Figure 4.8, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

Fig. 4.13. As in Figure 4.9, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

Fig. 4.14. As in Figure 4.10, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

Fig. 4.15 Composite soundings for back-building (red) and false-positive (grey) cases. Parcel path for a surface-based parcel is shown with a circle indicating the LCL. The difference between the back-building and false-positive composite wind profiles is shown in barbs on the exterior axis.

Fig. 4.16. Composite radar imagery from (left) 1925 UTC, and (right) 2320 UTC 19 July 2019. The approximate motion of convective cells is shown in the black arrows.

Fig. 5.1. Composite radar imagery depicting a back-building MCS in the Hudson Valley from 29 June 2021 at (a) 2005 UTC, (b) 2040 UTC, (c) 2125 UTC and (d) 2235 UTC.

Fig. 5.2. Area-averaged Hudson Valley sounding valid at 1800 UTC 29 June 2021. Taken from HRRR 0-hr analysis.

Fig. 5.3. Composite radar imagery showing convection in the Hudson Valley from 30 June 2021 at (a) 1835 UTC, (b) 1920 UTC, (c) 1950 UTC and (d) 2015 UTC.

Fig. 5.4. Sounding launched from KALY at 1800 UTC 30 June 2021. [Reproduced from SPC Observed Sounding Archive <https://www.spc.noaa.gov/exper/soundings/>]

1. Introduction

1.1 Motivation

Forecasting severe convection in the Northeastern United States is made difficult by complex topographical features, which can create mesoscale differences in surface environmental characteristics. The convective environment in valleys, such as the Hudson and Mohawk Valleys of Upstate New York (Figure 1.1), can be influenced by interactions between the terrain and background flow, including terrain channeling of surface winds and variations in boundary layer thermodynamic properties (Wasula and Bosart 2002, Freedman and Fitzjarrald 2017). Modification of the convective environment by terrain channeling can contribute to severe weather outbreaks by helping to create low-level moisture convergence (Riley and Bosart 1983). Katona et al. (2016) found that the High Resolution Rapid Refresh (HRRR) model produces localized convective environments that are heavily affected by the terrain features in Upstate New York.

The Hudson and Mohawk Valleys tend to have higher values of mixed-layer CAPE (MLCAPE) than the Catskill and Adirondack Mountains, especially under a southwesterly flow regime (Figure 1.2). Higher values of 0–1-km storm-relative helicity within the Mohawk and Hudson Valleys are also evident, owing to an increase in surface wind speed compared to the mountainous regions (Katona 2016). These spatial differences lead to convective storms and lightning occurring more frequently within the valleys of Upstate New York compared to the Catskill and Adirondack Mountains (Figure 1.3).

During August 2019, there were 11 days with damaging winds, hail or tornadoes in New York State. This was an unusually active month for severe weather, with 263 reports compared

to a climatological value of 76 reports within New York State according to the NOAA Storm Events Database (<https://www.ncdc.noaa.gov/stormevents>). Analysis of synoptic conditions using the National Center for Environmental Prediction's (NCEP) Climate Forecast System Reanalysis showed a reoccurring theme of weak synoptic forcing with low vertical wind shear and moderate to high CAPE. Weak synoptic forcing environments, with small height falls or even height rises occurring at 500 hPa, can still produce severe convection (Thompson and Grams 2008). The predictability of severe convection in the Northeast is impacted by the amount of vertical wind shear and CAPE, with low-shear, high-CAPE environments having lower predictive skill (Vaughan 2017).

Radar composites from August 2019 showed that the weak synoptic forcing favored discrete convection initiating over the Finger Lakes region and Catskill Mountains. The discrete convection that moved to the east often changed mode as it descended off of the Catskill Mountains into the Hudson Valley, growing upscale into multicellular storm modes. Some cells exhibited quasi-stationary or back-building behavior, which can extend the amount of time that a location is impacted by heavy precipitation. Conversations with National Weather Service forecasters from the Albany and Boston Weather Forecasting Offices (WFO) revealed that forecasting when and where changes in convective mode will occur can be a challenge in the Northeast. Thompson and Grams (2008) discussed how they think forecasters pay too much attention to the synoptic forcing. It is hypothesized that the terrain influences modifying the low-level pre-convective environment can also have a strong effect on convective evolution in marginal severe weather environments.

Back-building mesoscale convective systems (MCSs) have been studied extensively in the Great Plains and Midwest regions of the United States. Prior research has focused on their

structure (Parker 2000), their direction and speed of propagation (Corfidi 1996; 2003), and their connection with extreme rain events (Schumacher 2005). However, comparatively little research has been conducted on back-building MCSs in the Northeast. It has been found that back-building convection is responsible for 31% of non-tropical flash flood events in the Northeast (Jessup 2012), making this type of convective evolution important to investigate further.

1.2 Literature Review

1.2.1 Terrain Channeling

Terrain channeling is when the background flow is modified by terrain to change the direction or magnitude of the low-level flow. Terrain-channeled flow is often in a different direction than the flow above the terrain features. Figure 1.4 (Gross and Wippermann 1987) shows a cross section, viewed from the south, depicting a simulated terrain-channeled flow within the Rhine Valley of Germany. Despite the simulated large-scale winds being westerly, the surface winds within the valley are southerly, with the fastest winds located closer to the higher terrain feature in the east. There are several different physical mechanisms that can cause flow in a valley to be channeled. Four models for terrain-channeled flow have been established: thermally driven (Whiteman 1990), downward momentum transport, forced channeling, and pressure driven channeling (Whiteman 1993). Figure 1.5 shows phase spaces for each model that depict the direction of the synoptic (above-valley) wind and associated surface wind in a valley that runs from southwest to northeast.

Thermally driven winds are an ageostrophic wind that develops due to hydrostatic pressure gradients created by temperature differences forming along the valley. Thermally driven

winds are not dependent on the above-valley winds. Instead, thermally driven winds are driven by diurnal temperature cycles and the direction is reversed between the day and night (Figure 1.5a,b,c).

Downward momentum transport occurs when the above-valley winds are mixed down towards the surface. The resulting winds, which are strongly coupled with the above-valley winds, are nearly geostrophic, but increased friction in the surface layer causes a slight turning towards low pressure (Figure 1.5d). Downward momentum transport is made more likely with increasingly unstable conditions and is more common in wide valleys with small terrain features on either side.

Forced channeling occurs when the geostrophic above-valley winds are forced to follow the along-valley axis due to the terrain on either side of the valley. While the above-valley winds impact whether the surface flow is up or down the valley, the direction of forced channeling is always parallel to the along-valley axis (Figure 1.5e). Forced channeling is more common for narrow valleys (Carrera 2009) and unstable or neutral conditions.

Pressure driven channeling occurs when the large-scale pressure gradient has a component parallel to the valley, creating an ageostrophic wind within the valley. Pressure driven channeling can result in counter-currents, which occur when the large-scale pressure gradient is at an angle to the valley such that the pressure gradient force within the valley is opposite to the geostrophic wind direction. The resulting valley winds are in the opposite direction of the large-scale wind's along-valley component (Figure 1.5f). Counter-currents have been observed in the Rhine Valley of Germany, where a geostrophic wind from the north, east, or south most often coincides with a due-northerly surface wind in the valley (Figure 1.6).

The surface winds in the Hudson Valley are often impacted by the terrain on either side, the Catskill Plateau on the west and Taconic mountains on the east (Figure 1.1). The Hudson Valley is roughly 40 km wide with terrain mostly reaching 300 m, but as high as 1 km. This aspect ratio is similar to the Rhine Valley (Gross and Wippermann 1987) and Saint Lawrence River Valley (Razy 2012) studied in earlier research. An observational study found that the surface winds in the Hudson Valley between 30 May and 27 September 2004-2005 were often parallel to the valley, with southerly winds occurring most often and at the strongest magnitude (Figure 1.7). Southerly winds are dominant in the Hudson Valley when the above-valley winds range from southerly to westerly, caused by a combination of forced and pressure driven channeling (Sakai and Fitzjarrald 2004).

1.2.2 Back-Building Convection

Back-building convection was characterized by Bluestein and Jain (1985) as a line of convective cells where new cells are initiated upstream, with respect to the cell motion, of existing cells and merge with the existing cells. Numerous pieces of earlier literature described back building, often while discussing modes of linear squall lines in the Great Plains and Midwest regions of the United States (Abdullah 1954, Stout and Hiser 1955, Newton and Katz 1958, Newton 1963, Newton and Fankhauser 1964, Henz 1973). Back-building convection is multicellular by definition, as at any given time there are cells at different stages in their life cycle within a continuous line. Back-building MCSs most closely resemble the parallel stratiform precipitation archetype (Figure 1.8c; Parker 2000). Figure 1.9b (Schumacher 2005) shows a schematic of a back-building MCS, including the outflow boundary which is created by

the existing cells and acts as the lifting mechanism to initiate new cells. The new cells develop and mature while located in the most intense portion of the MCS, and then dissipate in the stratiform region as they continue their motion parallel to the line.

Training lines are similar to back-building MCSs in that they consist of multicellular convection with cell motion parallel to the line, however training lines do not propagate upstream and include an adjoining region of stratiform precipitation displaced perpendicular to the line (Figure 1.9a; Schumacher 2005). Back-building MCSs in the Northeast occur on days with other forms of MCSs and sometimes continue back building until another type of MCS propagates into the same area, causing the MCSs to merge (Jessup 2012). A schematic showing this evolution can be found in Figure 1.10.

The propagation of back-building MCSs is controlled by the balance between individual cell motion and new cells being initiated upstream, making a quasi-stationary propagation possible. A technique for estimating the short-term (3–6-h) motion of MCSs was developed originally in Corfidi et al. (1996). The vector technique, referred to as Corfidi Vectors, was the vector sum of the mean cloud-layer wind, meant to represent the individual cell motion, and a vector equal in magnitude and opposite in direction to the low-level jet (LLJ), meant to represent the effect of new cell initiation. The original Corfidi Vectors do not account for changes in the convective environment and thus would need to be recalculated often in a forecasting scenario. The original technique was useful for forecasting the motion of back-building MCSs, but was developed to be used in areas with little orography, as more complex terrain would make the representation of new cell initiation by the LLJ an incomplete assumption.

The Corfidi vectors were later revised in Corfidi et al. (2003) in an attempt to address the shortcomings of the original technique. The revised technique focused on improving the

assumptions made in representing the new cell initiation by taking the orientation and motion of the convective cold pool into account. The new Corfidi Vectors had separate forecasted motion vectors for upwind and downwind-propagating MCSs. The upwind-propagation vector remains the sum of the cloud-layer winds and the negative of the LLJ, while the downwind-propagating vector is the sum of the upwind-propagating vector and the cloud-layer winds (Figure 1.11; Corfidi 2003). This new vector technique has improved forecasting applications, as it can differentiate between upwind and downwind-propagating MCS's. However, it would still likely struggle in determining the motion of an MCS in complex terrain, where upslope and downslope flow can affect the speed of gust fronts through gravity currents and provide an added lifting mechanism for new cells.

Schumacher et al. (2005) described the environments associated with back-building MCSs, which occur along outflow boundaries from pre-existing convection, and compared them to training lines, which often occur along stationary fronts. The back-building MCS environment tended to have a surface wind shift and a strong theta-e gradient, but only within the immediate vicinity of the outflow boundary. The strongest low-level convergence was found on the upstream side of the outflow boundary, where the high theta-e air is lifted to initiate new cells. High relative humidity (RH) values were found near the surface at the initiation site and were tilted towards the stratiform precipitation region with height, consistent with the lifting of moist surface air to create the new convective cells which were trained along the outflow boundary. While the outflow boundaries were consistent in BB MCSs, the large-scale environment tended to have more variation, with common characteristics in the moisture and wind profiles. Back-building MCSs were found in environments with high precipitable water values, above 45 mm, which increased quickly in the preceding hours. Figure 1.12 is the composite wind profiles at the

center of the MCS for training lines with adjoining stratiform precipitation and back-building MCS's during the peak rainfall time. Back-building MCSs had a wind profile with weaker winds in the upper levels than training lines and directional shear contained in the lower-troposphere.

Though back-building MCSs tend to be relatively small systems, on the order of 100-150 km, they consist of deep convection forming and passing over the same location for an extended period of time making flash floods their most common hazard. In the Great Plains, back-building MCSs represent 19.7% of extreme rain producing MCSs, making them the second most common type behind training lines (31.6%) (Schumacher 2005). In the Northeast U.S., back building is even more often responsible for flash flood events. Back-building MCSs were solely responsible for 31% of non-tropical flash flood events that met the 24-h precipitation thresholds used by Schumacher et al. (2006) (Figure 1.13). Events where there were multiple back-building MCSs or back-building MCSs that merged with over convective lines made up another 22% of the flash flood events.



Fig 1.1. 3-D rendering of the topography in New York State. The major terrain features in eastern New York are labeled. [Adapted from <https://pixels.com/featured/new-york-state-usa-3d-render-topographic-map-border-frank-ramspott.html>]

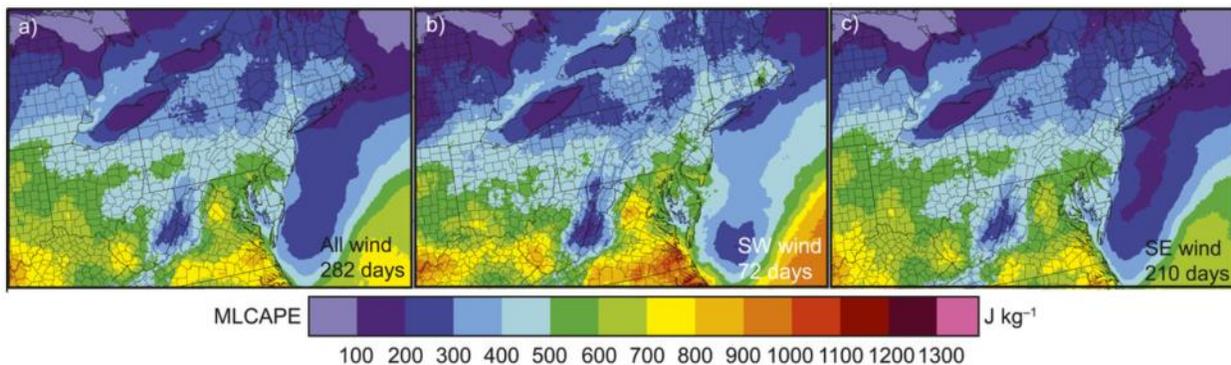


Fig. 1.2. Mean MLCAPE from HRRR model on convective days between 1 Feb and 30 Sep 2013-15 for (a) all days, (b) SW-flow days and (c) SE-flow days in the Northeast. [Figure 2 and caption adapted from Katona et al. (2016, section3).]

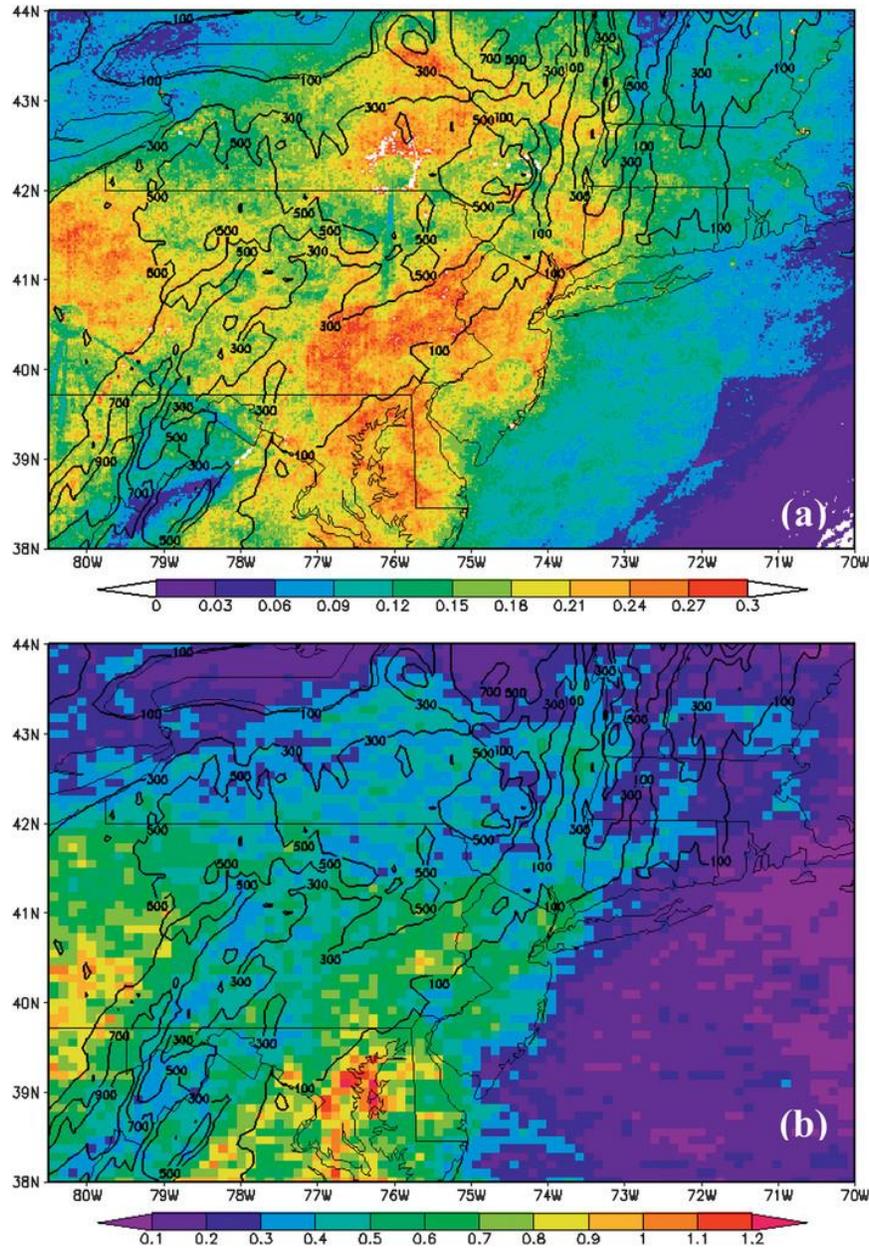


Fig. 1.3. (a) Storm frequency Pr_{storm} per 15-min interval (shaded every 0.03) for a composite reflectivity ≥ 45 dBZ from April through September during the warm seasons of 1996–2007. Values have been multiplied by 100 and therefore are equivalent to percentages. (b) The monthly cloud-to-ground lightning frequency per square kilometer (shaded every 0.10) during the warm season. Terrain height is contoured every 200 m starting at 100 m. [Figure 2 and caption recreated from Murray et al. (2011, section 3).]

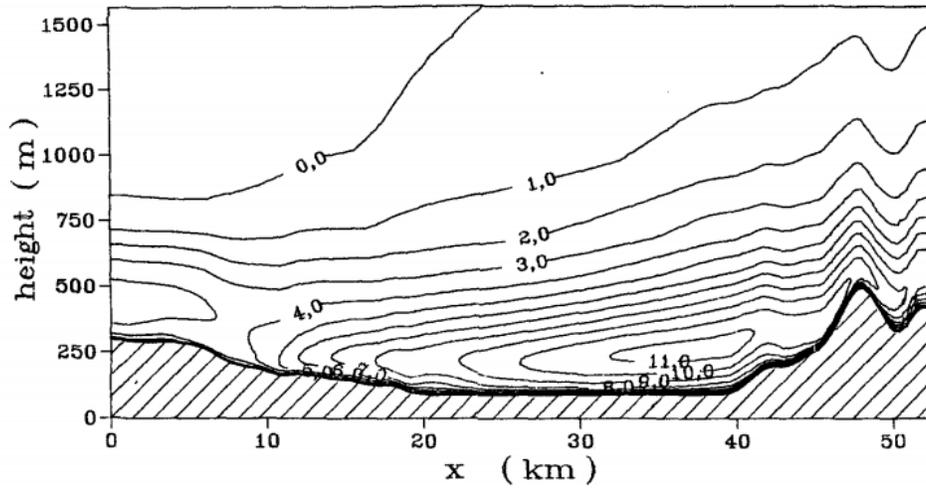


Fig. 1.4. Cross section from an idealized terrain channeling simulation of the velocity component v (parallel to the valley) for a geostrophic wind from the west (perpendicular to the valley) showing a pronounced channeling; intervals, 1 m/s. [Figure 11 and caption adapted from Gross et al. (1987, section 4).]

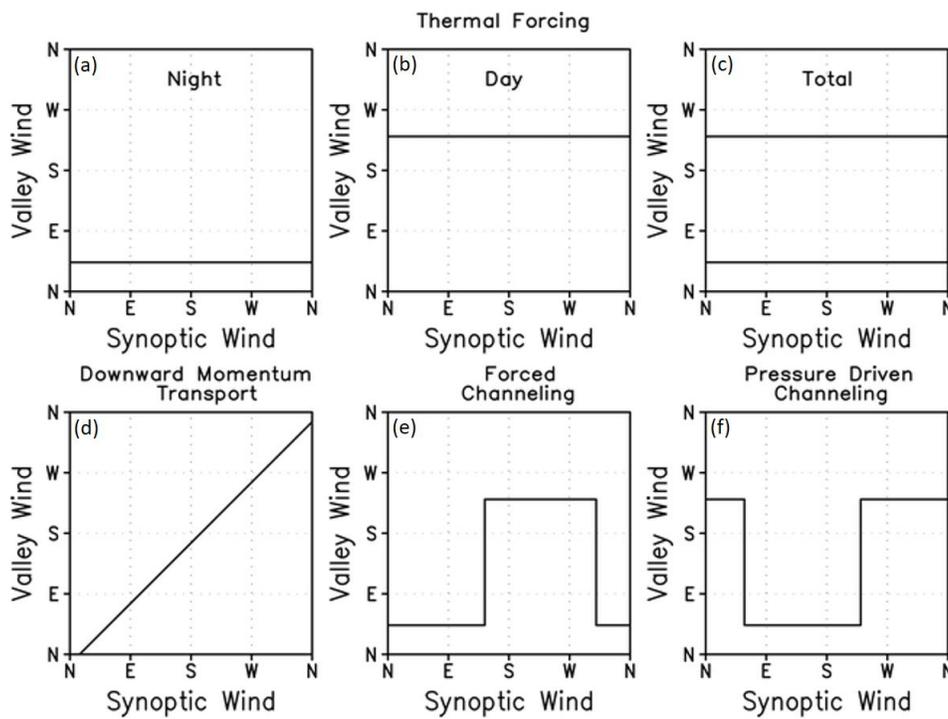


Fig. 1.5. Relationships between above-valley (geostrophic) and valley wind directions for four possible forcing mechanisms: thermal forcing, downward momentum transport, forced channeling, and pressure-driven channeling. The valley is assumed to run from northeast to southwest. [Figure 1 and caption reproduced from Whiteman et al. (1993, section 1).]

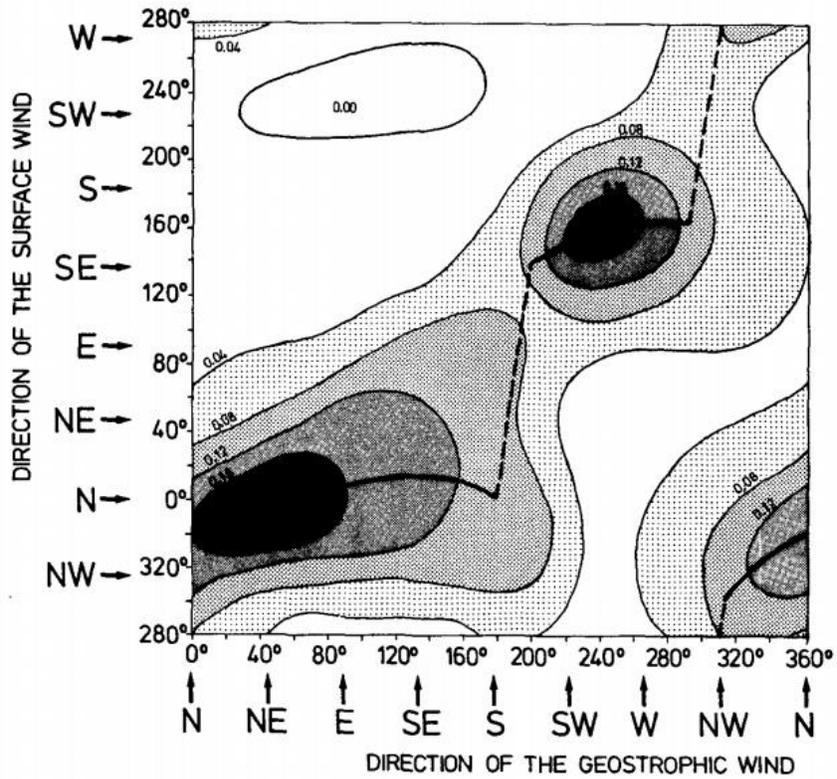


Fig. 1.6. Isolines for the relative frequency of the surface wind directions associated with a given direction of the geostrophic wind at the station Mannheim in the upper Rhine Valley. The heavy solid line connects the most frequent directions; data for stable stratification. [Figure 4 and caption reproduced from Gross et al. (1987, section 2).]

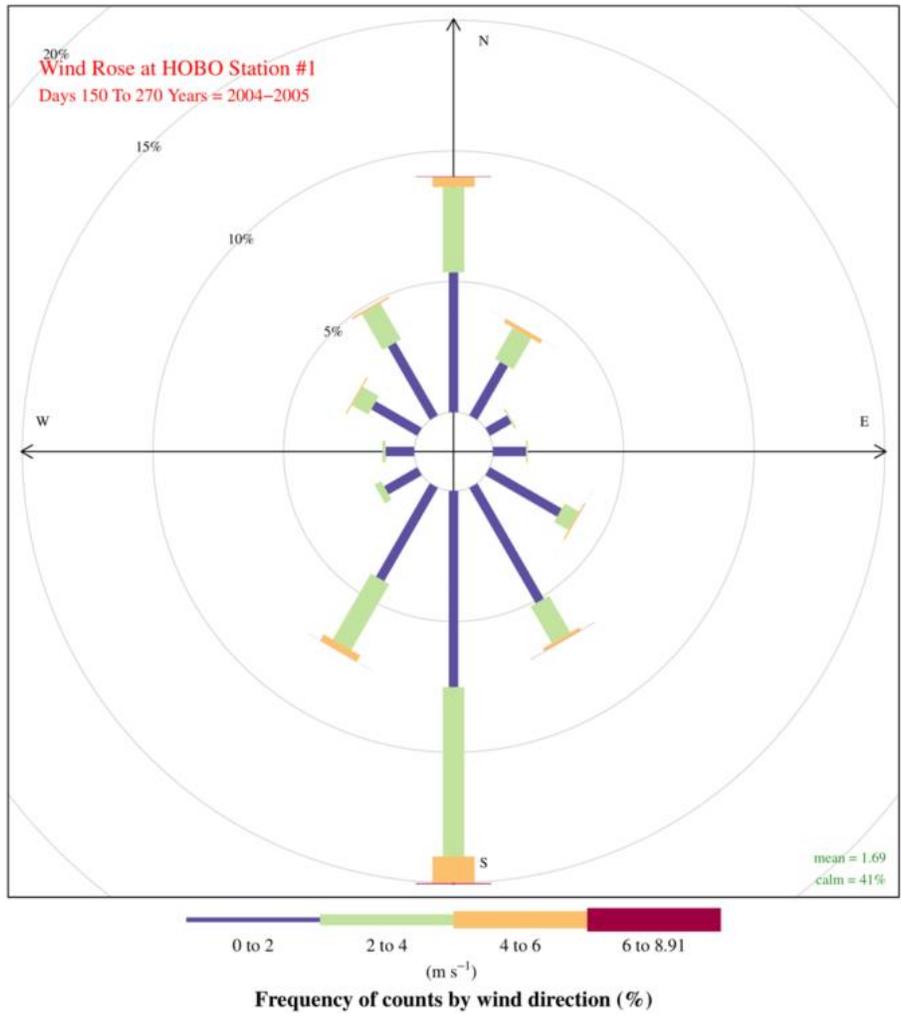


Fig. 1.7. Wind rose showing the frequency distribution of wind speed (U) and direction at HOBO station H1 (located on eastern edge of central Hudson Valley) for Days 150-270 (30 May – 27 September) during years 2004 and 2005. [Figure 3 and caption adapted from Freedman et al. (2017, section 3).]

Linear MCS archetypes

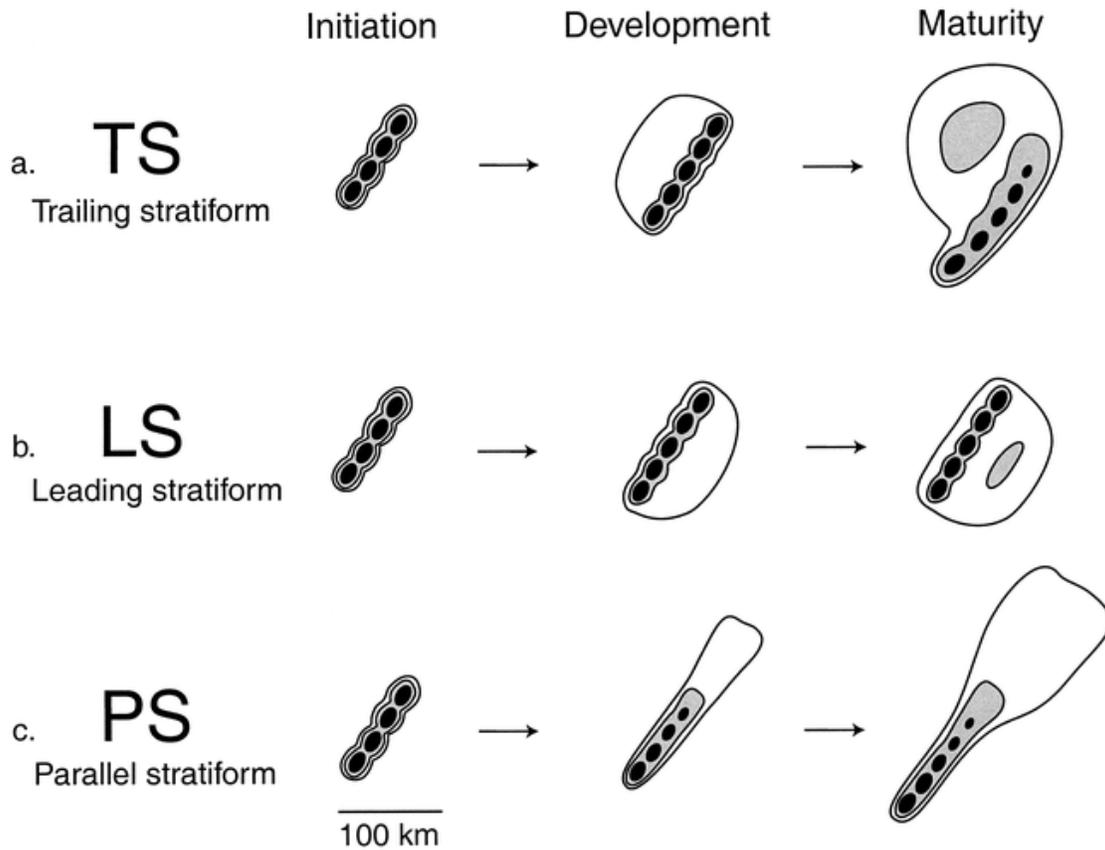
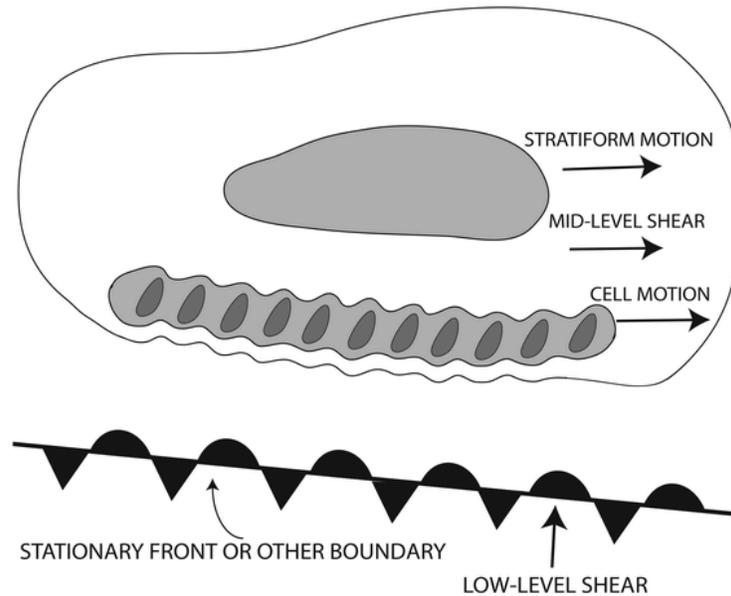


Fig. 1.8. Schematic reflectivity drawing of idealized life cycles for three linear MCS archetypes: (a) TS, (b) LS, and (c) PS. Approximate time intervals between phases: for TS 3–4 h; for LS 2–3 h; for PS 2–3 h. Levels of shading roughly correspond to 20, 40, and 50 dBZ. [Figure 4 and caption reproduced from Parker et al. (2000, section 4).]

A) TRAINING LINE -- ADJOINING STRATIFORM (TL/AS)



B) BACKBUILDING / QUASI-STATIONARY (BB)

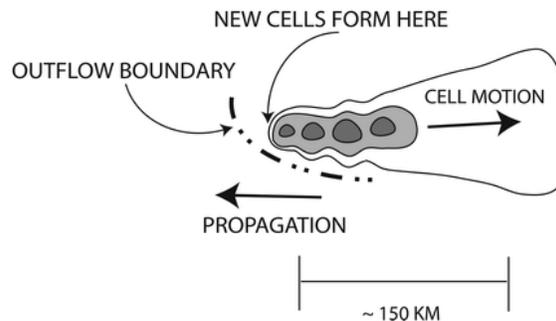


Fig. 1.9. Schematic diagram of the radar-observed features of the (a) TL/AS and (b) BB patterns of extreme-rain-producing MCSs. Contours (and shading) represent approximate radar reflectivity values of 20, 40, and 50 dBZ. In (a), the low-level and midlevel shear arrows refer to the shear in the surface-to-925-hPa and 925–500-hPa layers, respectively. The dash-dot line in (b) represents an outflow boundary; such boundaries were observed in many of the BB MCS cases. The length scale at the bottom is approximate and can vary substantially, especially for BB systems, depending on the number of mature convective cells present at a given time. [Figure 3 and caption reproduced from Schumacher et al. (2005, section 3).]

Back-building Events

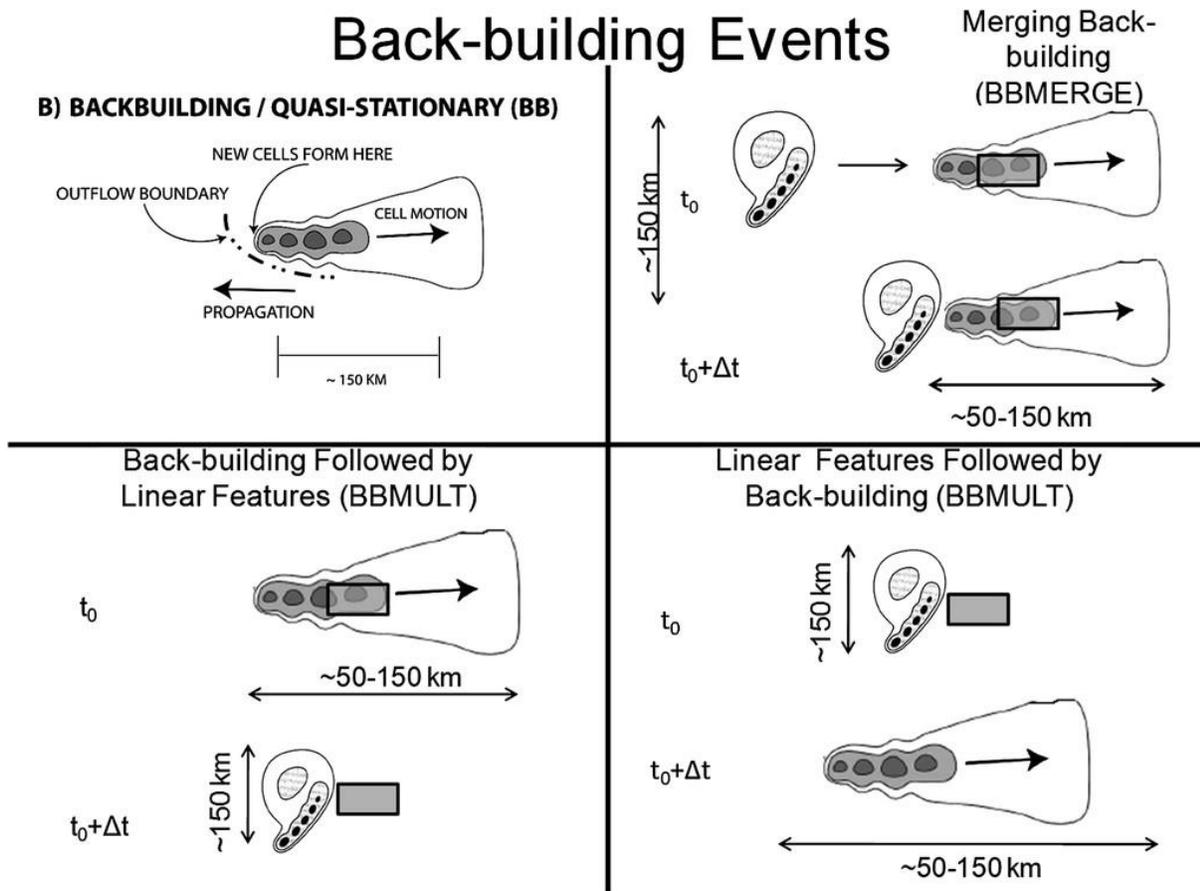


Fig. 1.10. Schematic diagram of the four back-building subclasses of events. Subclasses are determined by the relative motion of back-building cells and MCSs and their relative transits across the flood area. Gray areas indicate the areas where flash flooding is most likely. [Figure 2 and caption reproduced from Jessup et al. (2012, section 3).]

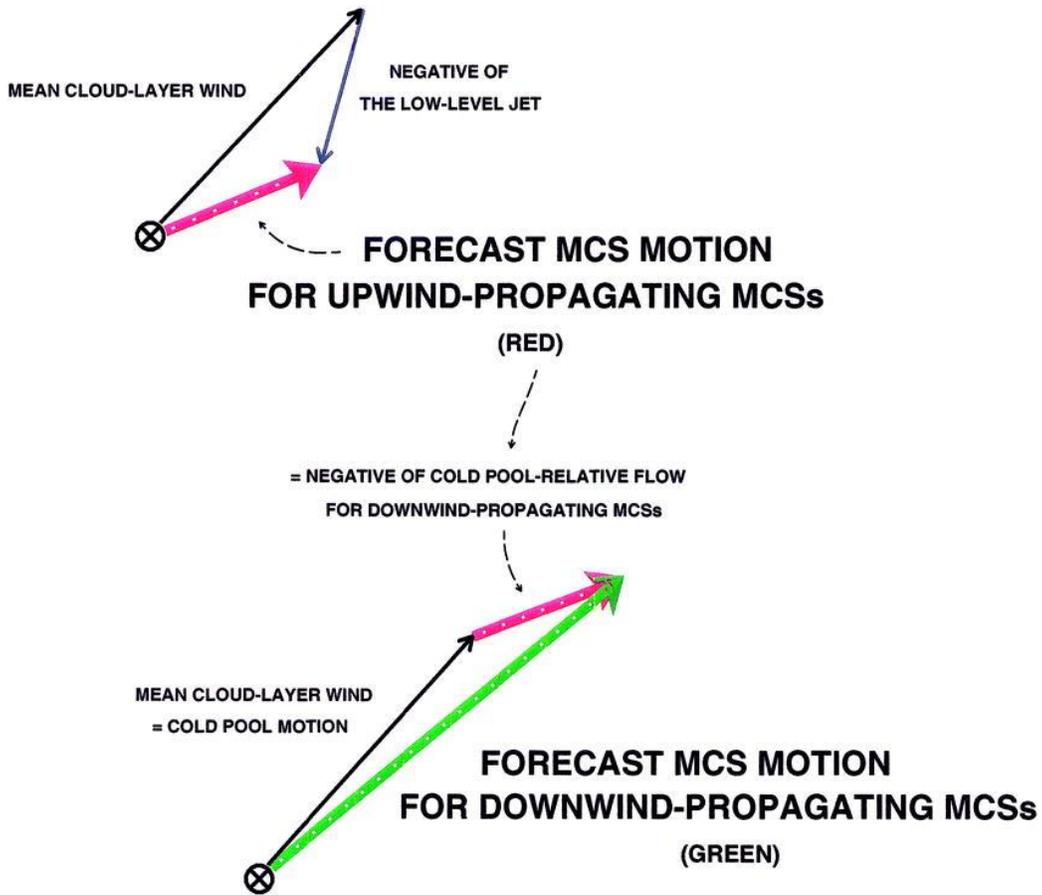


Figure 1.11. Comparison schematics of (top) original (upwind) and (bottom) downwind versions of the vector technique to forecast short-term motion of upwind-developing systems (MCS motion given by thick red arrows) and downwind-developing systems (MCS motion given by thick green arrow at bottom of figure), respectively. Vector lengths are proportional to wind speed; MCS centroids are denoted by the cross symbol. [Figure 6 and caption reproduced from Corfidi (2003, section 3).]

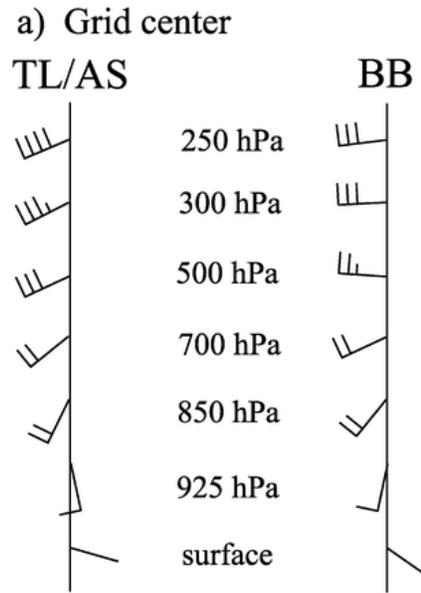


Fig. 1.12. Composite wind profile at (a) the grid center for TL/AS and BB MCS extreme rain events at the peak rainfall time. Wind barbs are plotted conventionally in kt. [Figure 13 and caption adapted from Schumacher et al. (2005, section 4).]

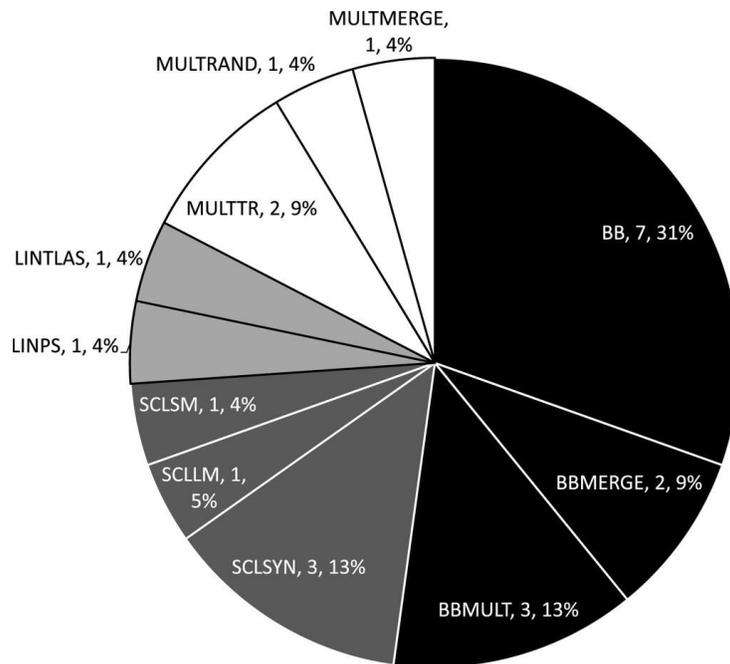


Fig. 1.13. Frequency of northeast U.S. flash flood events, 2003–07, by event type, with radar-estimated rainfall meeting the 24-h precipitation threshold of [Schumacher and Johnson \(2006\)](#). Both the total frequency and the relative frequency (%) are given for these 23 events. [Figure 7 and caption recreated from Jessup et al. (2012, section 4).]

2. Data and Methodology

2.1 Case Selection

Composite radar imagery from June, July, and August of 2015–2020 was analyzed to identify cases with convective storm modes in Upstate New York. The convective cases were then further separated into cases with and without back-building convection in the Hudson Valley. A case was considered to have back building if new convective cells were initiated adjacent to mature cells, creating a continuous radar echo signature which was quasi-stationary or propagated upstream with regards to the low-level flow. If a case exhibited back-building behavior for less than one hour, it was not considered to be a back-building case. The total number of convective cases identified was 70 cases with 15 cases being classified as back-building cases. In all back-building cases, the back-building behavior started between 1500 and 2300 UTC.

Composite radar imagery depicting an example of back-building convection from 23 July 2020 can be found in Figure 2.1. This case was classified as a back-building case, as new cells were initiated upstream of mature cells, causing a propagation down the Hudson Valley between 2000 UTC 23 July and 0000 UTC 24 July. An example of convection without back building from 13 August 2016 can be found in Figure 2.2. This case was not classified as a back-building case because new cells initiated in the vicinity of mature cells but there was no continuous radar echo and no upstream propagation or quasi-stationary behavior.

2.2 WRF Simulations

Operational models with a horizontal resolution of 3 km are able to resolve mesoscale environments induced by interactions between the terrain and background flow (Katona 2016). However, a goal of this research is to investigate the interactions between individual convective cold pools and channeled flow, necessitating the use of finer resolution than is used in current operational models. The Weather Research and Forecasting (WRF) Model version 4.03 was used to simulate cases of back-building convection in the Hudson Valley. A triple-nested domain configuration (Figure 2.3) was used with horizontal resolutions of 9 km, 3 km and 1 km respectively and 51 vertical levels with an upper boundary at 50 hPa. The resolved terrain in the inner-most domain, as seen on Figure 2.4, represents the valleys and mountains in Upstate New York well. The inner-most domain was centered over the eastern Catskill Mountains, the most common location for convective initiation, and the Hudson Valley, in order to focus on back-building convection in the Valley.

The WRF simulations were initialized and updated with boundary conditions using HRRR 0-hour analyses (a description of the HRRR can be found in section 2.3). The simulations were initialized at 0600 UTC, roughly twelve hours before the back-building convection occurred in the Hudson Valley. This initialization time was chosen to allow the model to spin up the mesoscale environments and channeled flow caused by the terrain before any convection formed. All simulations were run with a physics package that mimics the physics from the operational HRRR, which is often used to forecast the location and mode of convective storms. The Thompson ice-bearing microphysics scheme with the Mellor-Yamada-Nakanishi-Niino (MYNN) 2.5 level TKE PBL and MYNN surface layer schemes were utilized in all domains, while the Grell-Freitas

cumulus parameterization scheme was used on the coarsest domain. Variations in microphysics schemes were tested, but did not produce the convective evolution observed in composite radar imagery. The Rapid Update Cycle (RUC) land surface model was used, the same land surface model found in the operational HRRR.

2.3 HRRR Analysis Data

As part of NOAA's Collaborative Science, Technology and Applied Research (CSTAR) program, one of the goals of this research is to provide forecasters with information that will be useful in an operational setting. While gaining a deeper understanding of the physical mechanism responsible for back-building convection in the Hudson Valley (discussed in chapter 3) is useful for forecasters, a description of the environment associated with back building would also be useful. More specifically, the relevant questions are: What makes a pre-convective environment conducive to back-building convection? Are there threshold values or spatial patterns for specific environmental parameters that could help forecasters assess the risk of back-building convection in the Hudson Valley?

The HRRR 0-hour analyses was chosen to represent the pre-convective environment within the Hudson Valley on convective days, as HRRR model output is commonly used for forecasting convective storms in an operational setting. The operational HRRR, an hourly updated convection-allowing model (CAM) with 3 km grid spacing, has been run since 30 September 2014. The HRRR model is able to resolve mesoscale features, including both terrain influences (Katona 2016) and convective storms. Archived HRRR 0-hour analyses dating back to September 2014 were obtained from the Google Cloud Platform HRRR archive

(<https://console.cloud.google.com/marketplace/product/noaa-public/hrrr?project=python-232920&pli=1>).

The 1800 UTC 0-hour HRRR analysis surface files were collected for June, July and August of 2015–2020. These months were chosen because they have the highest frequency of high impact and missed severe weather events in the Northeast (Figure 2.5) and the highest number of storm reports (Hurlbut and Cohen 2014). The 1800 UTC analyses were chosen because flash flood events in the Northeast are maximized between 1500 UTC and 2100 UTC (Jessup 2012), and the first severe thunderstorm report on severe weather days in the Northeast is most commonly between 1800 and 2000 UTC (Hurlbut and Cohen 2014).

Area averages of several thermodynamic variables and severe weather indices within the Hudson Valley were calculated. The Hudson Valley was defined to be the area between latitudes 41.7 and 42.8 N and longitudes 73.55 and 74.0 W for this analysis (Figure 2.6). The area includes the portion of the Hudson Valley roughly between Troy, New York in northern Albany County and Poughkeepsie, New York in Dutchess County. This area was chosen as it is the portion of the Hudson Valley with the highest terrain features on either side, making it conducive for terrain-channeled flow.

A complete list of variables taken or calculated from the HRRR 0-hour analyses can be found in table 1. The variables that were taken directly from analyses include the wind components at 10 m above ground level (AGL), 850 hPa, and 500 hPa; the components of vertical wind shear between the surface and 1 km and 6 km AGL respectively ; the temperature and dewpoint at 2 m AGL, 850 hPa, and 500 hPa; and the surface-based CAPE (SBCAPE).

Dew point depression (DD) is defined as the difference between the environmental temperature and dew point (T_d) at a given level. DD was calculated at 2 m AGL, 850 hPa and 500 hPa. DD is used as a measure of the environmental saturation with respect to water vapor.

$$DD = T_{level} - T_{d_{level}}.$$

The Lifted Index (LI) is defined as the difference between the environmental temperature at 500 hPa and the temperature of a parcel lifted adiabatically from a certain level. The LI is often calculated using a parcel lifted from the surface while the Best LI, which was used in this study, is taken as the lowest value computed using parcels lifted from between the surface and 850 hPa. Lower values of LI indicate higher instability, with values between -3 and -9 representing moderate instability and values below -9 representing extreme instability.

$$LI = T_{env\ 500\ hPa} - T_{parcel\ 500\ hPa}.$$

The Total Totals (TT) is the sum of the Vertical Totals (VT), a measure of lapse rate in the mid troposphere, and Cross Totals (CT), a measure of the moisture content in the mid troposphere. TT has been used to assess the mid-level environmental conditions for thunderstorm occurrence. A value above 45 means thunderstorms are possible, while values above 55 indicate severe thunderstorms are more likely.

$$TT = VT + CT.$$

$$VT = T_{850\ hPa} - T_{500\ hPa}.$$

$$CT = T_{d_{850\ hPa}} - T_{500\ hPa}.$$

The mean sea-level pressure (MSLP) included in the HRRR 0-hr analyses is calculated using the Mesoscale Analysis and Prediction System (MAPS) sea-level pressure algorithm (Benjamin and Miller 1990). The MAPS sea-level pressure uses the temperature at 700 hPa to

estimate the temperature of the fictitious layer of air beneath the ground in order to calculate the thickness. This assumption makes the MAPS system a useful method for locations with high terrain, but hurts its accuracy over low terrain, where a measured or simulated temperatures below 700 hPa could be used. MSLP differences were taken between the northern-most and southern-most extent of the Hudson Valley area (Figure 2.6), such that positive values indicate higher MSLP to the south.

There were several variables, potentially relevant for back-building convection, that were not included in the analysis as they could not be calculated using the available fields in the HRRR surface files, including DCAPE and MUCAPE. Variables used to represent the strength of the convective cold pool were not included as it is unlikely that they were accurately represented in the HRRR surface files, including cold pool depth, pressure and temperature change across the gust front and horizontal vorticity along the gust front.

2.4 Decision Trees

Decision trees were used to compare the pre-convective environment from cases with and without back building. Decision trees are a type of flowchart used for testing the likelihood of a specific outcome based on sequential decisions (Figure 2.7). The typical decision tree terminology labels the variable being tested at each node as a feature. The value to which each feature is being compared is the threshold, and the outcome being predicted at each node is the target. The success of each leaf node is assessed by its impurity (defined below) with lower impurity corresponding to higher predictive skill.

$$\text{Impurity} = \frac{\text{Number of targets incorrectly predicted}}{\text{Number of targets}}.$$

Decision tree classifier algorithms were run using the Scikit-Learn software package. The decision tree classifier uses sequential feature selection (SFS), an iterative process in which each feature is greedily chosen from the set of features in order to eliminate the maximum number of cases with a negative outcome. The algorithm is greedy in that each feature chosen eliminates the most negative cases possible at that step without taking the full selection process into account. An example of a split made by SFS can be found in Figure 2.8. The decision tree classifier was also run with increased weights on back-building cases, which was achieved by creating multiple cases with the same values for each feature. The total set of cases ($n=70$) was randomly separated into training sets ($n=52$), which were used to train the decision trees, and testing sets ($n=18$), which were used to test the predictive skill of the decision trees. The decision tree classifier was run on 5000 random training sets to reduce the dependence on which cases were chosen.

The decision tree classifier was run on several different sets of features. A shorter list was initially used (short list in Table 1), which focused mainly on surface thermodynamic variables and typical severe weather forecasting variables including surface-based CAPE (SBCAPE) and vertical wind shear. The list of variables was later expanded to include more upper-air thermodynamic variables and severe weather indices, such as LI (expanded list in Table 1).

The decision tree classifier was run on the list of variables after they had been manipulated in several different ways to test to the robustness of results. First, the decision tree algorithm was run with different units for thermodynamic variables (SI, imperial and metric). Next, the values of each feature was normalized to have a mean of zero and a variance of one. Lastly, a principal component analysis (PCA) was used to reduce the dimensionality of the dataset. The PCA was performed after the dataset was rotated to cluster linear combinations of

separate features, resulting in thirty distinct factors. The decision tree algorithm was run on the first sixteen factors, which explained 33% of the variance. Though the resulting thresholds from each test varied, none of the changes yielded a qualitative difference in the environmental characteristics represented by the chosen features.

Variable	Used in short list	Used in expanded list	Taken from analysis	Calculated from analysis
10-m zonal wind	x	x	x	
10-m meridional wind	x	x	x	
10-m wind magnitude	x	x		x
10-m wind direction	x	x		x
850-hPa zonal wind		x	x	
850-hPa meridional wind		x	x	
850-hPa wind magnitude		x		x
850-hPa wind direction		x		x
0-1 km zonal shear		x	x	
0-1 km meridional shear		x	x	
0-1 km shear magnitude	x	x		x
0-1 km shear direction	x	x		x
0-6 km zonal shear		x	x	
0-6 km meridional shear		x	x	
0-6 km shear magnitude	x	x		x
0-6 km shear direction	x	x		x
2-m absolute humidity		x	x	
2-m relative humidity		x		x
2-m dew point depression	x	x		x
2-m dew point	x	x	x	
2-m temperature		x	x	
850-hPa dew point depression	x	x		x
850-hPa dew point	x	x	x	
850-hPa temperature		x	x	
500-hPa dew point depression	x	x		x
500-hPa dew point	x	x	x	
500-hPa temperature		x	x	
SBCAPE	x	x	x	
Lifted index		x		x
Total totals		x		x

Table 1. Variables populating the sets of features used in decision tree analyses. Variables marked as “taken from analysis” are fields found in 1800 UTC HRRR 0-hour analyses surface files, while variables marked as “calculated from analysis” were calculated using fields found in the surface files.

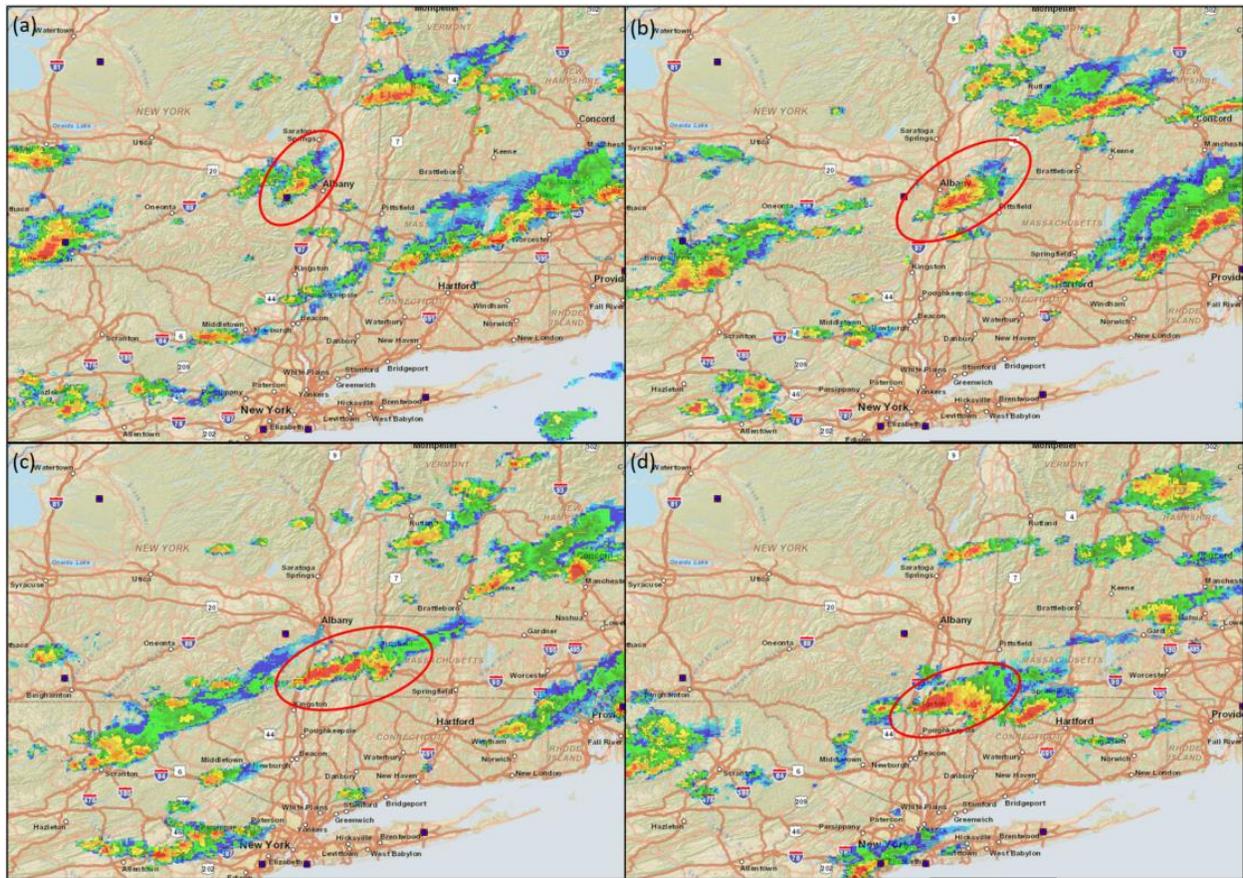


Fig. 2.1. Composite radar imagery showing an example of a back-building MCS (circled in red) in the Hudson Valley from 23 July 2020 at (a) 1930 UTC, (b) 2030 UTC, (c) 2155 UTC and (d) 2325 UTC.

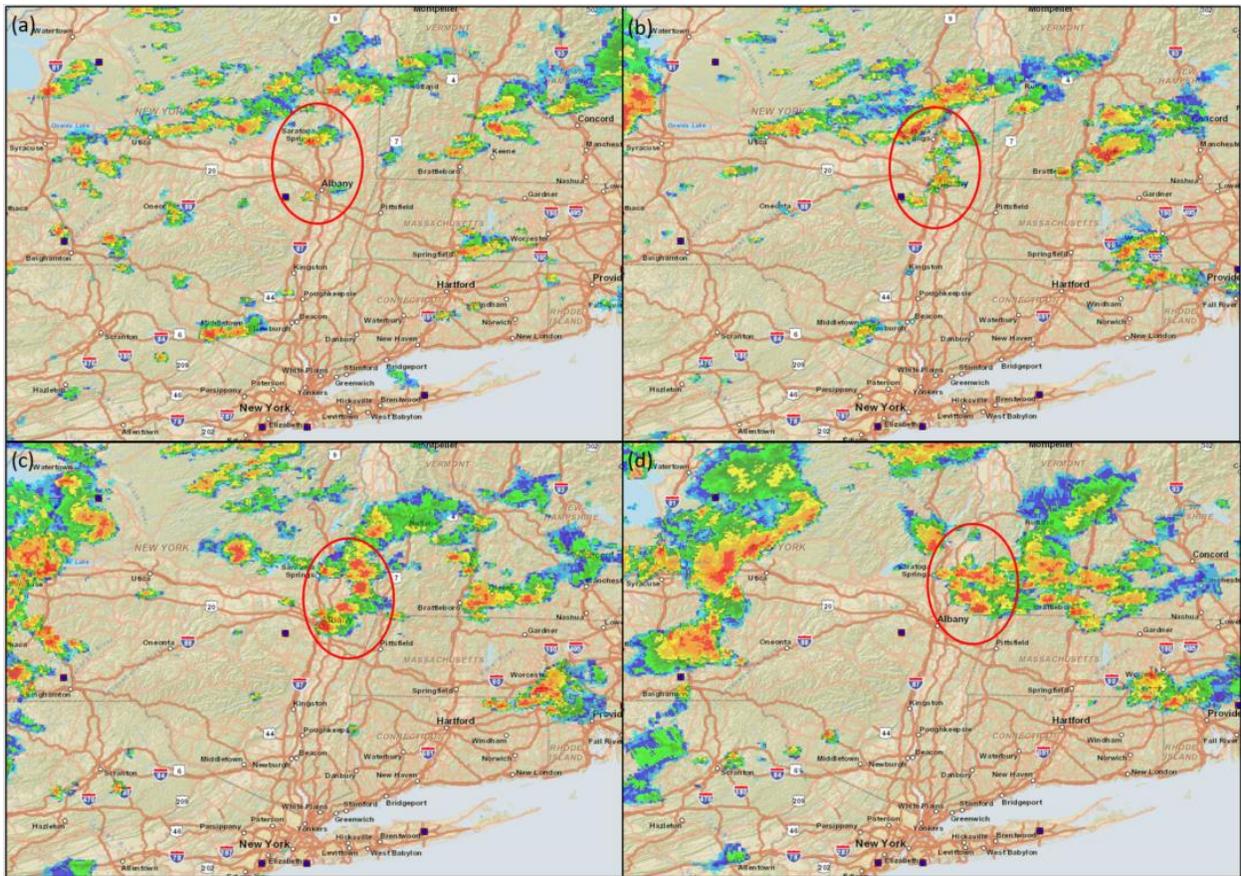


Fig. 2.2. Composite radar imagery showing an example of convection without back building (circled in red) in the Hudson Valley from 13 August 2016 at (a) 1830 UTC, (b) 1930 UTC, (c) 2020 UTC and (d) 2120 UTC.

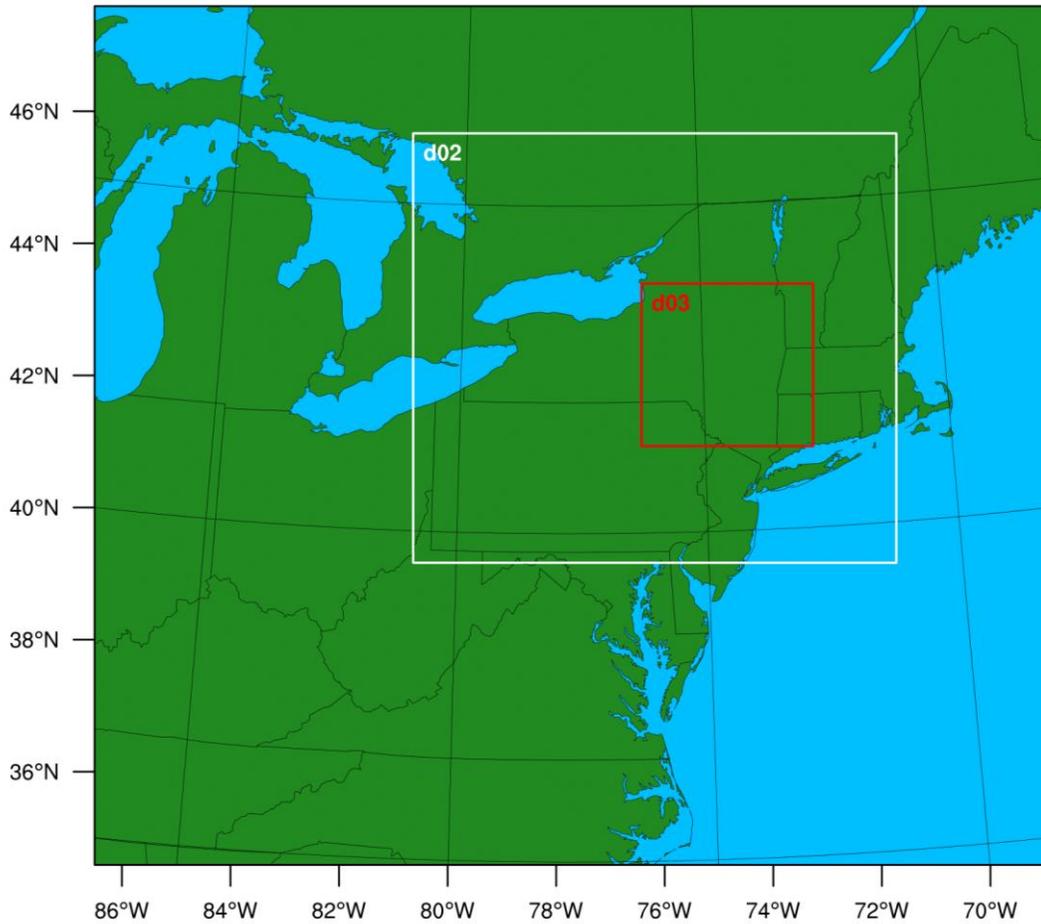


Fig. 2.3. Location of model domains used in WRF simulations of back-building convection. The inner most domain is centered over the Catskill Mountains, a common convective initiation location, and the Hudson Valley, a common back-building location. The horizontal resolution used in each domain is 9, 3 and 1 km respectively.

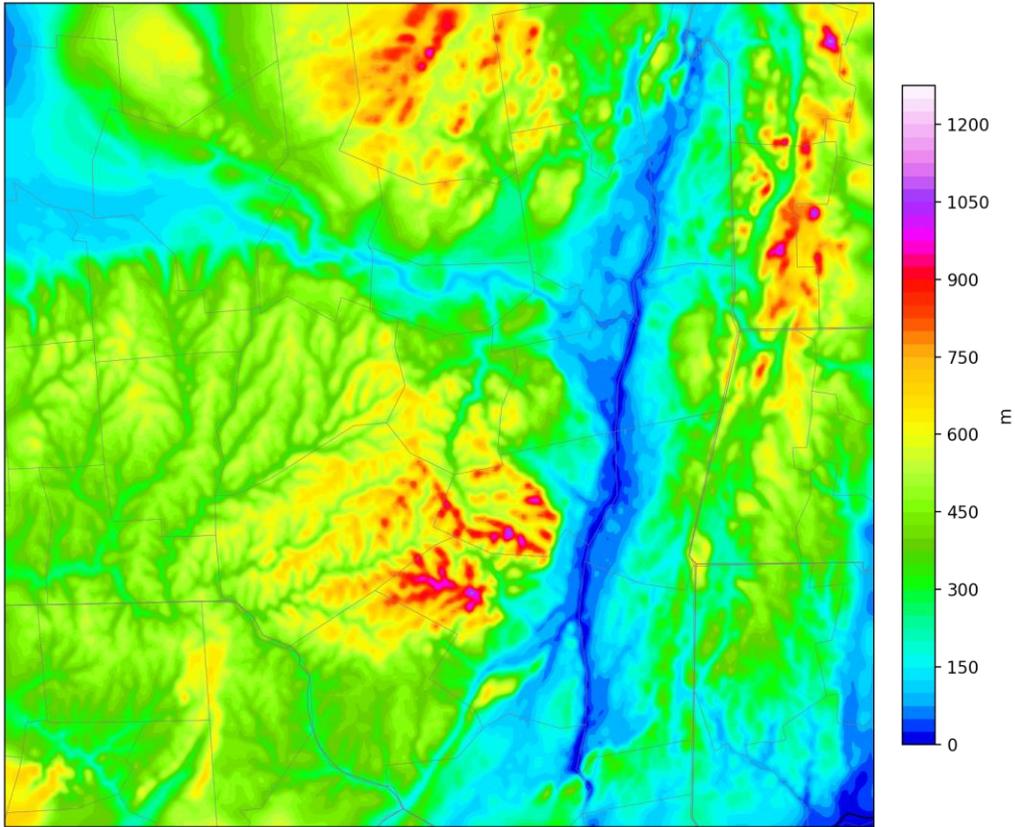


Fig. 2.4. Map of terrain height from the inner-most domain of the WRF simulations. Terrain height is filled at an interval of 25 m and county borders appear in grey.

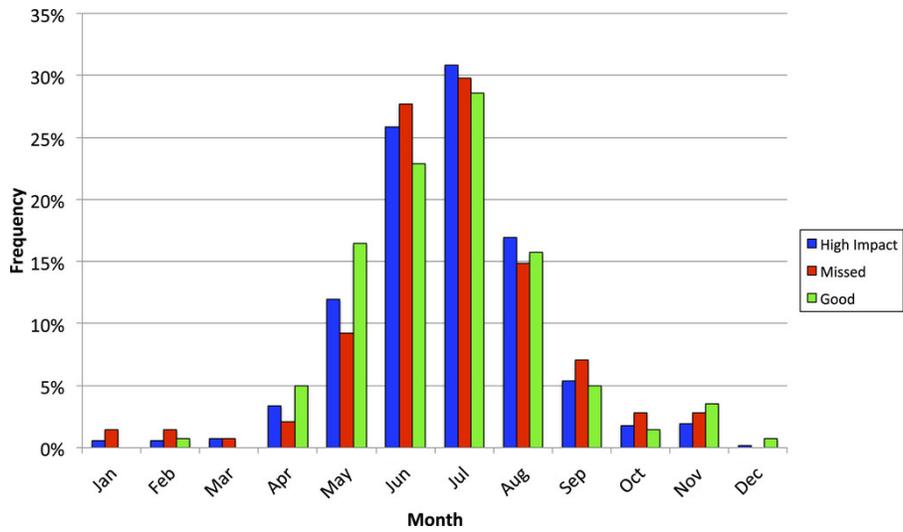


Fig. 2.5 Monthly frequencies of high-impact (blue), missed (red), and good (green) events from 1980 to 2013. [Figure 6 and caption reproduced from Vaughan et al. (2016, section 3).]

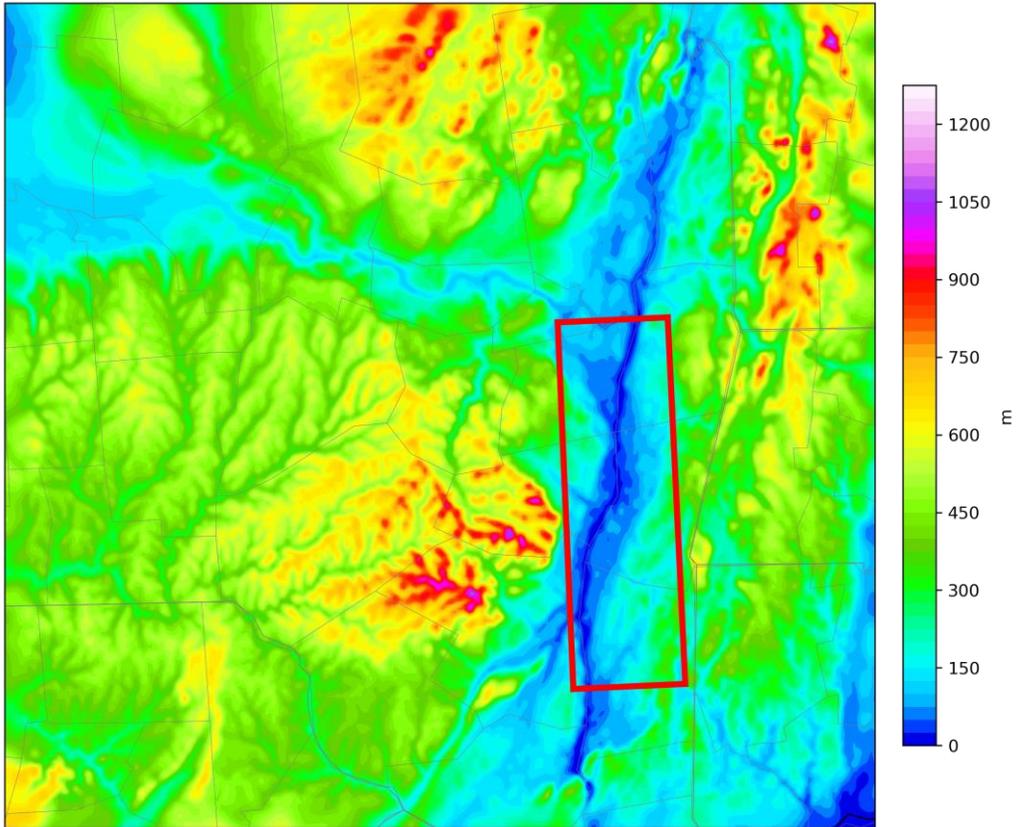


Fig. 2.6. Map of terrain height from inner-most domain of WRF simulations. Terrain height is filled at an interval of 25 m and county borders appear in grey. Red box denotes the area considered to be the Hudson Valley for area averages.

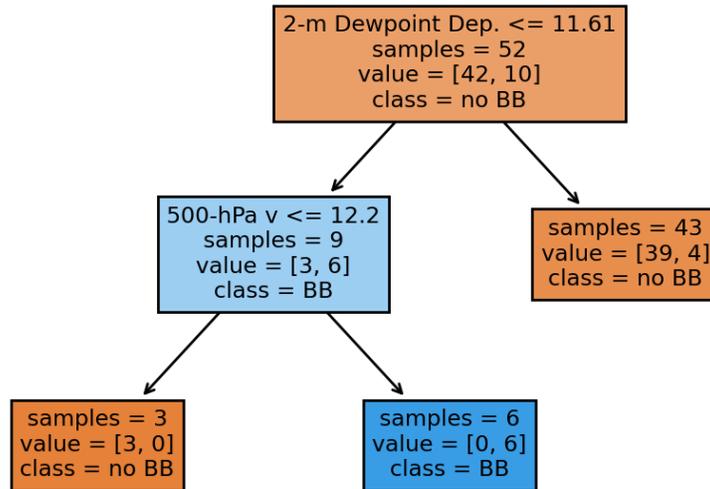


Fig. 2.7. Example decision tree that can provide guidance on the likelihood of back-building convection based on pre-convective parameters. Cases that have values lower than the threshold at each decision node enter the next lowest node on the left, while values higher than the threshold go to the right.

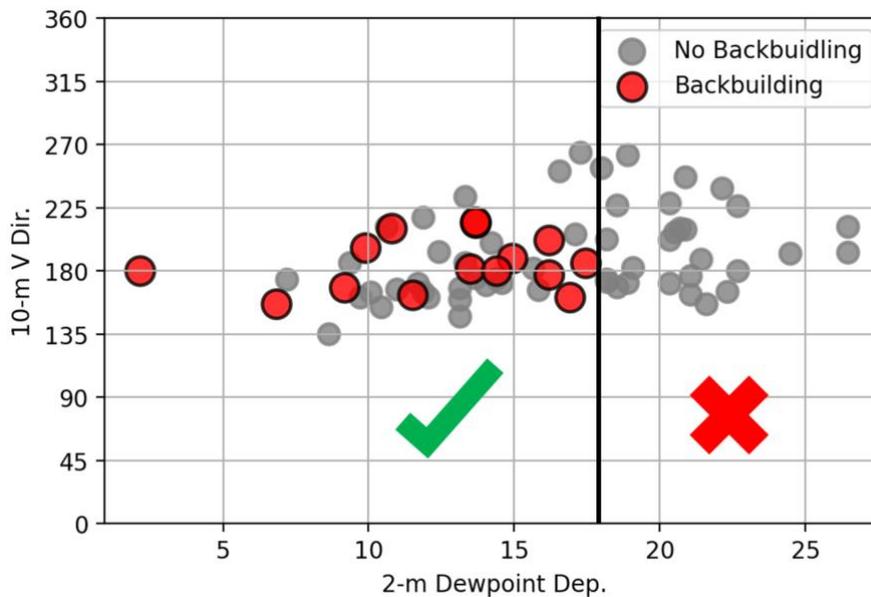


Fig. 2.8. Parameter phase space of 2-m dew point depression and 10-m wind direction. The thick black line shows an example of a threshold used in sequential feature selection to separate BB and no-BB cases. Cases with a DD below 17.5 would be kept and all other cases removed from the dataset.

3. Physical Mechanisms

3.1 Case Study 21 August 2019

3.1.1 Case Overview

On 21 August 2019, a discontinuous line of convective cells developed over the Mohawk Valley and northern Catskill Mountains at 1700 UTC, following several rounds of weaker precipitation passing through Upstate New York between 0700 UTC and 1500 UTC. The line moved east into Saratoga County, where the convection grew upscale to an MCS at 1900 UTC (Figure 3.1). The MCS began producing new cells along its southwestern edge, which travelled parallel to the gust front as they matured, allowing the MCS to back build. The MCS continued to back build as it propagated to the south within the Hudson Valley, until 2200 UTC, at which time the convective line moved off to the southeast.

Several watches and warnings were issued by the Storm Prediction Center (SPC) and Albany Weather Forecast Office (WFO) during this event. The SPC issued a severe thunderstorm watch for much of eastern New York and western New England at 1753 UTC, just as the convection was forming in the Mohawk Valley. A flood advisory was issued by the Albany WFO at 2004 UTC for Saratoga County and a series of severe thunderstorm warnings were issued between 1845 UTC and 2032 UTC (Figure 3.2). All severe thunderstorm warnings were issued with polygons that extended to the east of the back-building portion of the convective line. Each successive polygon was issued further to the southeast of the convective line as the propagation down the Hudson Valley became more apparent on radar.

The back-building MCS created multiple types of hazards in the Hudson Valley, with the most impactful being heavy rainfall. The back-building cells produced a localized precipitation

maxima of over 50 mm in Saratoga County as reported by the National Center for Environmental Protection's (NCEP) Stage IV Quantitative Precipitation Estimate (Figure 3.3). There were reports of flash flooding, damaging winds, and hail in Saratoga, Albany, Rensselaer, Greene, and Columbia Counties. Damaging winds were reported in Saratoga County at 1935 UTC and Greene County at 2140 UTC, making them chronologically consistent with the propagation of the back-building MCS (Figure 3.4). There were three tornadoes reported in the Mohawk Valley, Saratoga, and Bennington Counties on 21 August 2019, but none were in connection to the back-building MCS in the Hudson Valley.

3.1.2 Synoptic Environment

The synoptic environment observed during this case was representative of a pattern of weak synoptic forcing seen for much of August 2019. The polar jet stream was displaced northwards into southern Canada, leaving much of the United States in weak upper-level flow. A closed upper-level low had formed just south of Hudson Bay with a 300-hPa jet streak in the base of the trough over the western Great Lakes. To the east, large zonal ridge extended into the northern Atlantic (Figure 3.5). The trough and jet streak were too far upstream to provide strong upper-level forcing for ascent over eastern New York. A 500-hPa shortwave trough I passed over eastern New York at 1800 UTC, likely contributing to the initiation of the discrete convection through positive vorticity advection (forcing for ascent).

The shortwave trough is also evident at 700 hPa, as there is a pronounced westerly to southerly wind shift across the trough axis. The precipitable water field shows that there was higher moisture content along the East Coast than the interior United States, extending up into

eastern New York and New England, coinciding with the southerly winds. The temperature gradient at 850hPa in the Northeast was very weak, with slight warm-air advection into the New England. The NOAA Weather Prediction Center (WPC) 1800 UTC surface analysis indicated a surface low of 1009 hPa over Lake Huron (Figure 3.6). A cold front extended westward into Michigan, and a warm front extended across northern New York through the Adirondacks and northern New England. This left much of upstate New York in the warm sector of the cyclone, where observed surface pressure was around 1011 hPa.

3.1.3 Local Environment

Though the Northeast had a somewhat homogeneous large-scale environment within the warm sector, there were notable inhomogeneities on smaller scales. Figure 3.6 is a Weather Prediction Center (WPC) surface analysis valid at 1800 UTC 21 August 2019. The 2-m T_d , as reported on the surface analysis, were in the mid-70's in the Hudson Valley and New England. New York State Mesonet (NYSM) station in the Hudson Valley reported even higher T_d of 74°-75°F, while stations in the Catskills to the west reported T_d in the high 60s. Figure 3.7 is four meteograms showing how the local environment changed leading up to the event at the Schodack NYSM station, to the southeast of Albany, NY. The meteograms show T_d steadily increased between 0900 UTC and 1500 UTC, likely in connection to the passing showers in the morning. A final increase in T_d occurred right after 1500 UTC, coinciding with a change to southerly winds at the surface, representative of the background flow being channeled up the Hudson Valley (discussed in section 3.2.3). The high moisture content found in the Hudson Valley made

it a more favorable thermodynamic environment for convection than the Mohawk Valley and Catskill Mountains to the west.

3.2 WRF Simulation

3.2.1 Comparison with Observed Case

The WRF simulation was initialized at 0600 UTC on 21 August 2019. This initialization time was chosen as it allowed the model twelve hours to spin up before the upscale growth of convection was expected to occur. During that twelve hours, the model developed much of the same localized environments and flow characteristics found in the HRRR 0-hr analyses depicting the case. Figure 3.8 shows area-averaged soundings of the Hudson Valley, as defined in Figure 2.6, from the WRF simulation and the HRRR 0-hr analysis respectively. The soundings are valid at 1800 UTC, which was roughly an hour before the first convection started forming in the Hudson Valley. The HRRR sounding has a surface T_d of 23°C and a nearly dry adiabatic lapse rate at 850–900 hPa, resulting in an area average SBCAPE of 2037 J/kg. The wind profile has nearly unidirectional southwesterly winds with weak surface winds and a 0-6 km shear value of 17.8 m/s. The WRF sounding has a surface T_d of 20°C and a conditionally unstable lapse rate at 700–900 hPa. These small changes near the surface produce a higher lifted condensation level and an area average SBCAPE of 1232 J/kg. The wind profile shows veering between the surface and 600 hPa, indicative of warm air advection, and strong southerly surface winds created by terrain channeling (discussed in section 3.2.3). The upper-level winds are also stronger in the simulated wind profile, resulting in a 0-6 km shear of 22 m/s.

Despite the differences in the pre-convective environment found in the simulated case, the observed convective evolution was sufficiently represented in the simulation. The simulation reproduced the initiation site along the Mohawk Valley at 1730 UTC but initiated the convection over the Catskill Mountains earlier than it was observed (not shown). Between 1730 UTC and 1900 UTC, the simulation produced discrete convective cells over southeastern New York. Figure 3.9 shows the simulated reflectivity from the WRF simulation alongside the observed reflectivity at 1855 UTC and 2100 UTC. The overall extent of the convection is greater in the simulation, and the cells did not organize into a squall line in the eastern mouth of the Mohawk Valley, as seen in observed radar. The observed convection started back building in the Hudson Valley at 1925 UTC. The simulated reflectively also produced discrete convective cells in the upper Hudson Valley at 1925 UTC, but had not started initiating new cells along their outflow. The simulated convection started back building in the Hudson Valley by 1940 UTC (not shown).

Once the simulated convection started back building, the evolution was similar to the composite radar imagery evolution. An MCS formed, oriented from southwest to northeast across the Hudson Valley with an area of stratiform precipitation over the Taconic Mountains to the northeast (Figure 1.1). New convective cells initiated on the southwestern edge of the existing cells and were trained parallel to the line. The simulated line propagated down the Hudson Valley, continuing to back build until 2130 UTC. Once the weak line of convection over the Catskills entered the Hudson Valley, it conglomerated with the MCS, and the squall line moved to the east over the Taconic Mountains.

Figure 3.10 shows the simulated and observed total precipitation between 1500 UTC and 2300 UTC 21 August 2019. The simulated and observed total precipitation maxima have a similar orientation but the simulated precipitation has a smaller spatial extent. The difference in

spatial extent may be due to the simulation having a horizontal resolution of 1 km compared to the observed precipitation's 4-km resolution. The local maxima of precipitation are oriented from southwest to northeast, parallel to the cell motion within the back-building MCS. This signature is consistent with observed total precipitation fields from other back-building MCS cases, where cells are initiated and trained over the same location for an extended period of time. The WRF simulation was able to recreate the back-building convection in the Hudson Valley as observed on the composite radar imagery. The remaining sections in this chapter will focus on how the simulated cold pool interacted with the terrain and low-level flow to initiate new convective cells, and how the background flow being channeled by the terrain influenced the pre-convective environment and convective evolution. Investigating these mesoscale processes will further the understanding of how back-building convection is able to sustain itself in the Hudson Valley.

3.2.2 Convective Cold Pool

The process causing the upstream propagation of a back-building MCS is driven by the initiation of new cells along the outflow boundary of existing convection. The dynamics involved with initiating those new cells is largely driven by the interaction between the convective cold pool's gust front and the low-level flow. In the case of the back-building MCS that occurred on 21 August 2019, examining how the gust front formed and interacted with the southerly terrain channeled flow within the Hudson Valley can help one understand the simulated convective evolution.

The discrete convection that grew upscale into the back-building MCS was not initiated in the Hudson Valley, but rather over the eastern Catskill Mountains. Figure 3.11 shows the simulated reflectivity, equivalent potential temperature (θ_e), vertical motion and winds at 150 m from 1855 UTC, about 45 minutes before the back building began. There is a weak line of confluence over the eastern Catskills, where high- θ_e air, transported northward by the southerly winds, is meeting the lower θ_e air advected from the southwest. Discrete cells are initiated along this line between 1855 and 1910 UTC, but do not a strong cold pool until they have descended off of the Catskills. Once the convective cells enter the Hudson Valley, they intensify in the favorable thermodynamic environment, forming a distinct cold pool seen as a bull's eye of low θ_e in northern Albany County (Figure 3.11e).

Once the cold pool is formed, the convection starts to back build on the western edge of the Hudson Valley. The strongest ascent at 1940 UTC appears on the southern edge of the θ_e gradient, allowing new convective cells to initiate just to the southwest of the existing cells (Figure 3.12). The reflectivity expands parallel to the gust front, growing upscale into an MCS by 2000 UTC. At this time, there is a region of upward vertical motion on the southeastern edge of the gust front, showing the updrafts of the convective cells, and downward vertical motion behind the gust front representing the downdraft of evaporatively cooled air. Figure 3.13 shows a cross section of the potential temperature (θ) and vertical motion (solid lines indicating ascent) running south to north across the MCS. The upward vertical motion seen in this cross section is associated with the updrafts of mature convective cells on the northern end of the MCS. The updraft starts at about 400 m above the ground, just ahead of the convective cold pool, and tilts back over the cold pool with height. The aforementioned downdraft also appears on this cross section on the cold side of the gust front.

The MCS strengthens further while propagating down the Hudson Valley between 2000 UTC and 2100 UTC (Figure 3.14). As the cold pool continues to be reinforced with evaporatively cooled air, the orientation of the gust front rotates to become more perpendicular to the valley. This west-to-east orientation further increases the convergence along the gust front, as the northerly winds within the cold pool meet the southerly flow from the lower Hudson Valley. The depth of the cold pool (~500 m) stays relatively constant during this time while the volume of the coldest air (< 294 K) increases (Figure 3.15). However, the strongest vertical motion is no longer directly above the coldest portion of the cold pool, but rather located at the leading edge of the cold pool, where the strongest θ gradient is located. This change suggests that the temperature gradient along the leading edge of the gust front may be a more important variable for lifting potential than the coldest temperature found in the cold pool.

The back building ultimately ends at 2120 UTC, when a separate line of convection enters the Hudson Valley south of the MCS (Figure 3.16). This new line of convection in the southern Hudson Valley intensifies when it enters the favorable thermodynamic environment and forms its own cold pool. This new cold pool cuts off the channeled flow coming up the valley, which was acting as the inflow for the back-building MCS. Without a continued source of high θ_e air to be lifted over the gust front, new cells were no longer initiated upstream of the existing cells, which move off to the east and out of the Hudson Valley.

3.2.3 Terrain Channeling

An important feature of the WRF simulation was the presence of southerly terrain channeling in the Hudson Valley (Figure 3.17). The terrain channeling appeared in the

simulation starting at ~1600 UTC and continued until it was disrupted by the second line of convection entering the lower Hudson Valley (~2100 UTC). The channeling was represented by a decoupling between the winds within the valley for the larger-scale flow seen above the valley. When the terrain channeling was at its strongest, the wind speed within the valley had two distinct maxima of 12 m/s and 13 m/s and the above-valley wind was 9 m/s. The synoptic scale flow in Upstate New York was south-southwesterly but the flow within the valley was southerly, with even a slight easterly component on the western side.

The increased southerly flow in the Hudson Valley played an important role in creating the favorable thermodynamic environment upstream of the back-building MCS. Figure 3.18 shows zonal and meridional cross sections of the horizontal water vapor flux and meridional winds in the Hudson Valley at 1800 UTC, about an hour before the back building began. There is increased vapor flux within the upper Hudson Valley at this time, which is collocated with the maximum in southerly winds created through channeling. This vapor flux up the Hudson Valley caused the 2-m T_d to increase from 69.4 °F at 1600 UTC to 72.1 °F at 1800 UTC, allowing the low-level instability to increase before the back building occurred.

Figure 3.19 shows the vapor flux cross sections at a more southern location, ahead of the back-building MCS while it was at its most intense (2055 UTC). The meridional wind has become stronger by this time, resulting in an increased horizontal vapor flux upstream of the MCS. The maxima in horizontal vapor flux descends as it approaches the gust front, collocating with the strongest vertical motion as seen on Figure 3.15. It was ultimately this moist southerly flow that was lifted over the gust front to initiate new convective cells in the back-building MCS.

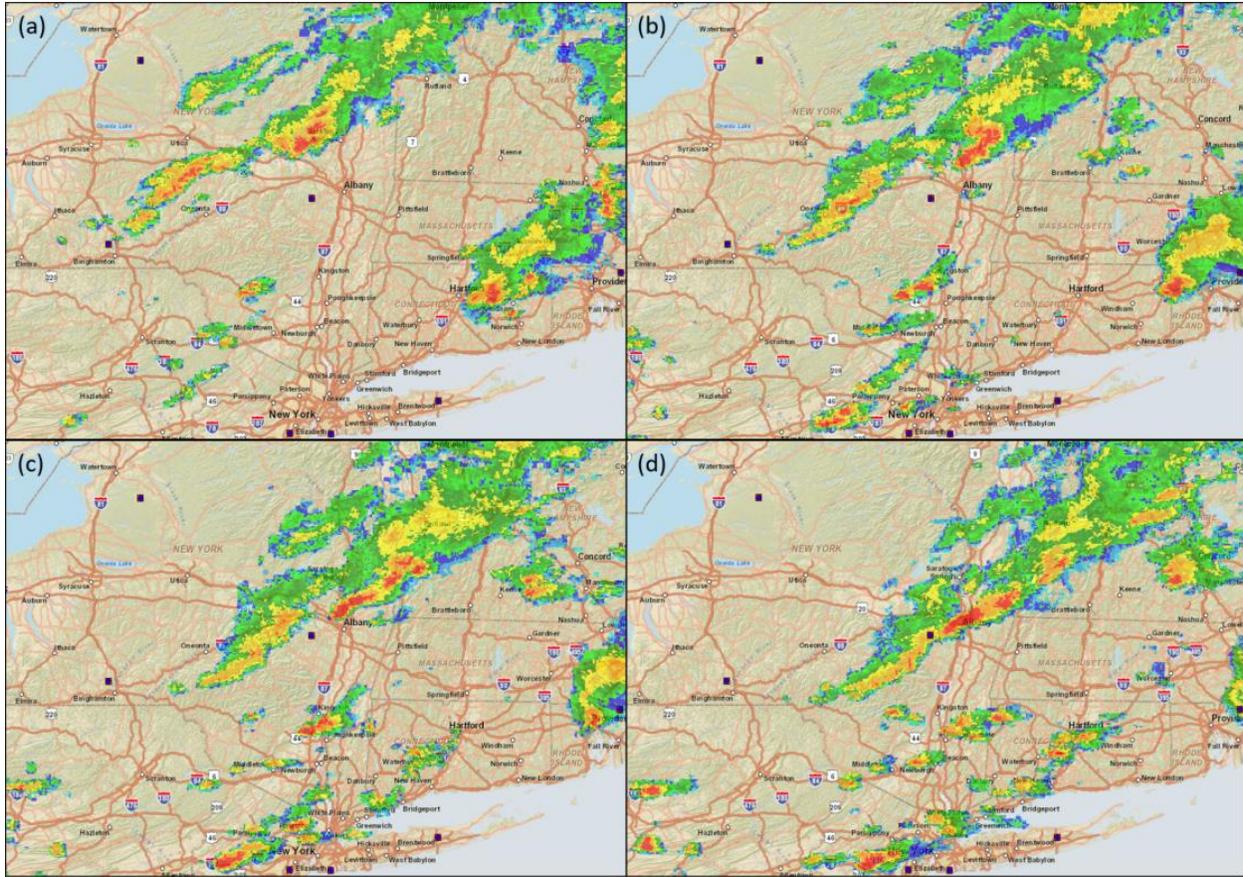


Fig. 3.1. Composite radar imagery showing a back-building MCS in the Hudson Valley from 21 August 2019 at (a) 1855 UTC, (b) 2000 UTC, (c) 2030 UTC and (d) 2100 UTC.

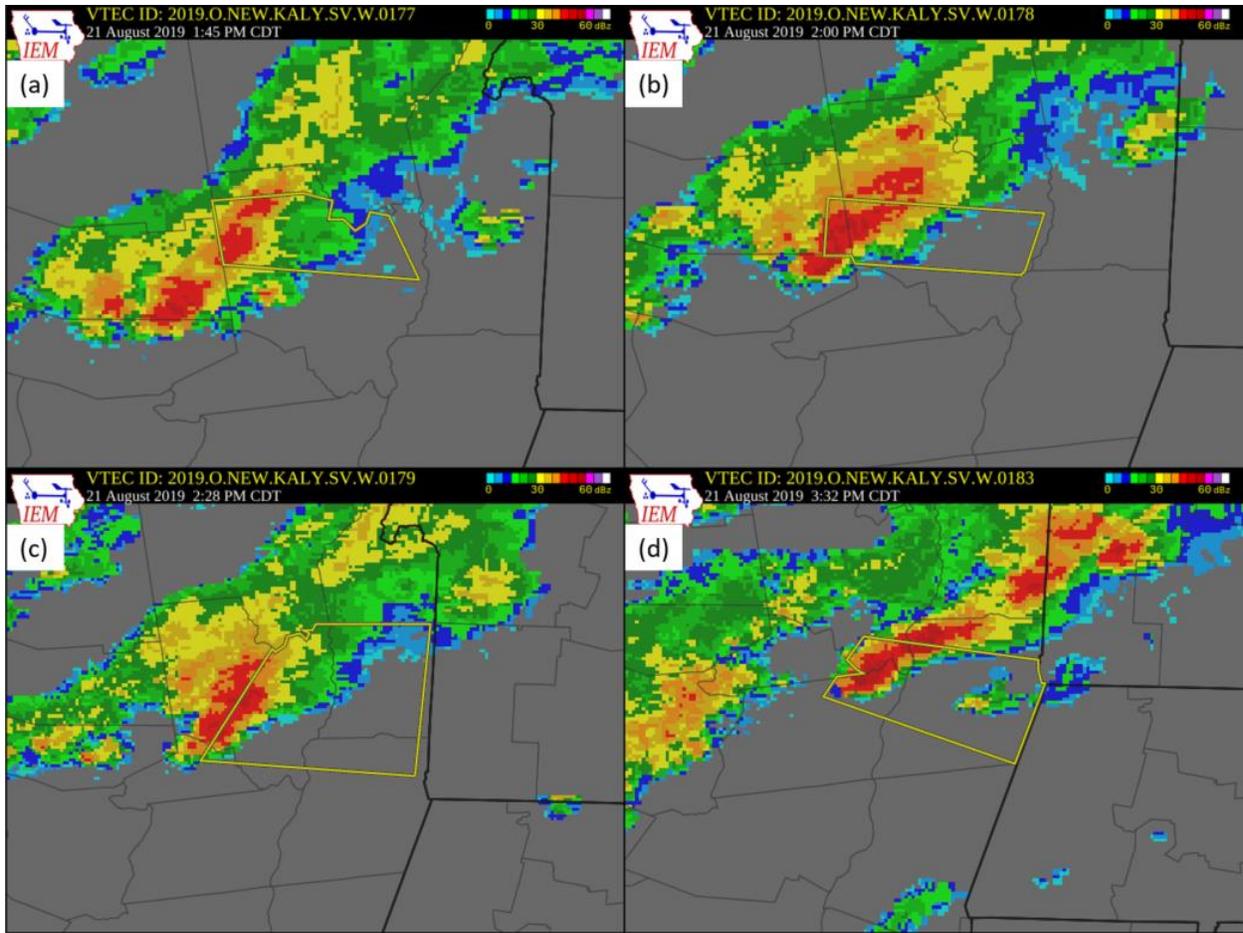


Fig. 3.2. Composite radar imagery depicting the back-building MCS at (a) 1845 UTC, (b) 1900 UTC, (c) 1928 UTC and (d) 2032 UTC on 21 August 2019. Yellow lines denote severe thunderstorm warning areas issued by the Albany WFO at each time.

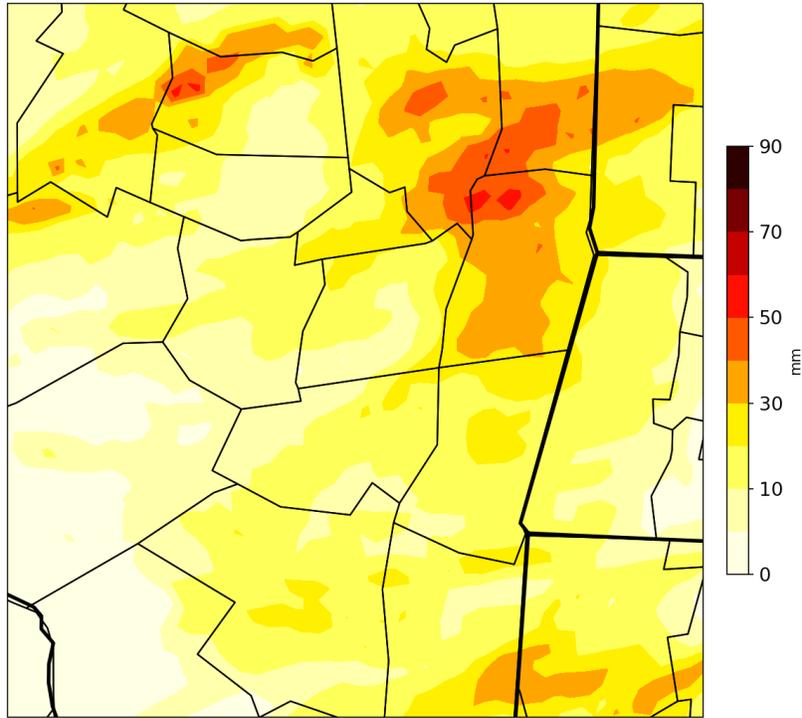


Fig. 3.3. Storm total quantitative precipitation estimate map valid at 2300 UTC 21 August 2019. Data is taken from NCEP Stage IV QPE with a horizontal resolution of 4 km. Thin black lines denote county borders and thick black lines denote state borders.

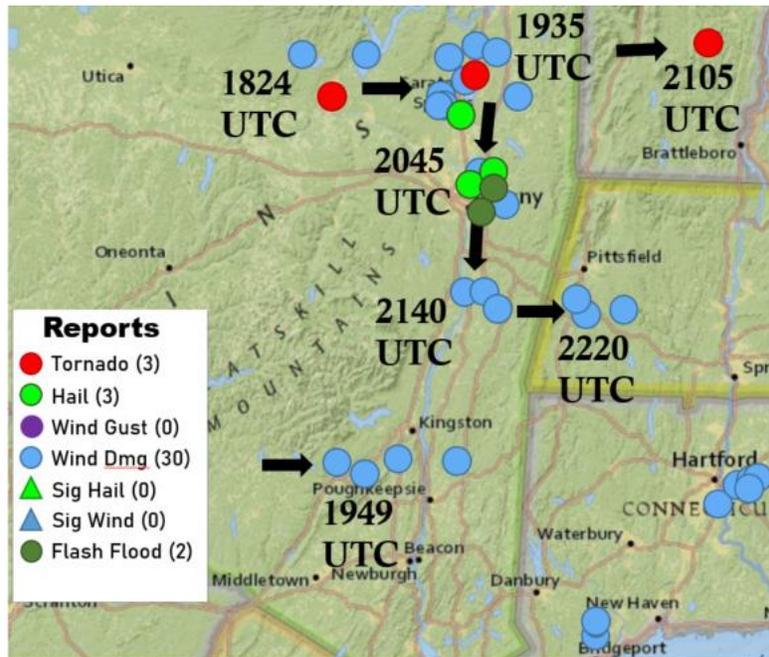


Fig. 3.4. Map of severe weather and flash flood reports from 21 August 2019. Text showing the approximate time of each cluster of reports and arrows denote the chronological evolution of the responsible convection. Adapted from NOAA Interactive Local Storm Reports.

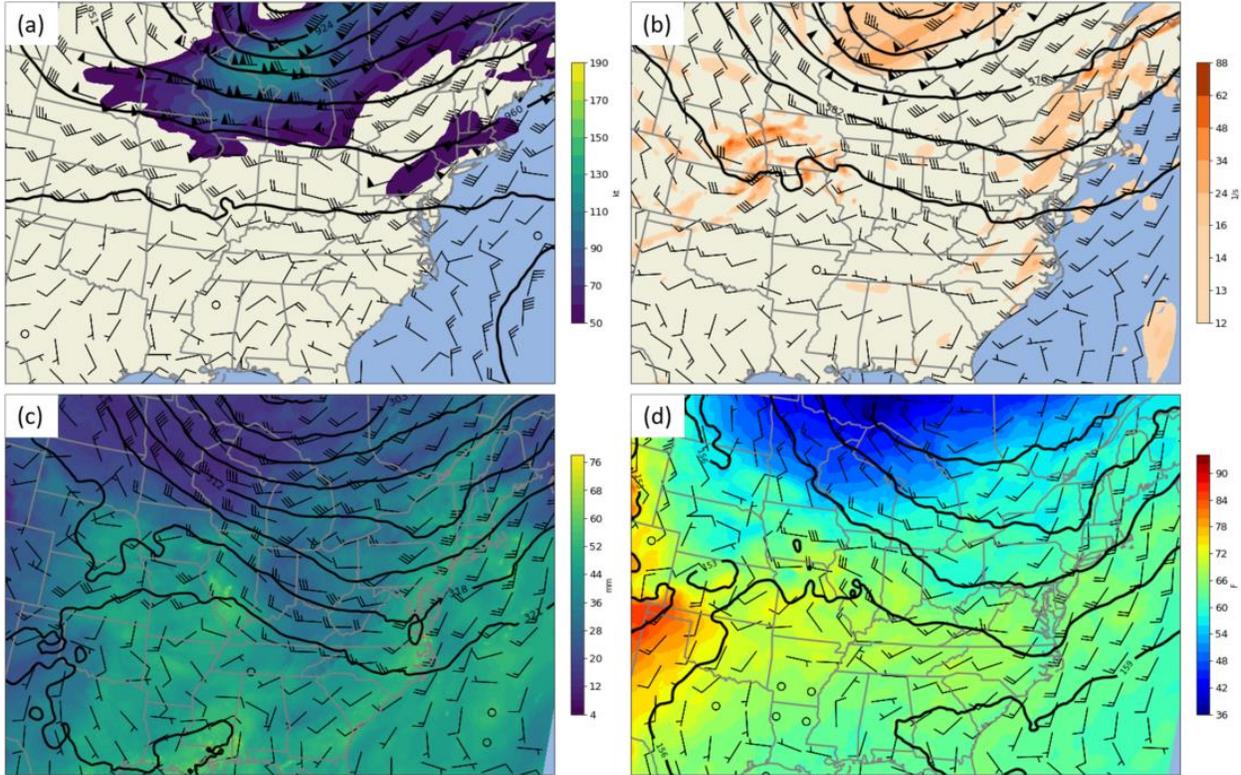


Fig. 3.5. Maps of (a) 300-hPa geopotential height (dam), isotachs (kt) and wind (kt); (b) 500-hPa geopotential height (dam), relative vorticity ($10^{-5} s^{-1}$) and wind (kt); (c) 700-hPa geopotential heights (dam), precipitable water (mm) and wind (kt); (d) 850-hPa geopotential heights (dam), temperature (F) and wind (kt). All maps valid at 1800 UTC 21 August 2019 and data taken from HRRR 0-hr analysis.

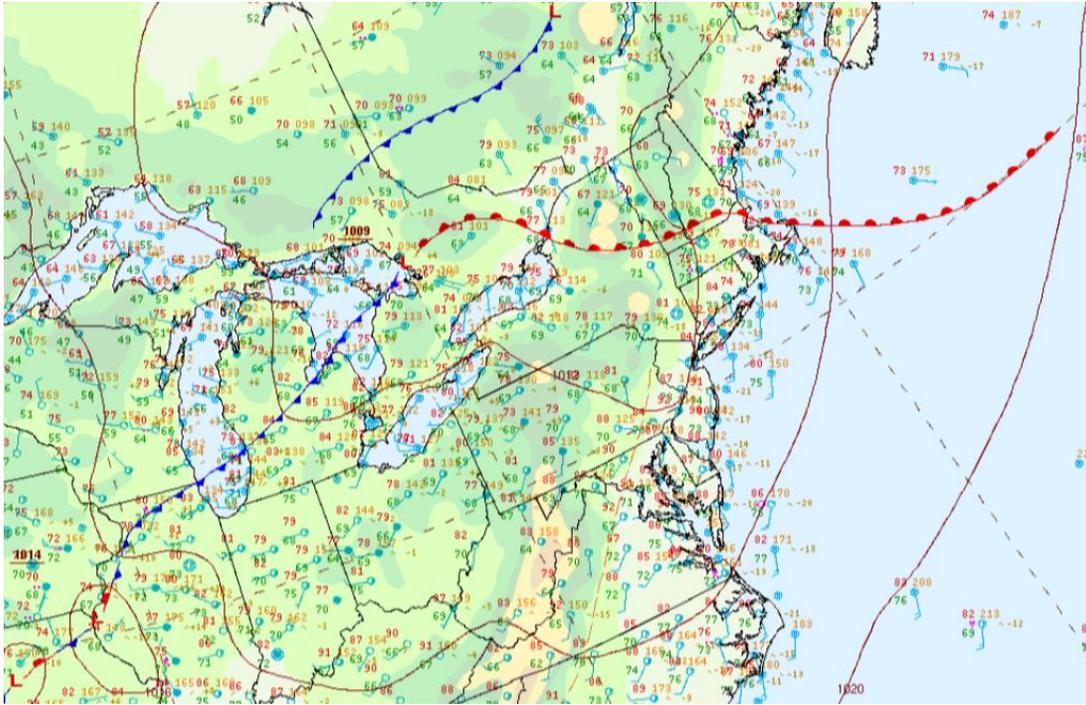


Fig. 3.6. Weather Prediction Center (WPC) surface analysis valid at 1800 UTC 21 August 2019.

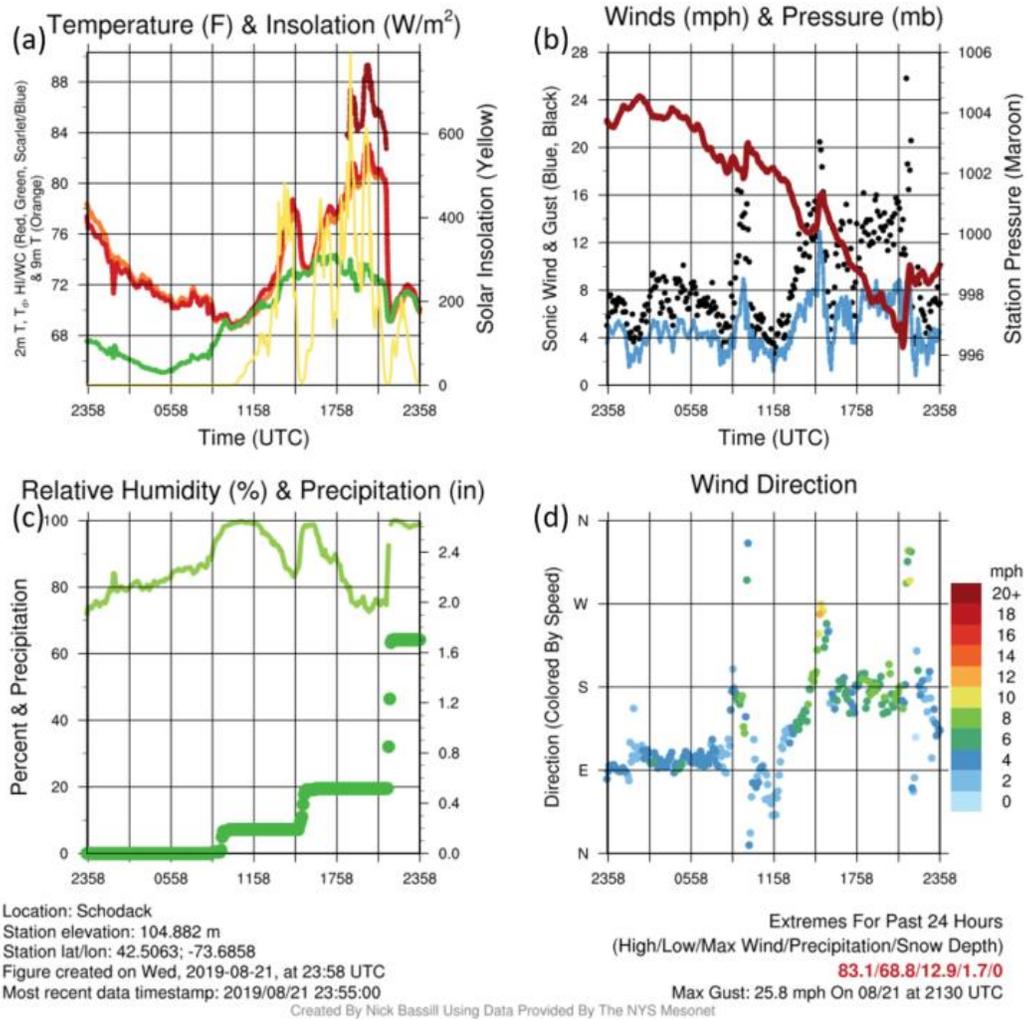


Fig. 3.7. Meteograms showing (a) temperature (F) and solar insolation (W/m^2), (b) wind (mph) and pressure (hPa), (c) relative humidity (%) and observed precipitation (in), (d) wind direction and magnitude (mph). Data taken from New York State Mesonet (NYSM) Station in Schodack, NY.

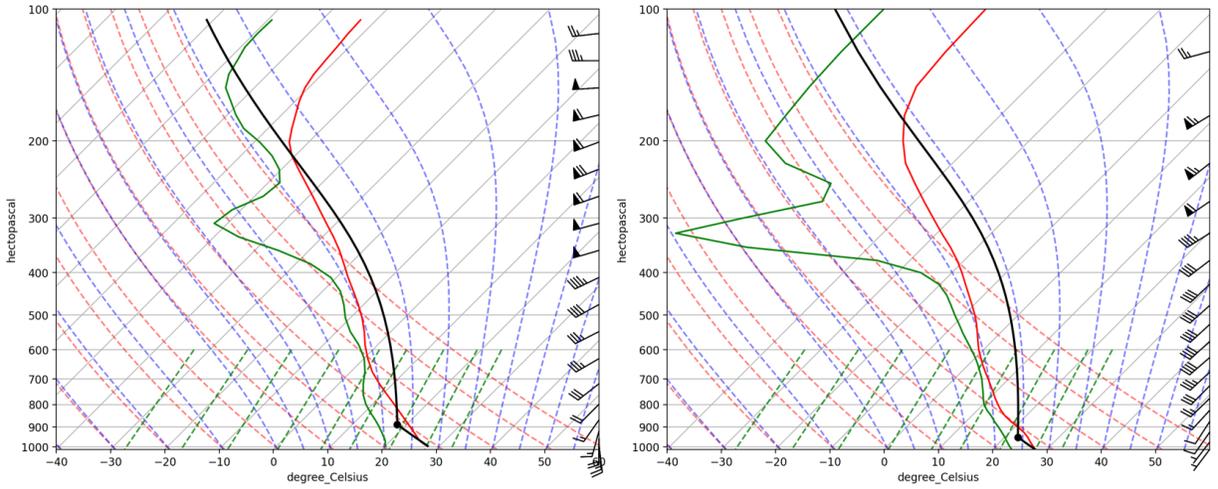


Fig. 3.8. Area-averaged Hudson Valley model soundings valid at 1800 UTC 21 August 2019 from (left) WRF simulation and (right) HRRR 0-hr analysis.

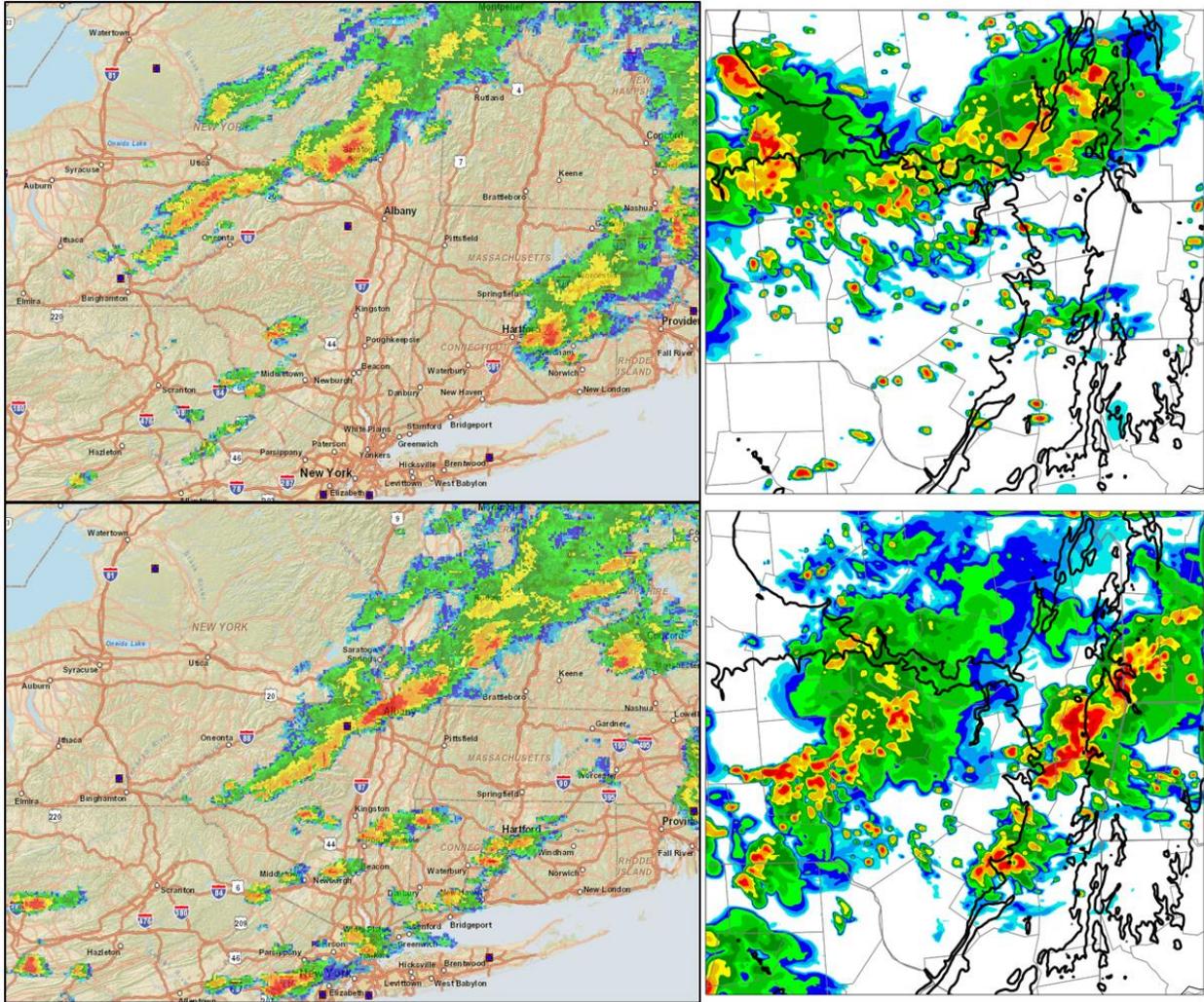


Fig. 3.9. Left panels showing composite radar imagery from 1855 UTC (top) and 2100 UTC (bottom) 21 August 2019. Right panels showing simulated reflectivity from WRF simulation of 21 August 2019, valid at 1855 UTC (top) and 2100 UTC (bottom). The 125-m terrain height contour is included to roughly outline the Mohawk and Hudson Valleys.

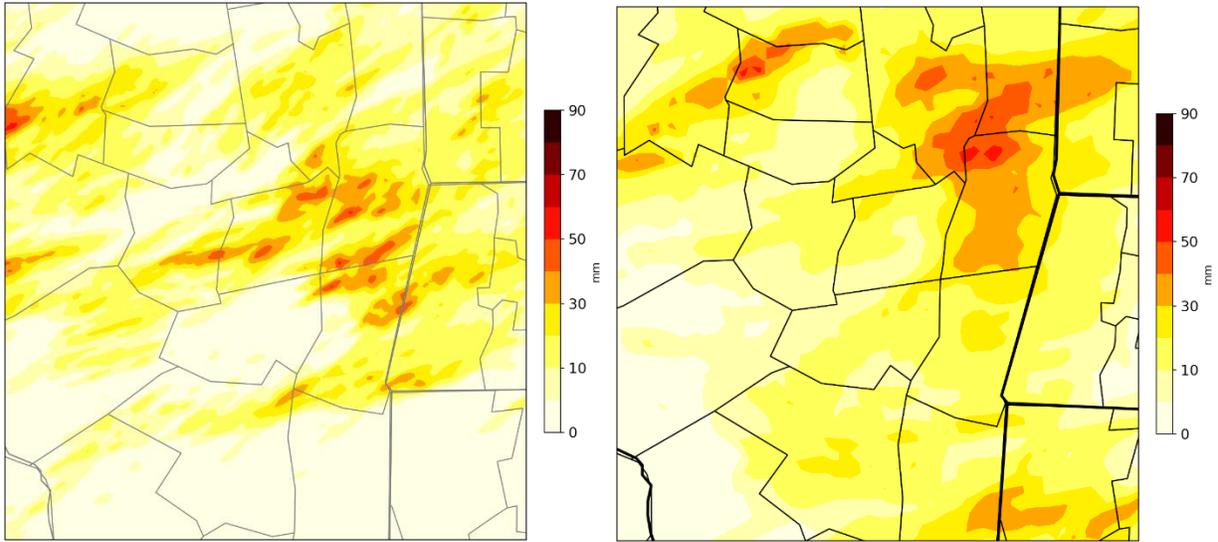


Fig. 3.10. Left panel showing simulated storm total precipitation from WRF simulation of 21 August 2019, valid at 2300 UTC. Right panel showing observed storm total precipitation from 21 August 2019, valid at 2300 UTC, taken from NCEP Stage IV QPE.

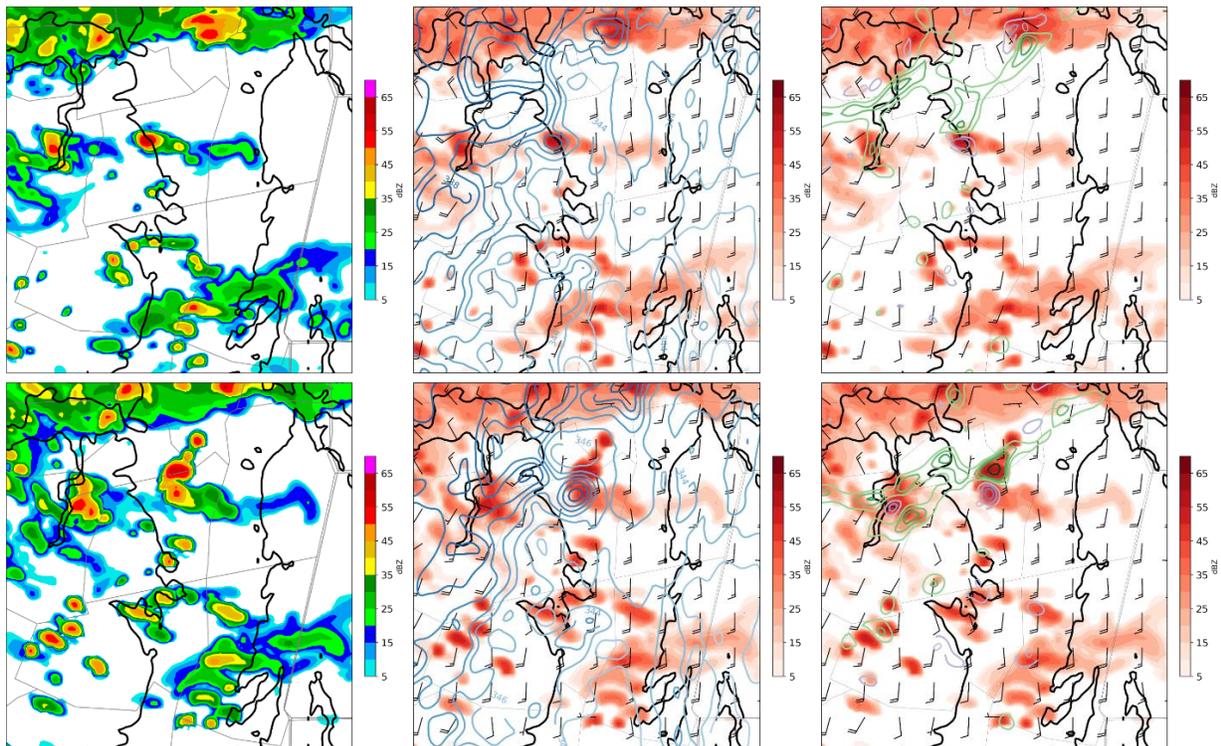


Fig. 3.11. Maps of (left) simulated reflectivity, (center) filled reflectivity with contoured θ_e every 1 K, and (right) filled reflectivity with upward (green) and downward (purple) vertical motion contoured at 0.5, 1, 1.5 and 3 m/s. The 125-m terrain height contour is included to roughly outline the central and upper Hudson Valley. All maps valid at (top) 1855 UTC and (bottom) 1910 UTC.

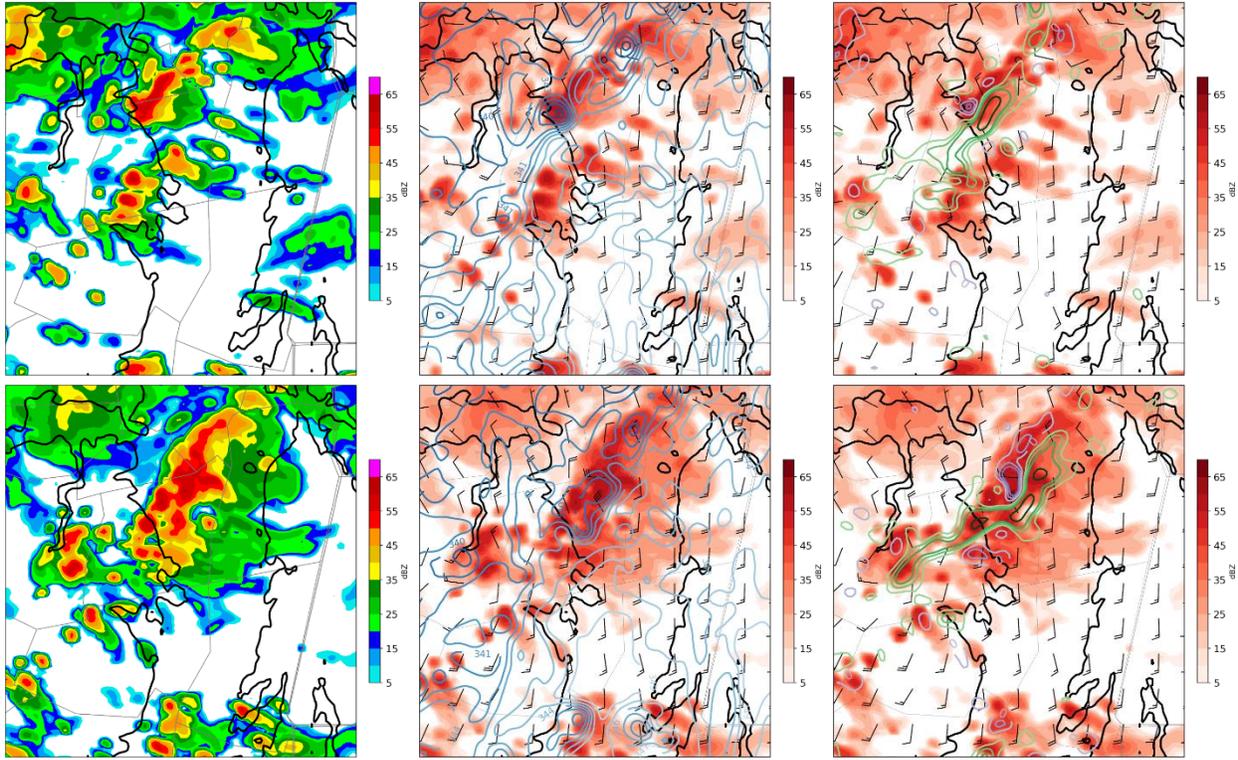


Fig. 3.12. As in Figure 3.11, but maps valid at (top) 1940 UTC and (bottom) 2000 UTC.

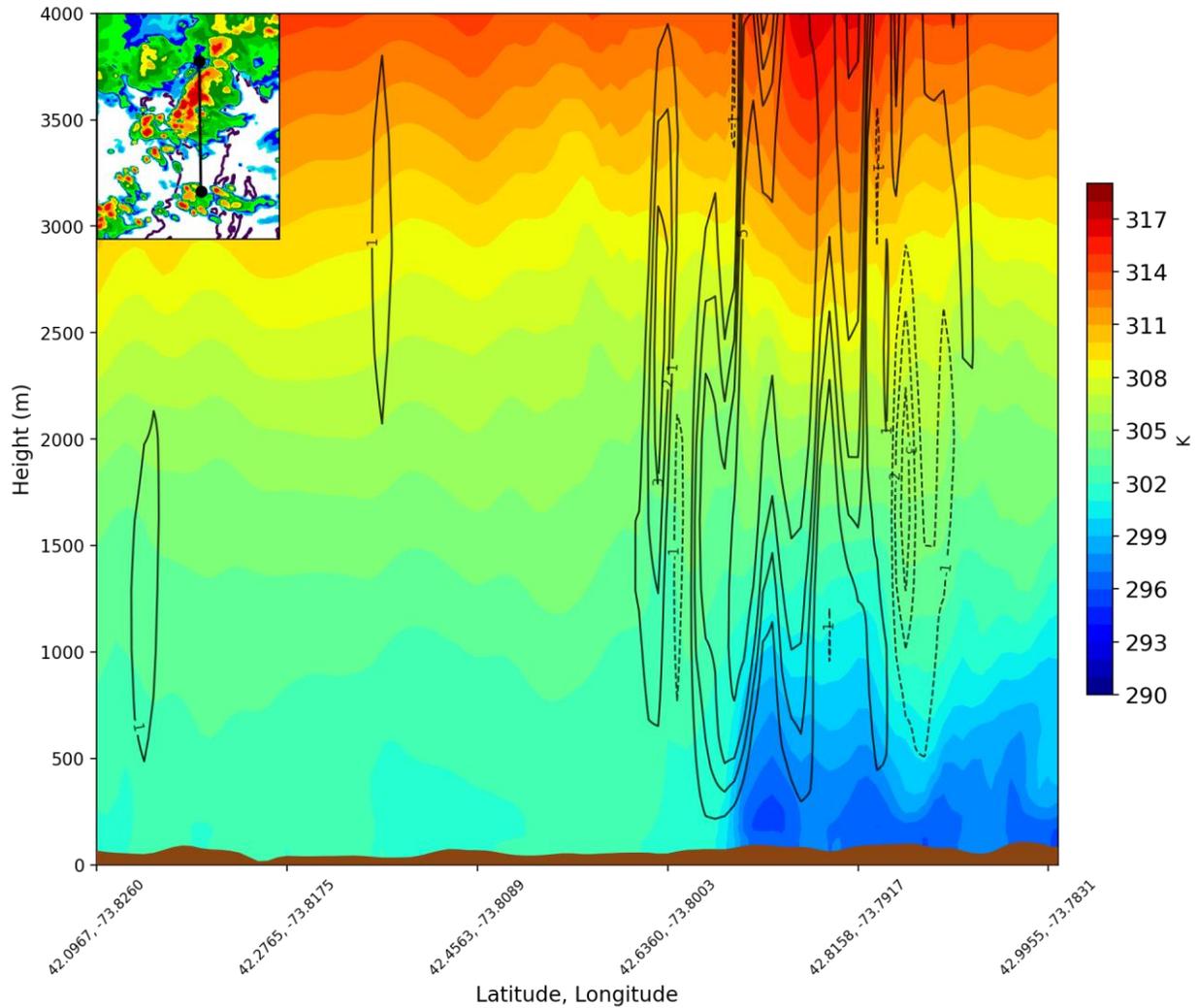


Fig. 3.13. Cross section from WRF simulation across the gust front of the back-building MCS showing filled potential temperature and vertical motion contoured at 1, 3, 5, 8 and 12 m/s. Cross section valid at 2000 UTC. Inset shows simulated reflectivity, 125-m terrain height contour and location of the cross section (black line).

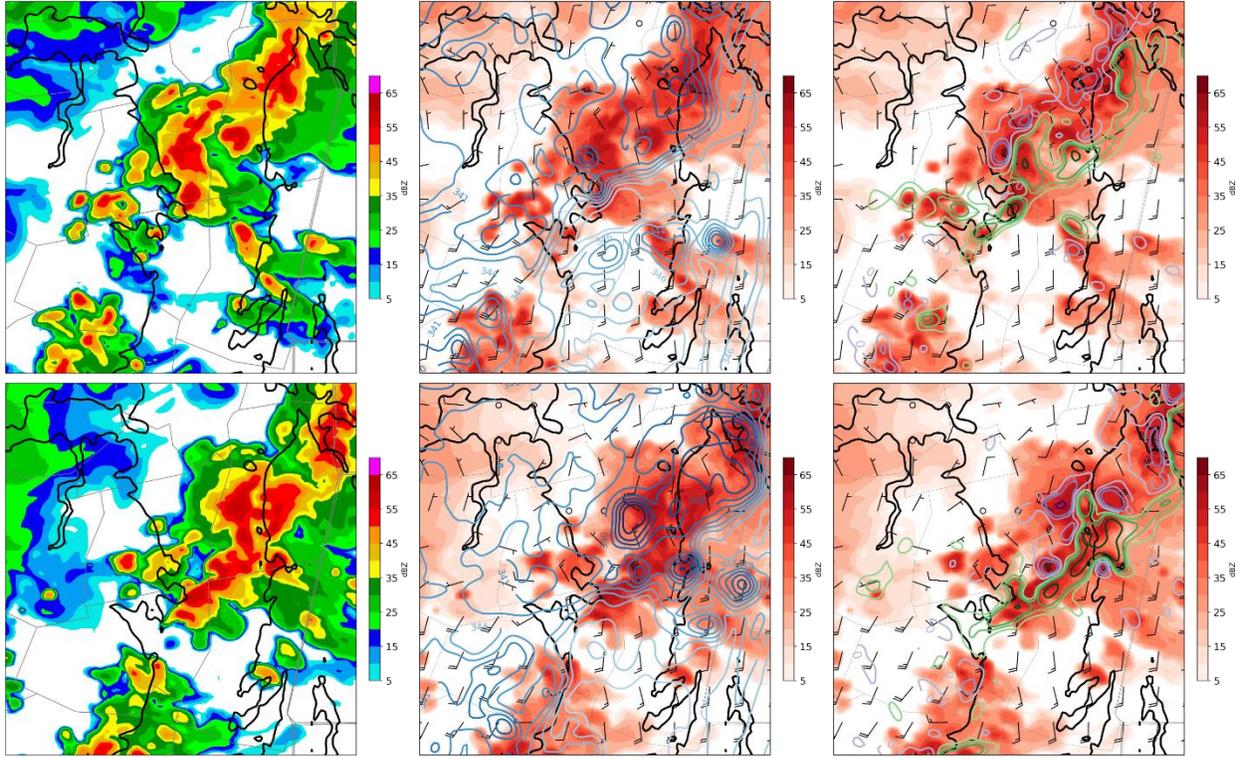


Fig. 3.14. As in Figure 3.11, but maps valid at (top) 1940 UTC and (bottom) 2000 UTC.

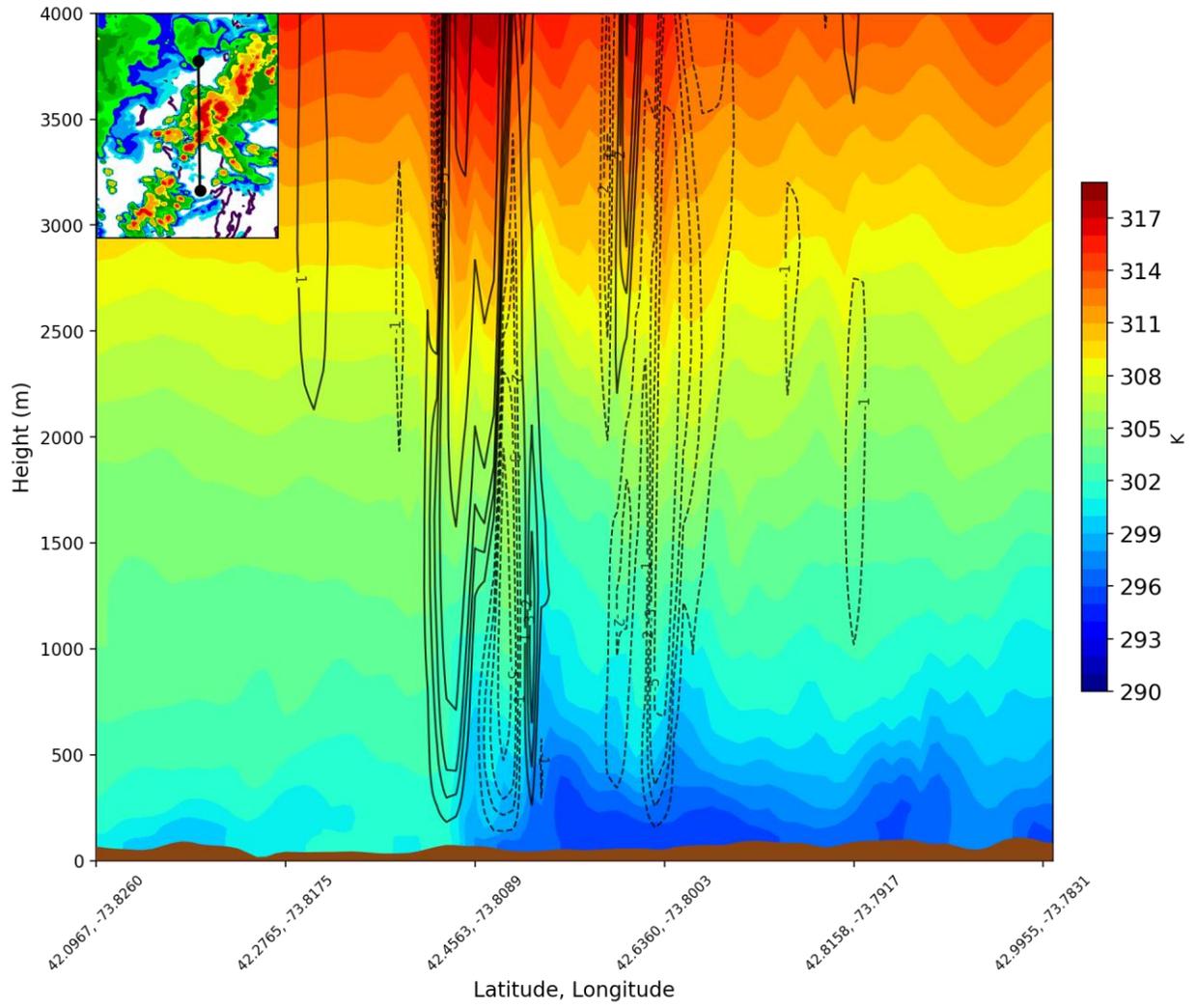


Fig. 3.15. As in Figure 3.13, but valid at 2050 UTC.

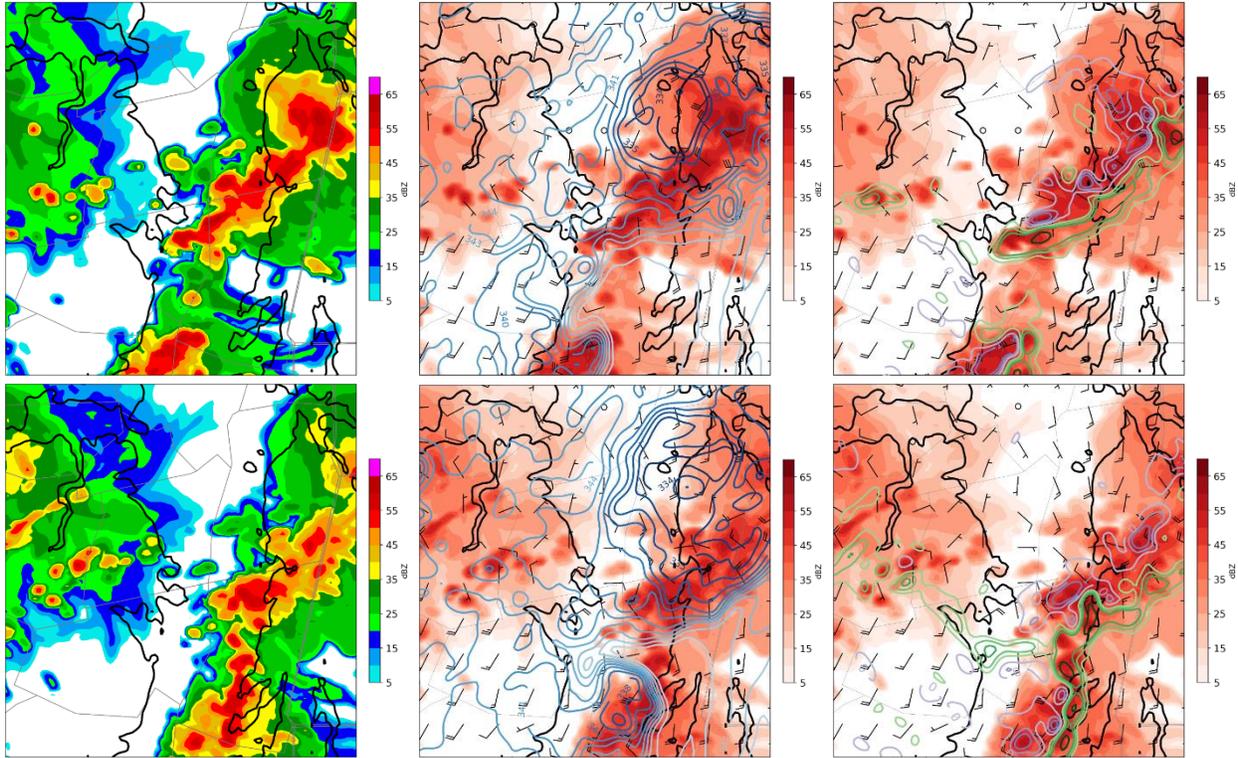


Fig. 3.16. As in Figure 3.11, but valid at (top) 2120 UTC and (bottom) 2140 UTC.

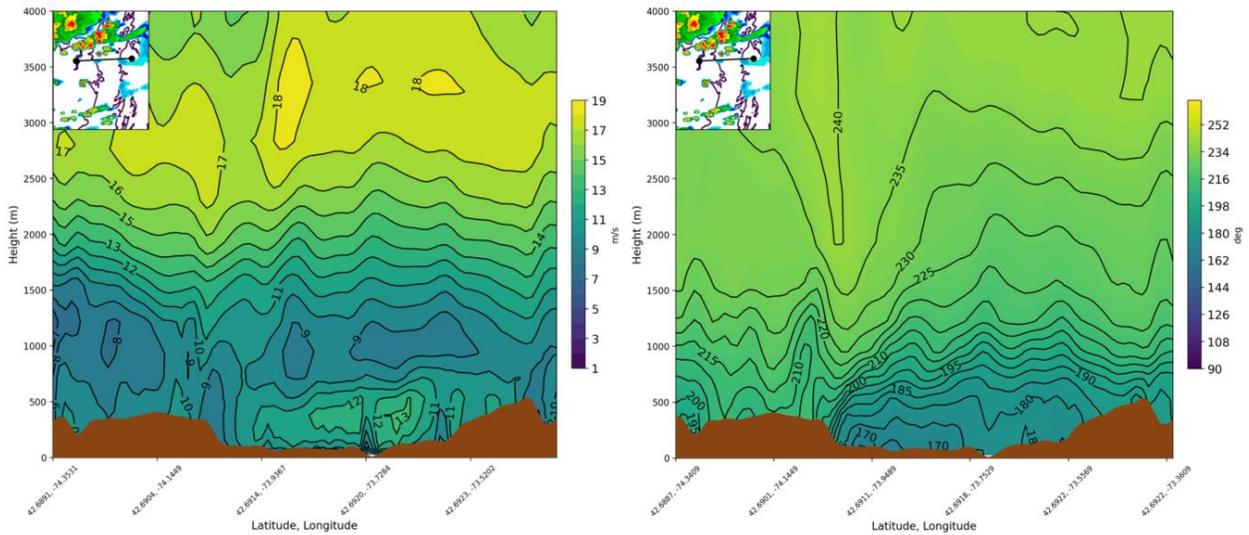


Fig. 3.17. Cross sections from WRF simulation across the Hudson Valley showing (left) wind speed (m/s) and (right) wind direction (deg), valid at 1800 UTC. Insets show simulated reflectivity, 125-m terrain height contour and location of cross section (black line).

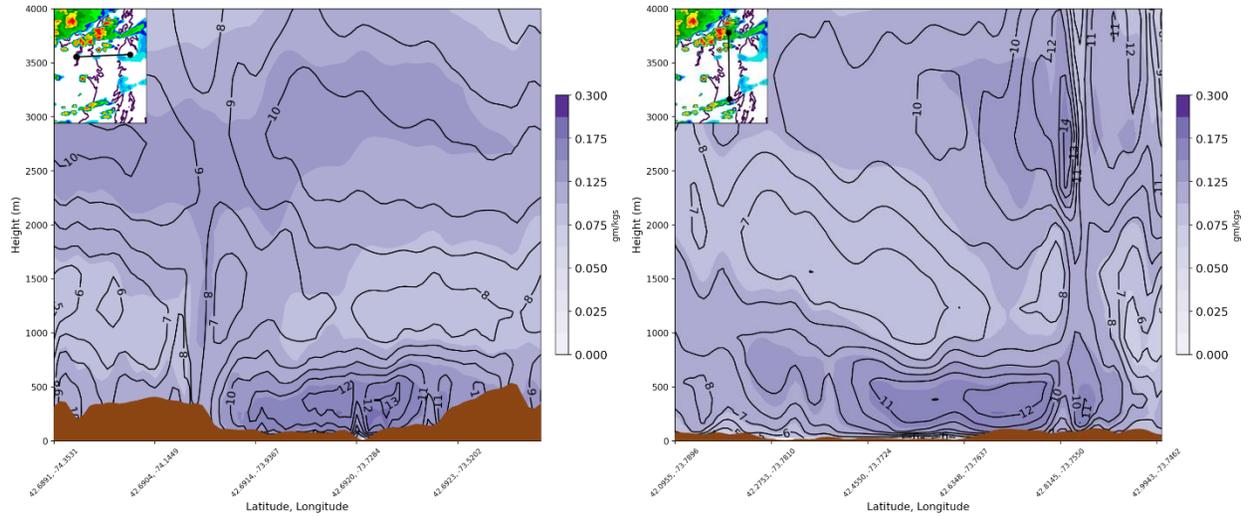


Fig. 3.18. Cross section from WRF simulation across the Hudson Valley (left) and along the Hudson Valley (right) showing horizontal vapor flux (filled) and wind speed (contoured), valid at 1800 UTC. Inset shows simulated reflectivity, 125-m terrain height contour and location of cross section (black line).

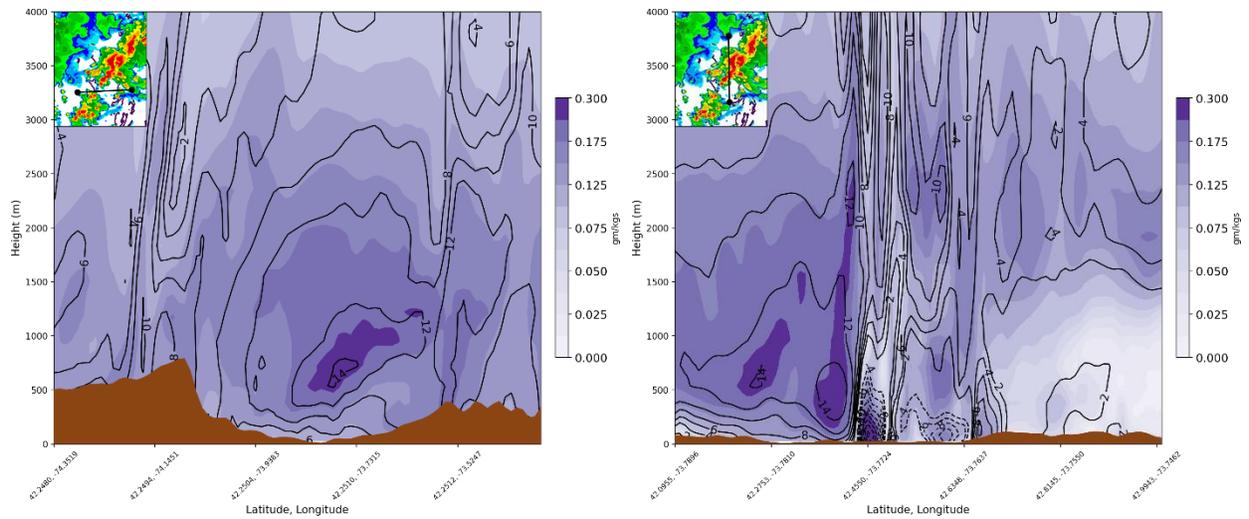


Fig. 3.19. As in Figure 3.18, but valid at 2055 UTC, the time at which the back-building MCS was most intense.

4. Prediction

4.1 Decision Tree Results

The results of the decision tree analysis can be best described by discussing which features were most commonly chosen for the 5000 random sets of cases. The most commonly selected features were a function of both the set of variables used in the decision tree classifier and the weight given to each class of outcomes. The thresholds that were used to distinguish between likely and unlikely back-building environments were also a function of the weight. When higher weights were given to the back-building cases, the thresholds tended to be less restrictive, resulting in higher impurities but fewer missed events (discussed in section 6.2).

The decision tree classifier was initially run with the short list of variables, which included surface wind variables; thermodynamic quantities at the surface, 850 hPa, and 500 hPa; as well as common severe weather indices (CAPE and shear) (Ch. 2, Table 1). Figure 4.1 is a histogram showing how many times each feature was chosen by one of the 5000 decision trees. The two most commonly chosen features were the 2-m DD and the SBCAPE. These two features were closely followed by the 500-hPa T_d and 10-m wind magnitude. When higher weights were given to the back-building cases, there was little change in the features chosen, with 2-m DD becoming slightly more common and SBCAPE slightly less common.

The threshold that was used for 2-m DD stayed constant at 17.75°F regardless of the weight assigned to back-building cases, but the SBCAPE threshold was not constant. Figure 4.2 shows the back-building and non-back-building cases plotted as a function of SBCAPE and 2-m DD, where the shaded area represents the combination of variables most conducive to back building. When the back-building cases were assigned a weight of 1 or 2, the threshold for

SBCAPE was more restrictive, with a value above 2375 J/kg indicating an increased likelihood of back building. When the weight of back-building cases was increased to 3 or 4, the threshold decreased to a value above 1275 J/kg. The less restrictive threshold allowed more back-building cases to fall inside the shaded area, but caused the impurity to increase from 0.38 with the lower weights to 0.57 with the higher weights.

The results were not as clear when the decision tree classifier was run on the expanded list of variables, which included more wind variables from the mid troposphere, surface thermodynamic variables, and severe weather indices (Ch. 2, Table 1). The most commonly selected feature by far was the LI, which was chosen regardless of the weight applied to back-building cases (Figure 4.3). The threshold chosen at all weights for the LI was values below -5.25°F (Figure 4.4). When the back-building cases were given a weight of 1, the 850-hPa T_d was the second most common feature, while a weight of 2 or 3 shifted it to the 2-m T. Applying a weight of 4 to the back-building cases yielded 4 different features that were chosen a similar number of times: the 2-m T, DD, and RH, followed by the 10-m wind magnitude. The thresholds for the 850-hPa T_d and 2-m T were values above 56.2°F and below 88.5°F respectively, while the threshold for the 10-m wind magnitude was values below 4.25 m/s (Figure 4.4).

While there was little direct overlap between the features chosen from the short and expanded lists, the physical meaning behind the results is largely the same. Both analyses had a measure of deep instability for a surface-based parcel, whether that be in the form of the SBCAPE or LI. The second feature chosen for the decision trees with a back-building weight of 3 was the 2-m DD, which had an upper limit for their threshold. A small DD implies that a parcel lifted from the surface will quickly become saturated and start cooling at the moist adiabatic lapse rate as it rises, maximizing the CAPE. If the DD is higher than the threshold, it is also more

likely that the cold pool will become too strong and outrun the convection, decreasing the likelihood of back building.

The physical significance of the other chosen features, a low 850-hPa T_d , 2-m T and 10-m wind magnitude respectively, are less obvious. The presence of dry air at 850-hPa, in the form of a low T_d , could conceivably lead to a stronger cold pool through the entrainment of dry air near the base of the cloud to encourage evaporative cooling. However, the dry air being entrained into the cloud near the base would decrease the buoyancy, making the growth of an updraft into a deep convective cell less likely. The 2-m T having a low value may make the 2-m DD lower, but would also decrease the temperature along the surface-based parcel's path and create less SBCAPE. A low 10-m wind magnitude would likely cause there to be less low-level convergence along the gust front, leading to a weaker updraft.

In order to further test the robustness of the results, we applied PCA on the variables in the dataset. This PCA resulted in a number of new principal component variables (or factors) that were linear combinations of the original area average variables. When the decision tree classifier was instead run with the PCA factors as inputs, there were two factors that were used more commonly than any others (Figure 4.5). The first factor (F4) was most closely tied to the 2-m RH, DD, and T, which will be referred to as the “low-level thermodynamics” factor. The second factor (F10) was a combination of the TT, a measure of the mid-level temperature and moisture lapse rates, and 500-hPa T, which will be referred to as the “mid-level instability” factor. The PCA factors represent another way to measure the same physical quantities that were chosen by the other decision trees: ample low-level moisture and deep instability.

4.2 Synoptic Environment

The back-building cases that fulfilled the thresholds identified in the decision tree analysis can be considered to be correctly-identified cases. A further analysis of the cases that were correctly identified by the decision trees using the expanded list of features and a back-building weight of 3 will be given in this section. The back-building weight of 3 was chosen as it represents a balance between the lowest impurity and lowest number of missed events. The decision tree controlled for a LI below -5.25°F and 2-m T below 88.5°F ; and identified 11 back-building cases.

The synoptic environment associated with back-building convection in the Hudson Valley is characterized by weak synoptic forcing in the Northeast. The 300-hPa isotach composite shows a pronounced ridge-trough pattern over the U.S., with the trough axis located over the Great Lakes region (Figure 4.6). The height gradient is small over the Northeast, producing southwesterly winds of only 30 kt downstream of the trough. A jet streak is present on the downstream side of the trough, but is too far north to produce strong forcing for ascent in eastern New York. The 500-hPa vorticity composite shows there is no strong vorticity advection, on average, during the back-building cases. Figure 4.7 is a spaghetti plot with the 582-dam contour on the 500-hPa level for each correctly-identified back-building case and the mean 582-dam contour in bold. Three of the cases have a trough that is considerably deeper than the other cases, causing the mean flow to be more amplified than it would be otherwise.

Figure 4.8 shows 700-hPa geopotential height and winds over the precipitable water field for the eleven correctly-identified back-building cases. There is confluent flow over the East Coast, where the westerly winds from the interior U.S. meet the southerly winds associated with

an anticyclone off the Southeast Coast. The southerly flow has a slightly higher value of precipitable water than the westerly flow, transporting moisture into the Northeast. The 850-hPa level composite also shows the trough collocated with a thermal trough. There is a temperature gradient between the northern Great Lakes region and upstate New York, representing the average location of a cold front. Individual cases had a stronger temperature gradient and wind shift, but their cold front position varied slightly.

Some of the localized features discussed in section 3.1.3 appeared in the composite of MSLP, 10-m wind and 2-m T_d . A weak temperature gradient and slight wind shift indicate a cold front stretching along the eastern Great Lakes. Western NY has south-southwesterly winds between 5 and 10 kt, while the Hudson Valley and New England have southerly onshore flow. There is a slight decrease in the MSLP within the Hudson Valley and to the east of the Taconic Mountains, possibly caused by flow descending off of the mountains creating a lee trough.

The 2-m T_d has a clear dependency on the terrain features in Upstate NY, with higher values over the Atlantic Ocean and Great Lakes and lower values over the Catskill and Adirondack Mountains. A local maxima in T_d exists within the central Hudson Valley, where values reach the mid-70s. It is unclear if the higher T_d in the Hudson Valley are a direct result of terrain channeling. Another possible explanation for this feature could be higher surface temperatures within valleys compared to the higher terrain, allowing a higher moisture content for a given relative humidity.

Figure 4.10 shows the area-average sounding for the correctly-identified back-building cases from the decision tree analysis, meaning cases with LI below -5.25°F and 2-m temperature below 88.5°F . The sounding could be described a high-CAPE, moderate-shear sounding for the Northeast. There is a layer of steep lapse rates between 850 and 950 hPa, which is largely

responsible for the deep instability for a surface-based parcel. The temperature profile above 850 hPa has a conditionally unstable lapse rate which decreases up to the tropopause. The surface-based parcel is more unstable than a parcel lifted from any other level, which favors convective cells that are initiated from convergence along the gust front.

The T_d profile has some interesting features at both the surface and mid troposphere. At the surface T_d is 23°C or 73.5°F. This moist surface environment allows the surface-based parcel to reach saturation at an LCL pressure of 939.6 hPa. The mid-levels are drier on average due to a low T_d on average. The drier mid-levels may be indicative of less cloud coverage allowing more shortwave radiation to reach the surface, though the HRRR estimated total cloud cover was not chosen in any decision trees.

The wind profile is consistent with the weak height gradient shown on the planar-view composites. The mean tropospheric wind for the correctly-identified cases is 20 kt, resulting in a slow individual cell motion. Slow-moving cells will not move out of the Hudson Valley as quickly, giving them more time to form a strong cold pool through evaporative cooling. There is very slight veering in the profile, indicative of weak warm-air advection, which can contribute to upward vertical motion and height falls.

4.3 False Positives

There were a number of null cases that became apparent during the decision tree analysis. These cases, which will be referred to as false-positive cases, fulfilled the threshold values of surface thermodynamic and instability variables established by the decision trees but did not result in a back-building MCS in the Hudson Valley. There were 10 false-positive cases

identified by the decision trees. A similar analysis of the synoptic-scale features common in the false-positive cases provides information that can be useful in forecasting back-building convection.

The composite of 300-hPa isotachs shows a similar height field to the back-building composite (Figure 4.11). There is a weak ridge over the West Coast and a trough over the Great Lakes region, but the trough is displaced further to the west. The false-positive composite has a stronger jet streak over Michigan with the equatorward exit region of the jet streak over Eastern New York. However, the acceleration out of the jet streak is weak so there is still no strong forcing for vertical motion associated with it. The wind at 300-hPa and 500-hPa is southwesterly, but stronger than the back-building composites as a result of the stronger height gradient.

Figure 4.12 shows the false-positive composites for the 700-hPa height and precipitable water; and the 850-hPa height and temperature. These composites show a well-defined extratropical cyclone in southeastern Canada that was not present in the back-building composites. There is a cold front from the low-pressure center to the southwest, ending around Illinois. The precipitable water field and 700-hPa winds suggest that the main axis of moisture transport is downstream of the cold front, as opposed to the coastal moisture transport found in the back-building cases. The height gradient at both 700 and 850-hPa is much stronger, producing stronger mid and lower tropospheric winds.

The spatial patterns in the 2-m T_d are very similar to those found in the back-building composite. There are lower T_d in the Catskill and Adirondack Mountains, however the maxima within the valleys is more widespread, with the Mohawk Valley appearing as moist as the Hudson Valley (Figure 4.13). The height gradient in the false positive cases is much stronger than for back-building cases, indicating a deeper surface cyclone in Quebec. There is still no single location that is favored for the cold front in western New York, but there is a small T_d

gradient. The lee trough found in the Hudson Valley for the back-building cases is still present for the false positive cases, but it is not quite as deep.

The composite sounding for false positive cases reflects many of the differences seen in the planar-view composites (Figure 4.14). The wind profile has mostly westerly winds, however there is more veering in the lower levels, indicating stronger warm air advection. The mean tropospheric wind speed is stronger which may cause the individual cell motion to be faster than for back-building cases (discussed in section 4.1.4). The thermal profile is very similar to that found in the back-building cases (Figure 4.15). The false positive cases tend to have higher T_d at all levels of the troposphere. The drier profile seen in the back-building cases may contribute to stronger cold pools through increased entrainment of dry air and evaporative cooling below the cloud base.

4.4 Failure Modes

The nature of back-building convection in the Hudson Valley is such that there are two necessary conditions that need to be met, a favorable environment and the presence of slow-moving discrete or multicellular convection. There are several characteristics of the pre-existing convection that will not allow the second necessary condition to be met, which will be referred to as failure modes. These failure modes can be observed on radar imagery, making them a powerful tool for forecasters in an operational setting, short-term forecasting or now-casting scenario.

The first failure mode has to do with the location where the discrete convection is initiated. If the discrete convection is initiated too far upstream, it will grow into mature cells and

dissipate before it is able to enter the Hudson Valley. An example of this failure mode is illustrated by the composite radar imagery from 19 June 2019 (Figure 4.16). The convective environment within the Hudson Valley on this day fulfilled the surface thermodynamic and instability thresholds established by the decision tree analysis. Discrete convection was first initiated over the Finger Lakes and western New York at 1925 UTC and had a due eastward cell motion. At 2320 UTC, the cells had gone through their life cycle and were dissipating over the Catskills Mountains. The dissipating cells failed to create a cold pool in the Hudson Valley to act as a lifting mechanism, so no back building occurred.

The next failure mode is the wrong intensity or spatial extent of the pre-existing convection, which could be thought of as a “Goldilocks zone”. If the pre-existing convection entering the Hudson Valley is too weak to form a sufficiently strong cold pool before moving off to the east, then the lifting mechanism may not be strong enough to initiate new convective cells. If the convection is too widespread, including discrete cells initiating in the lower Hudson Valley, then the surface-based instability may be consumed as the thermodynamic environment becomes contaminated with cold air. Furthermore, the individual cold pools forming in the lower Hudson Valley can disrupt the southerly terrain-channeled flow, which helps facilitate up-valley moisture flux and lower-level convergence along a potential gust front.

Finally, cell motion from any direction other than the southwest or due west cannot result in back building. A thought experiment can be used to logically deduce that back building will not occur with any other cell motion. Imagine a cluster of discrete cells that forms over the lower Hudson Valley and moves north. The Hudson Valley has southerly flow at the surface and a southerly mean-tropospheric wind, producing the southerly cell motion. In order for new cells to be initiated upstream of the mature cells, there would need to be low-level convergence along the

trailing side of the cold pool. The motion of the cold pool, estimated as the mean cloud-layer wind (Corfidi 2003), would be to the north, making strong convergence along the southern gust front very unlikely.

Next imagine a scenario where the mean tropospheric wind is easterly, producing cell motion from the east. If the surface winds are from the south, there is deep backing of the winds, consistent with cold air advection. Synoptic scale cold-air advection will not likely produce the necessary thermodynamic environment in the Hudson Valley and will cause subsidence, depressing the updrafts of any potential new cells. If the surface winds are from the north, the air being transported into the valley will not have a high moisture content, as the source region will be southeastern Canada.

Lastly, imagine a convective day with cell motion from the north. Southerly flow at the surface would create convergence on the downstream of any convection that forms, not on the upstream side. If a line of convection were to grow upscale due to a favorable environment in the Hudson Valley, it may grow into a linear MCS, but would not have any upstream propagation and therefore will not exhibit back building.

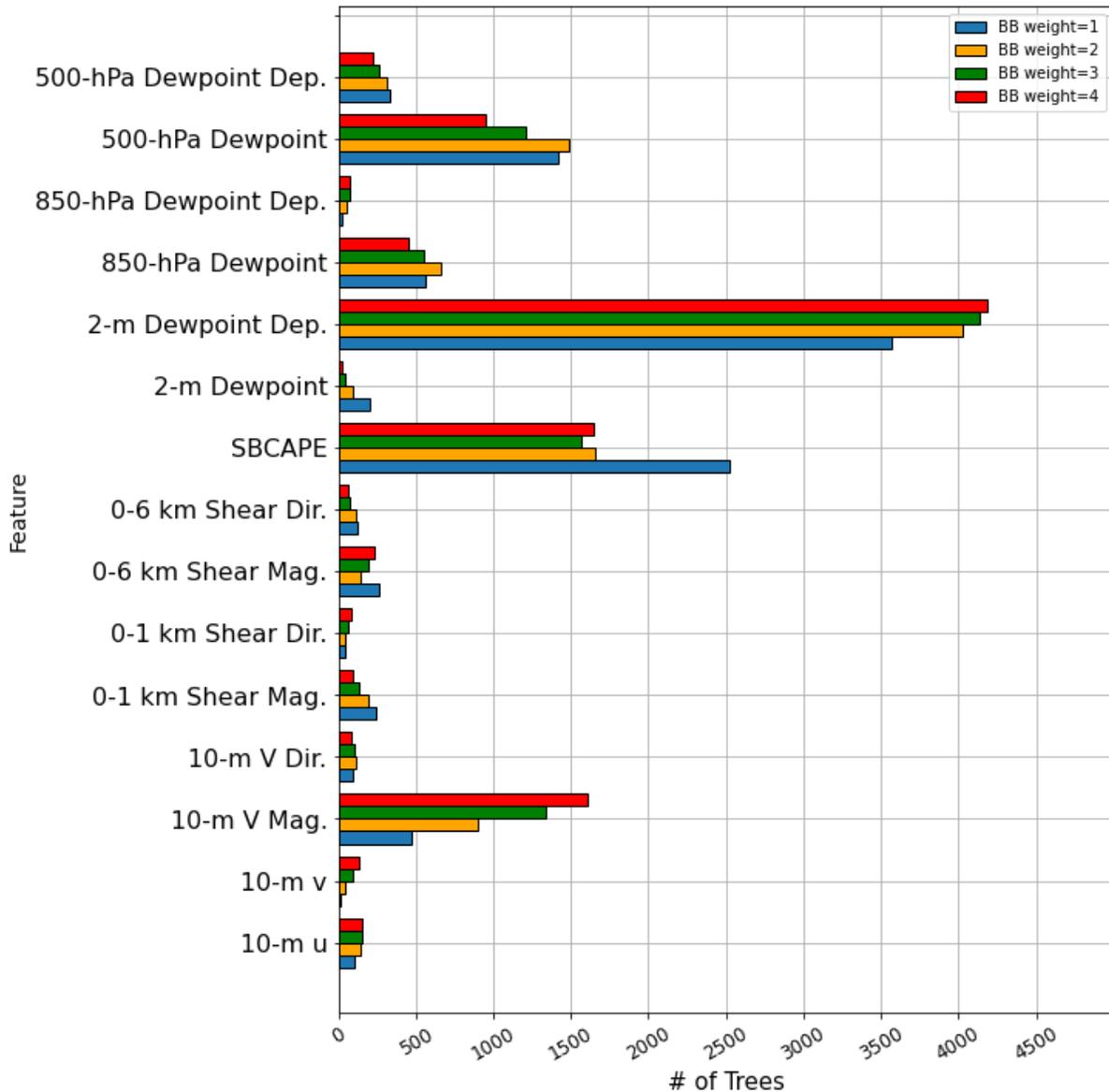


Fig. 4.1. Histogram showing the number of decision trees in which each feature was chosen from the short list of variables. Colors indicate the weight that was assigned to back-building cases in the decision tree classifier. Features represent the area average in the Hudson Valley taken from the HRRR 1800UTC 0-hr analyses.

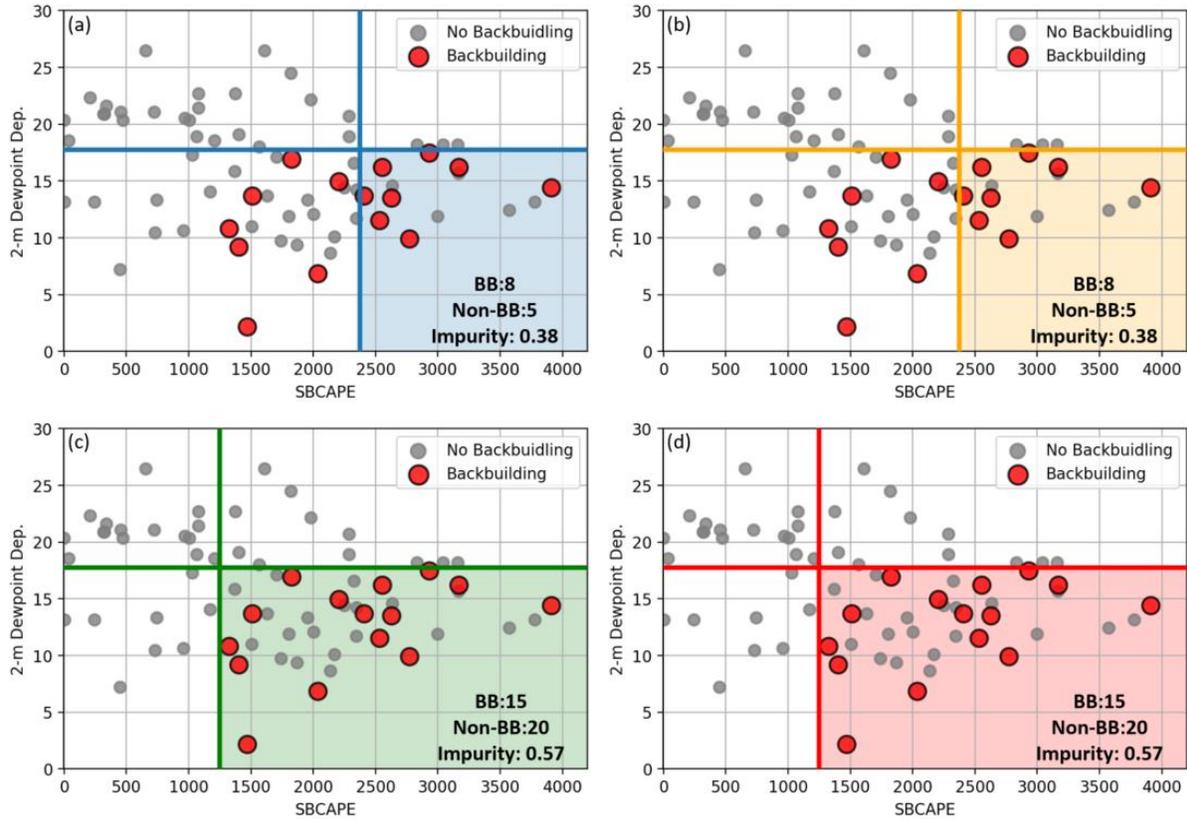


Fig. 4.2. Parameter space of back-building (red) and non-back-building (gray) cases. Parameters are the Hudson Valley area-averaged variables that were most commonly chosen from the short list by the decision trees. Back-building cases have been given a weight of (a) 1, (b) 2, (c) 3, and (d) 4 respectively. SBCAPE (J/kg) and 2-m DD (F) were chosen for all weights. Bold lines indicate the “best splits” from the decision tree, with the shaded area indicating the combination of parameters that yields the best identification of back-building versus non-back-building cases.

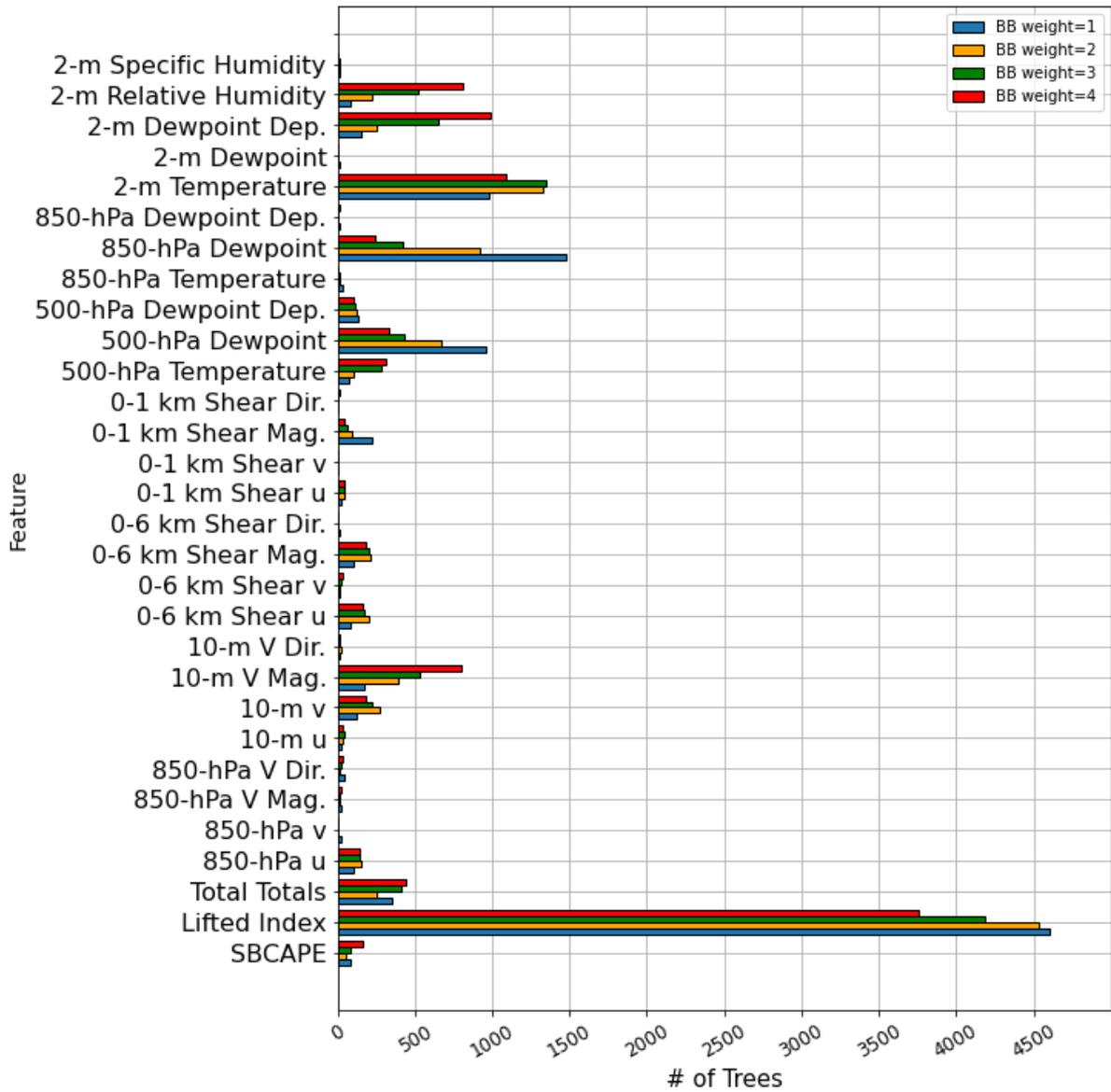


Fig. 4.3. As in Figure 4.1, but with the expanded list of variables.

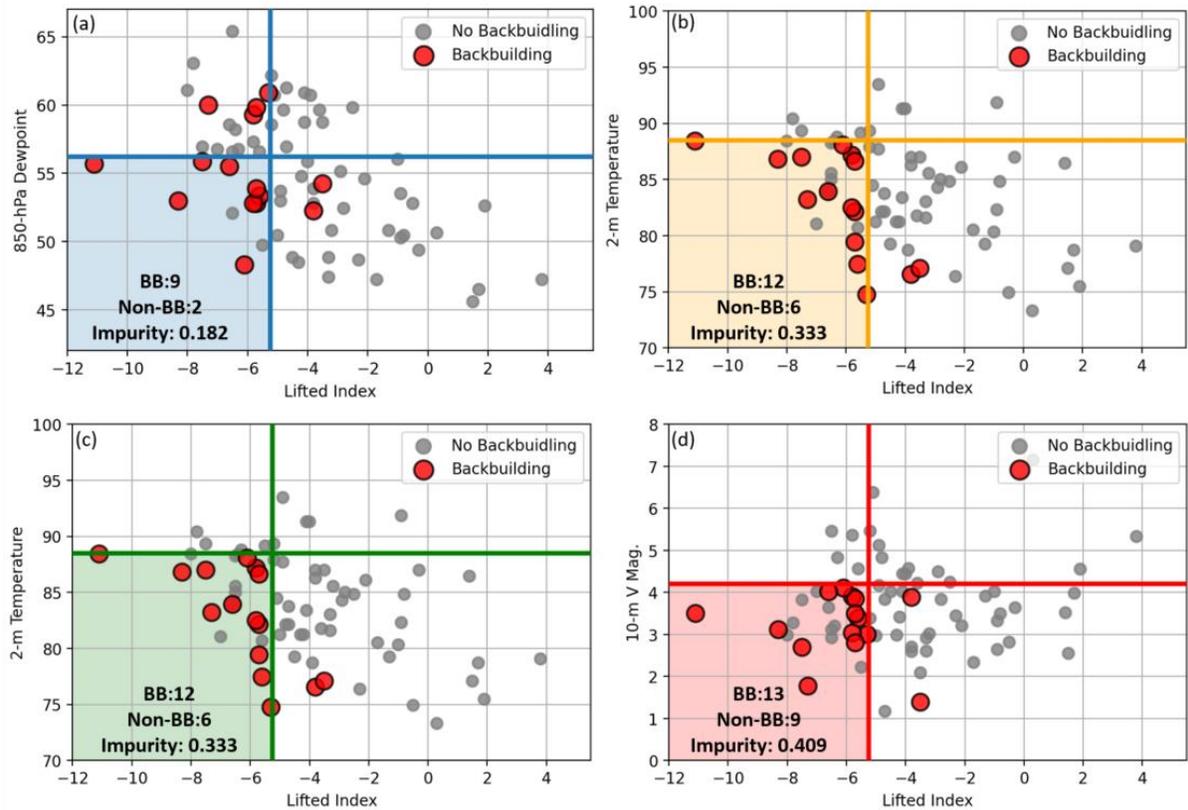


Fig. 4.4. Parameter space of back-building (red) and non-back-building (gray) cases. Parameters are the Hudson Valley area-averaged variables that were most commonly chosen from the expanded list by the decision trees. Back-building cases have been given a weight of (a) 1, (b) 2, (c) 3, and (d) 4 respectively. Lifted index (F) was chosen for all weights, with the second selected features being (a) 850-hPa dew point (F), (b) 2-m temperature (F), (c) 2-m temperature (F), and (d) 10-m wind magnitude (m/s). Bold lines indicate the “best splits” from the decision tree, with the shaded area indicating the combination of parameters that yields the best identification of back-building versus non-back-building cases.

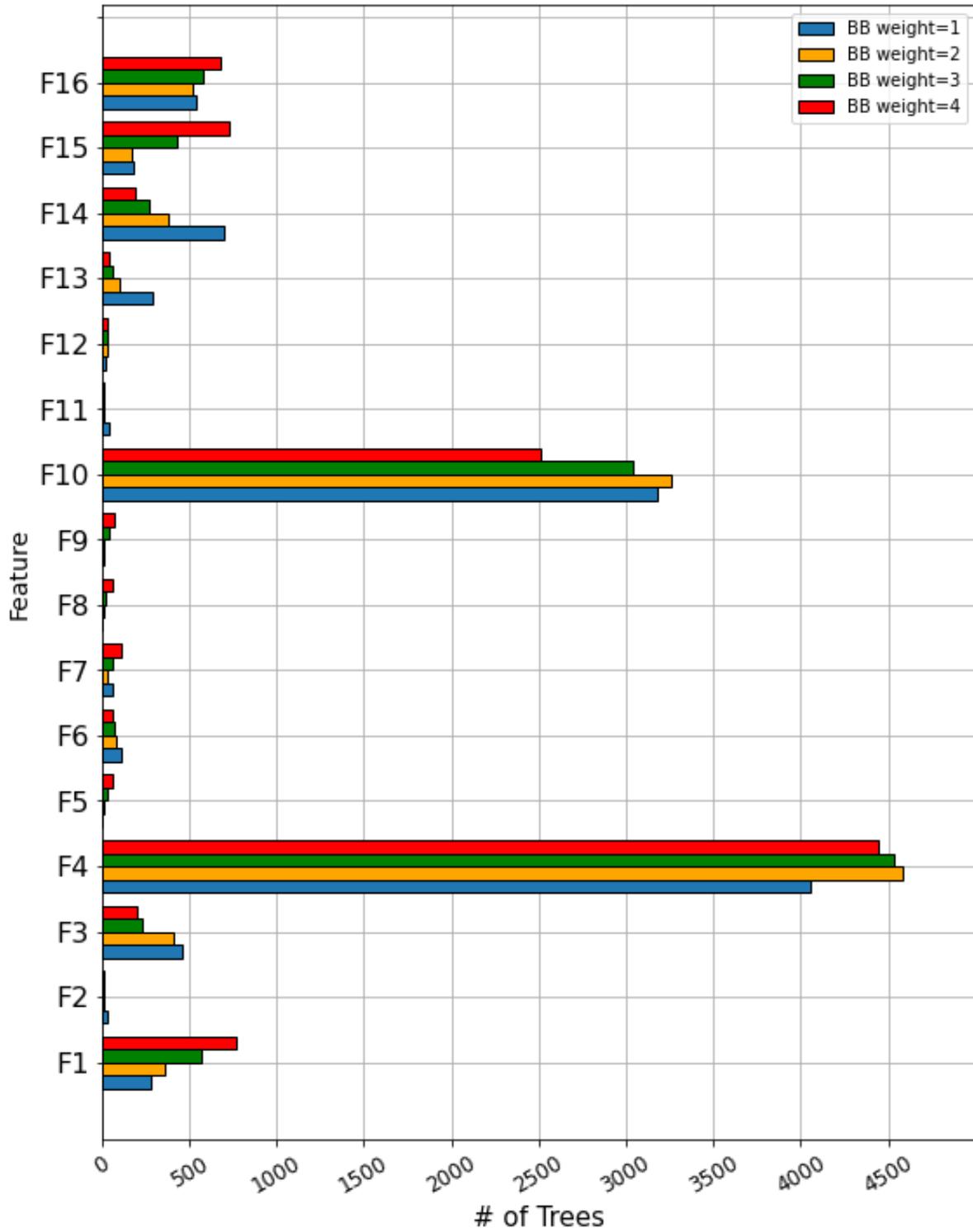


Fig. 4.5. As in Figure 4.1, but with factors identified through the PCA.

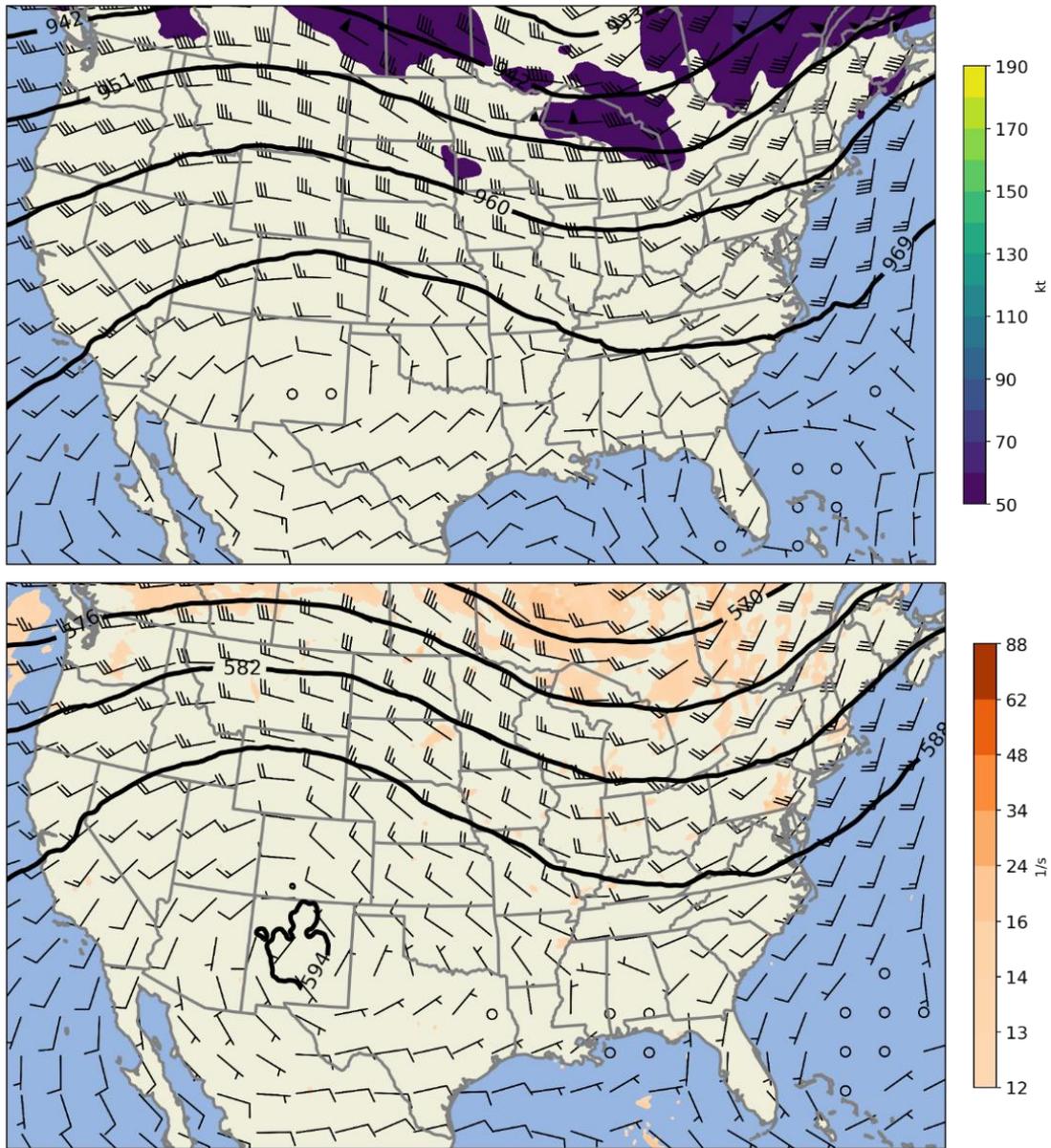


Fig. 4.6. Composite maps of back-building cases identified by the decision tree analysis with the expanded list of variables. Maps show (top) 300-hPa geopotential heights (dam), isotachs (kt) and wind (kt); and (bottom) 500-hPa geopotential heights (dam), relative vorticity ($10^{-5} s^{-1}$) and wind (kt) from 1800 UTC HRRR 0-hr analyses.



Fig. 4.7. Spaghetti plot of 500-hPa 582 dam geopotential height contours from back-building cases identified by the decision tree analysis with the expanded list of variables. The thick contour is the mean position of the 582 dam contour. Map taken from 1800 UTC HRRR 0-hr analyses.

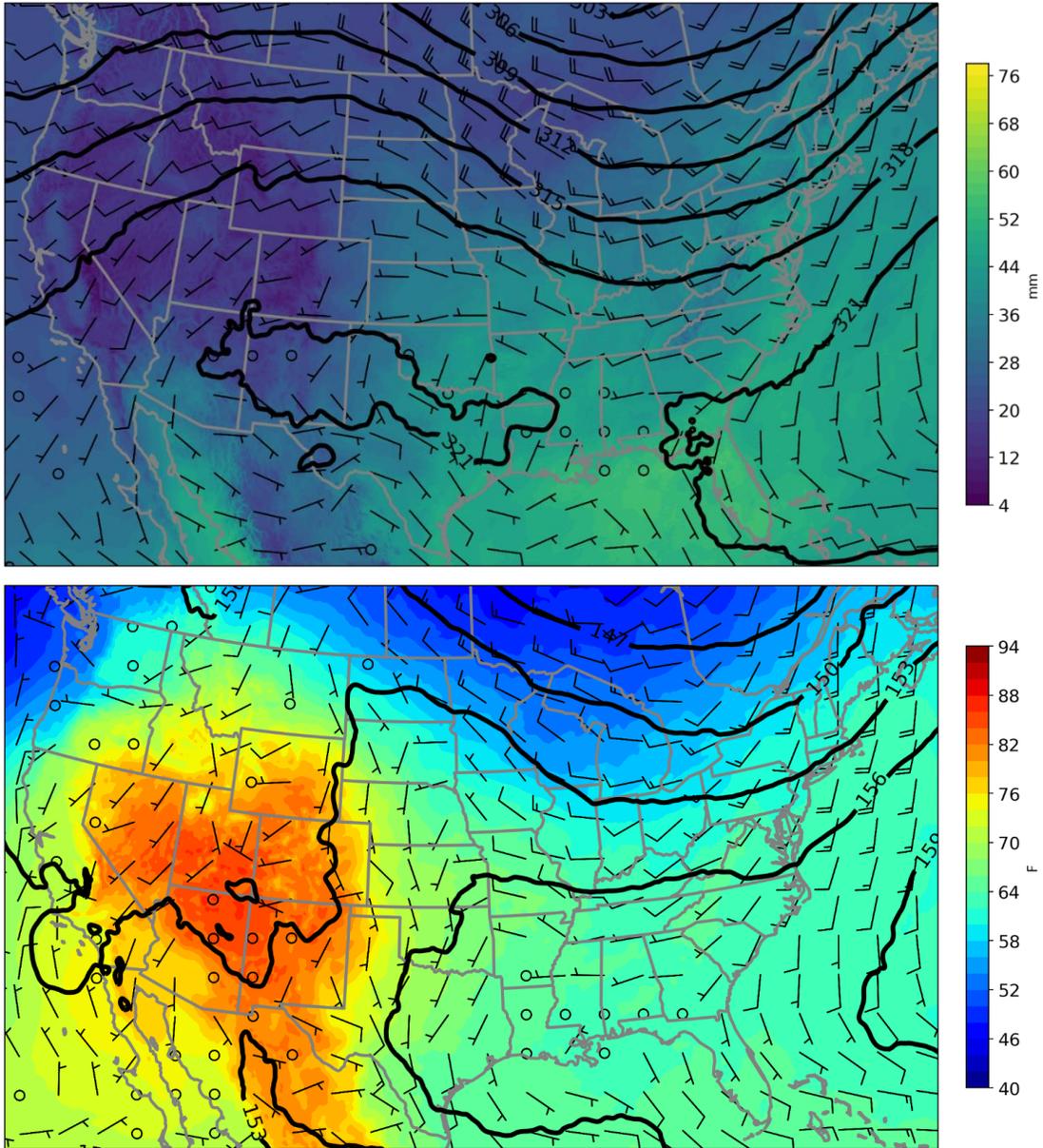


Fig. 4.8. Composite maps of back-building cases identified by the decision tree analysis with the expanded list of variables. Maps show (top) 700-hPa geopotential heights (dam), precipitable water (mm) and wind (kt); and (bottom) 850-hPa geopotential heights (dam), temperature (F) and wind (kt) from 1800 UTC HRRR 0-hr analyses.

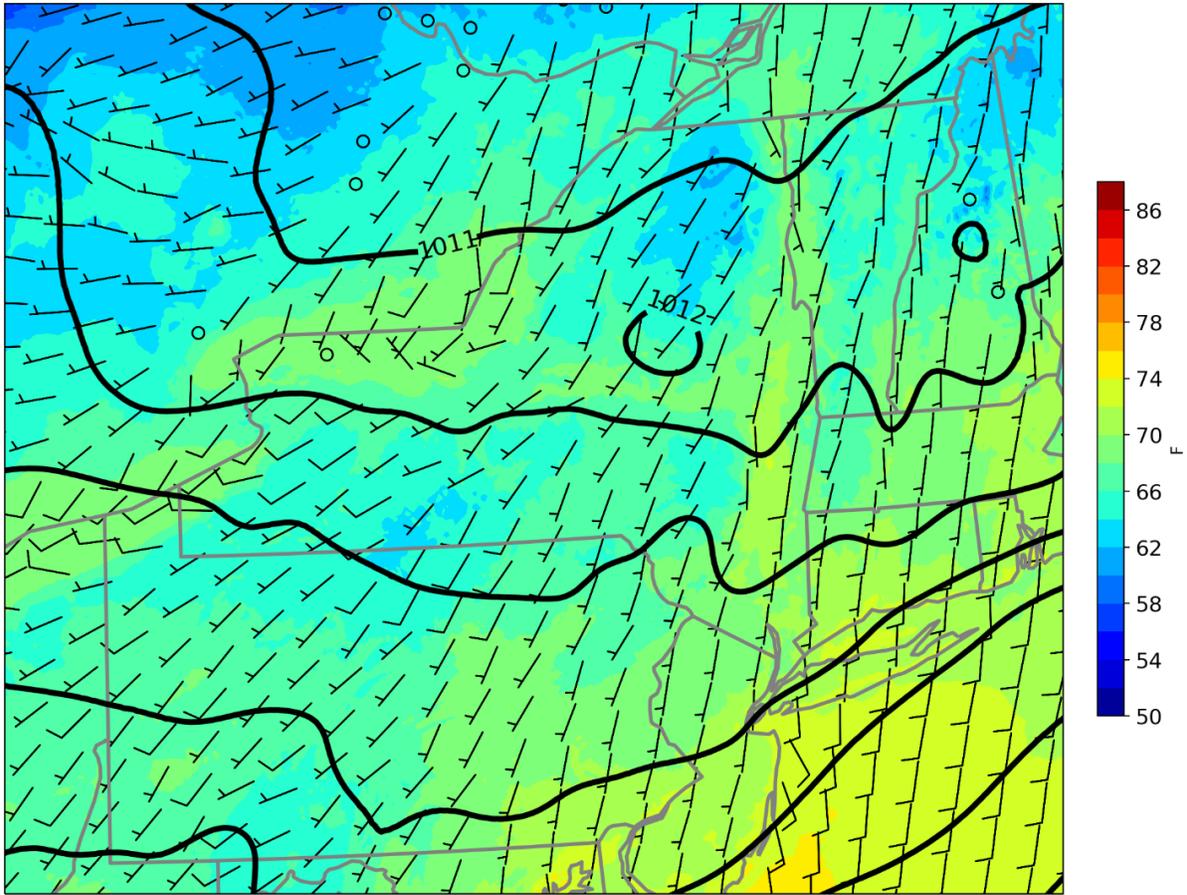


Fig. 4.9. Composite map of back-building cases identified by the decision tree analysis with the expanded list of variables. Map shows the MSLP (hPa), 2-m dew point (F) and 10-m AGL wind (kt) from 1800 UTC HRRR 0-hr analyses.

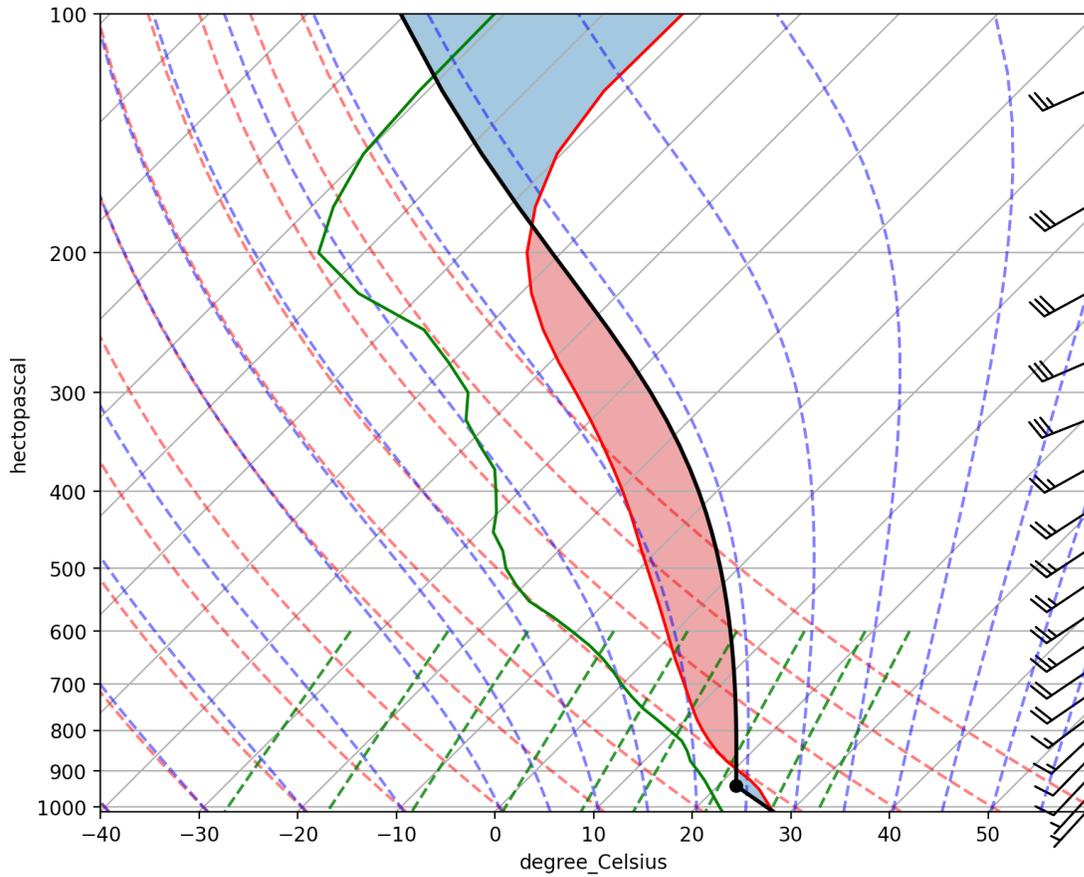


Fig. 4.10. Area averaged Hudson Valley sounding from back-building cases identified by the decision tree analysis with the expanded list of variables, taken from 1800 UTC HRRR 0-hr analyses.

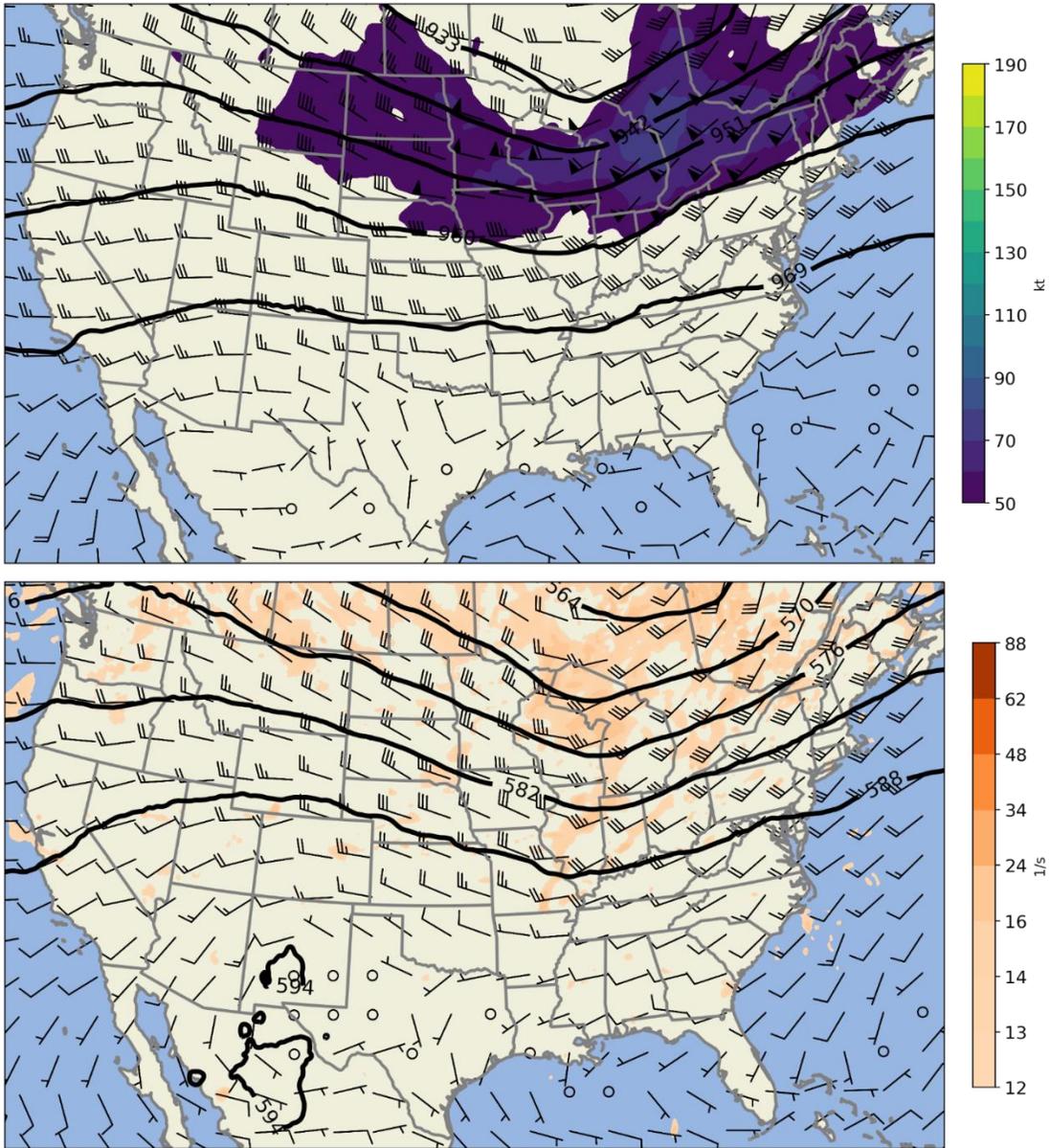


Fig. 4.11. As in Figure 4.6, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

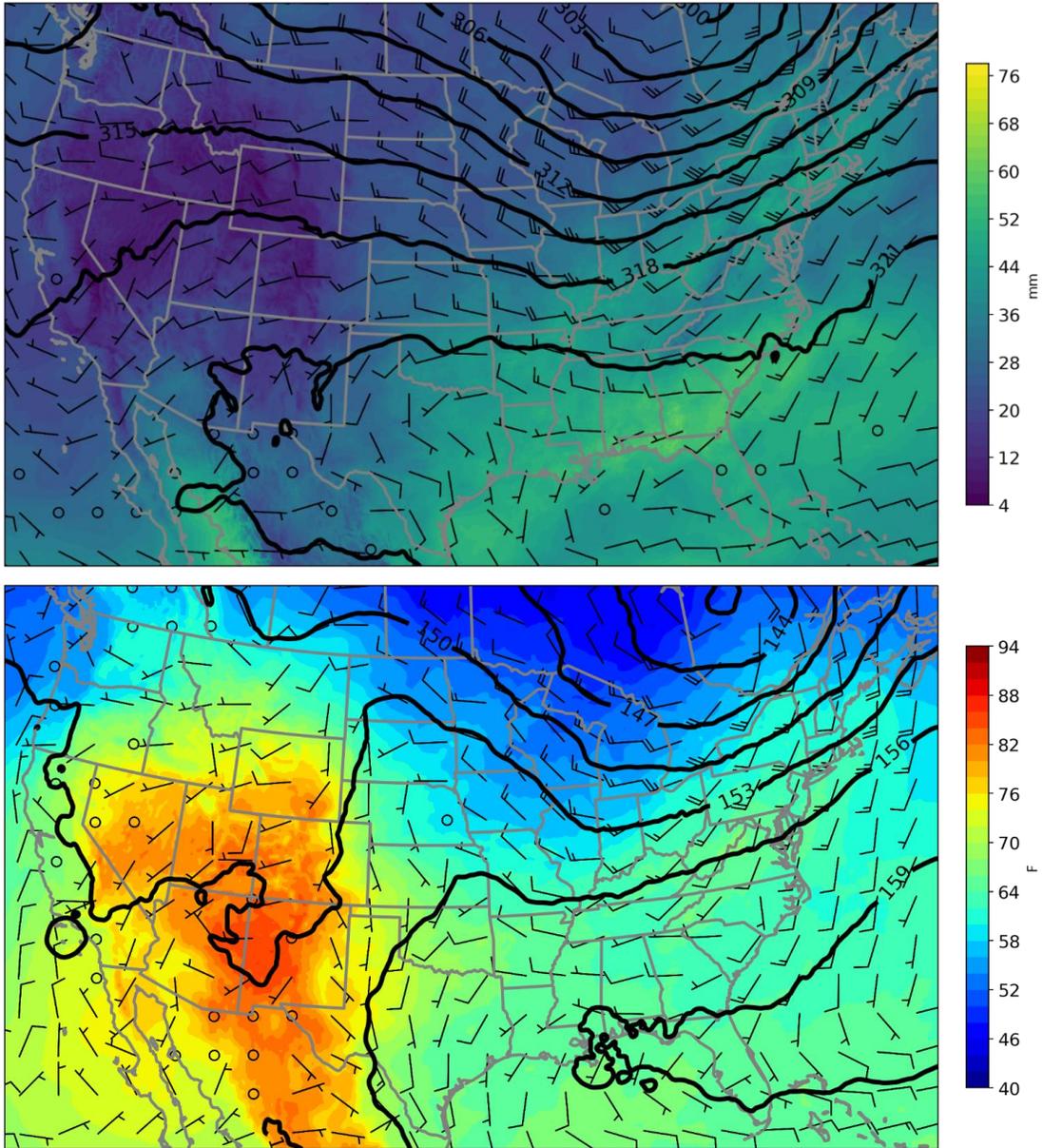


Fig. 4.12. As in Figure 4.8, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

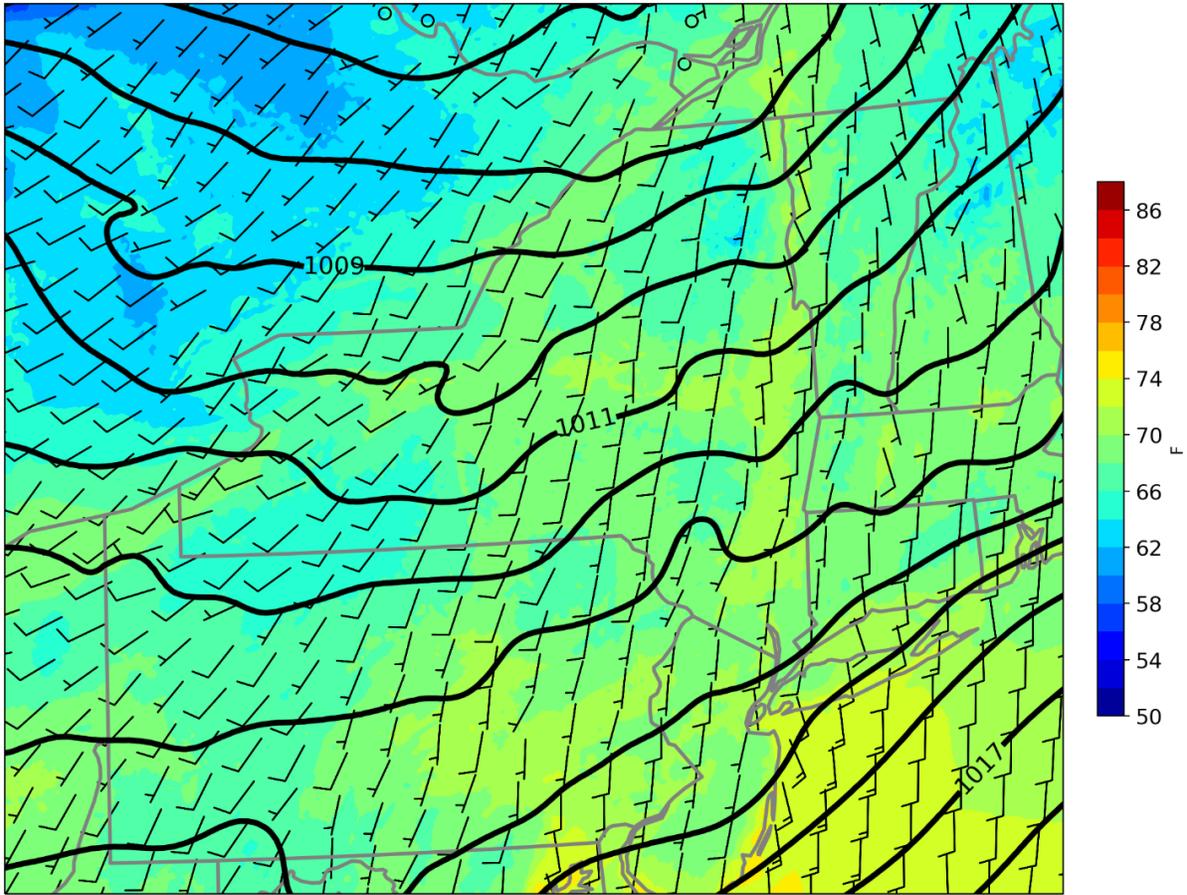


Fig. 4.13. As in Figure 4.9, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

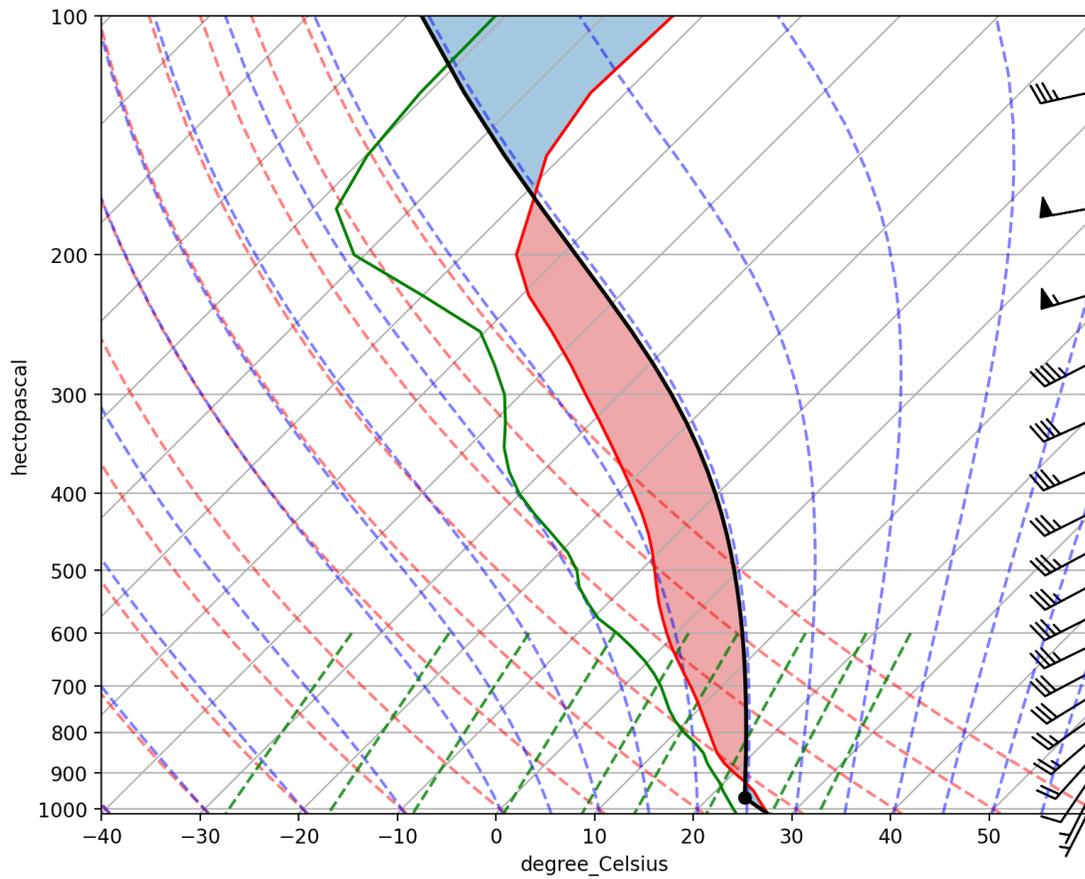


Fig. 4.14. As in Figure 4.10, but for false-positive cases identified by the decision tree analysis with the expanded list of variables.

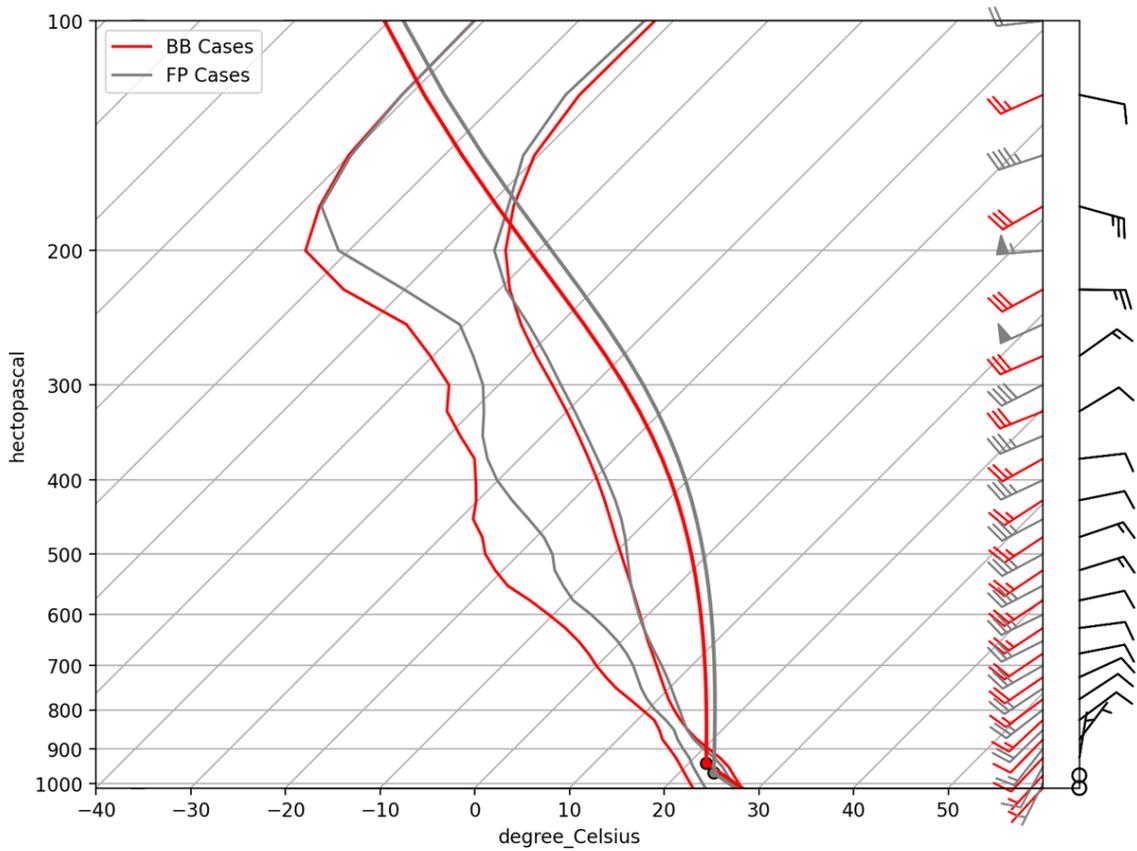


Fig. 4.15 Composite soundings for back-building (red) and false-positive (grey) cases. Parcel path for a surface-based parcel is shown with a circle indicating the LCL. The difference between the back-building and false-positive composite wind profiles is shown in barbs on the exterior axis.

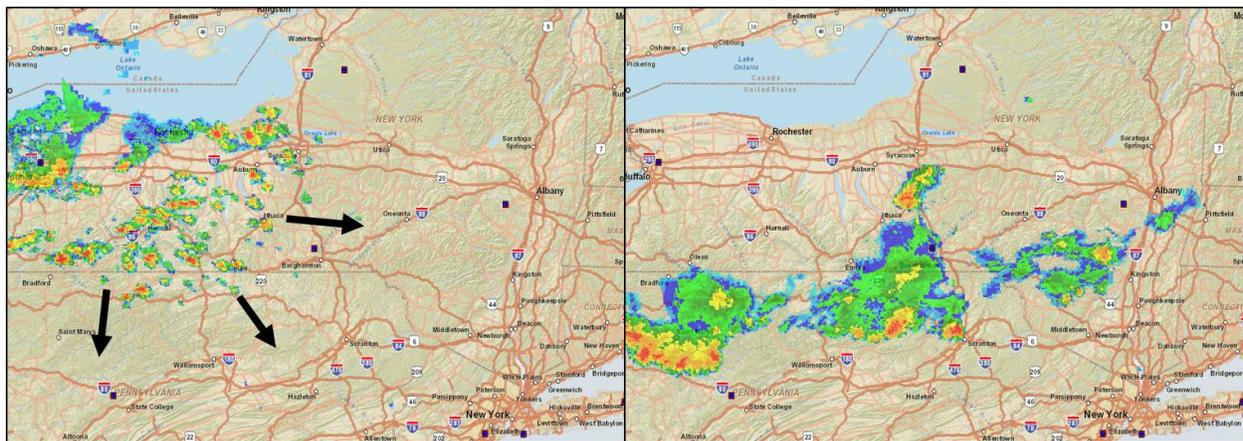


Fig. 4.16. Composite radar imagery from (left) 1925 UTC, and (right) 2320 UTC 19 July 2019. The approximate motion of convective cells is shown in the black arrows.

5. Illustrative Cases

5.1 Overview

During the writing of this thesis, there were several convective days in which the phenomenon of interest serendipitously occurred. Two of these cases, 29 and 30 June 2021, fulfilled some of the surface thermodynamic and instability parameters that were found to be conducive to back building. However, they had differences that fortuitously resulted in two different outcomes: a back-building MCS on 29 June and no back-building convection on 30 June. Examining the differences in environmental characteristics that yielded differences in convective evolution will illustrate how a forecaster may use the information provided by this research to assess the likelihood of back-building convection in the Hudson Valley.

5.2 29 June 2021

The convection on 29 June 2021 started at 2000 UTC, when a weak convective cell entered the Hudson Valley, just south of Albany (Figure 5.1). The cell quickly intensified and formed an outflow boundary on its southern edge. New cells were initiated along the boundary and the convection quickly grew upscale to form an MCS by 2040 UTC. The MCS propagated down the valley between 2040 and 2245 UTC, producing observed precipitation totals of 0.4 and 0.6 in at the Kinderhook and Schodack NYSM sites respectively. While this event was not as impactful as some of the other back-building cases studied, the pre-convective environment had some interesting overlap with the decision tree results.

The thermodynamic and vertical wind profiles found in the 1800 UTC HRRR 0-hr analysis from 29 June fulfilled most of the parameters found to be conducive to back-building convection. Figure 5.2 is the Hudson Valley area average sounding from 1800 UTC. The modeled SBCAPE was 3415 J/kg, which more than surpassed the threshold of at least 2375 J/kg. The high surface pressure and warm air advected from the southwest created 2-m temperatures as high as 94°F, as measured by the Schodack Mesonet station. If this temperature was representative of the average in the Hudson Valley, then the 2-m $T \leq 88^\circ\text{F}$ threshold was not satisfied for this case. The very high surface temperature meant that the threshold for 2-m DD ($\leq 17.5^\circ\text{F}$) was only met with the highest spikes in T_d , such as the 77°F measured by the Kinderhook Mesonet station at 1830 UTC. The wind profile was due westerly at all levels below 300-hPa, with a mean tropospheric wind magnitude of 20 kt, matching almost exactly the composite value of 20 kt.

The Albany WFO issued an area forecast discussion at 1521 UTC that acknowledge some of the features found to be conducive to back-building convection, but largely focused on the threat of high heat indexes across Upstate New York. “Synoptic forcing today is weak as the main trough is still well to our west and forcing today will be limited to [surface] convergence from terrain or lake breeze boundaries... Instability will likely be quite impressive owing to high dew points so any taller storms will be capable of producing damaging winds and heavy downpours.”

5.3 30 June 2021

On 30 June 2021, there was more intense widespread convection than 29 June, but there was no back-building convection in the Hudson Valley (Figure 5.3). A cluster of convective cells

moved from the Mohawk Valley into the upper Hudson Valley at 1835 UTC and intensified, forming an outflow boundary, similar to what was observed on 29 June 2021. However, the cells moved to the east faster than the day before. New convective cells did form along the outflow boundary, but the mature cells were already ascending the eastern edge of the Hudson Valley by the time the new cells formed. The initiation of new cells in the valley was not able to keep up with the cell motion, so the cluster of convection moved off to the east without ever becoming quasi-stationary.

The thermodynamic profile on 30 June 2021 was very similar to that found on 29 June, meeting some of the surface thermodynamic and deep instability characteristics identified in the decision tree analysis. A special sounding was launched from KALY at 1800 UTC, recording a surface temperature of 92°F and DD of 18°F, both of which exceeded the maximum value for back building according to the decision tree analysis (Figure 5.4). The measured SBCAPE in the Hudson Valley was 4659 J/kg, almost double the minimum value (2375 J/kg). Though there was little change in wind direction from 29 June 2021, the magnitude of the wind increased at all levels. The measured cloud-layer wind was 35 kt and the HRRR analysis mean tropospheric wind was 31 kt, well exceeding the typical value for back-building cases (20 kt). The increased wind speed caused the cell motion to be faster, which seemed to be the most meaningful difference for the convective evolution on 29 and 30 June 2021.

The area forecast discussion from the Albany WFO at 1432 UTC shows that forecasters were expecting a greater risk of severe convection on 30 June 2021, due to the stronger synoptic forcing and higher shear. “Forecast hodographs suggest multi-inflection points for clusters of convection as mode-types this afternoon.” Heavy rainfall was no longer one of the main concerns as the faster cell motion precludes any cell staying in one location for an extended

period of time. The main threat expected on this day was instead the mixing of strong winds down to the surface.

5.4 Summary

29 and 30 June 2021 had very similar surface environments within the Hudson Valley, but differences in their wind profile that ultimately lead to different outcomes. The high T_d at the surface on both days allowed the convection to intensify and form an outflow boundary once entering the valley. 29 June had a mean tropospheric wind of 20 kt, which resulted in the individual cell motion being slow enough for the initiation of new cells on the upstream side to cancel it out. The convective line exhibited back-building behavior such that the convection propagated down the Hudson Valley. The faster winds (31 kt) on 30 June resulted in a cell motion that was too fast for the initiation of new cells to cancel out and so the convection quickly moved to the east and out of the Hudson Valley. These cases highlight the fact that the surface thermodynamics and instability parameters are an indicator of the potential for back building, but may not be enough information on their own to forecast back building. Forecasters must also take the wind profile into account in order to improve their situational awareness on days with potential for back-building convection.

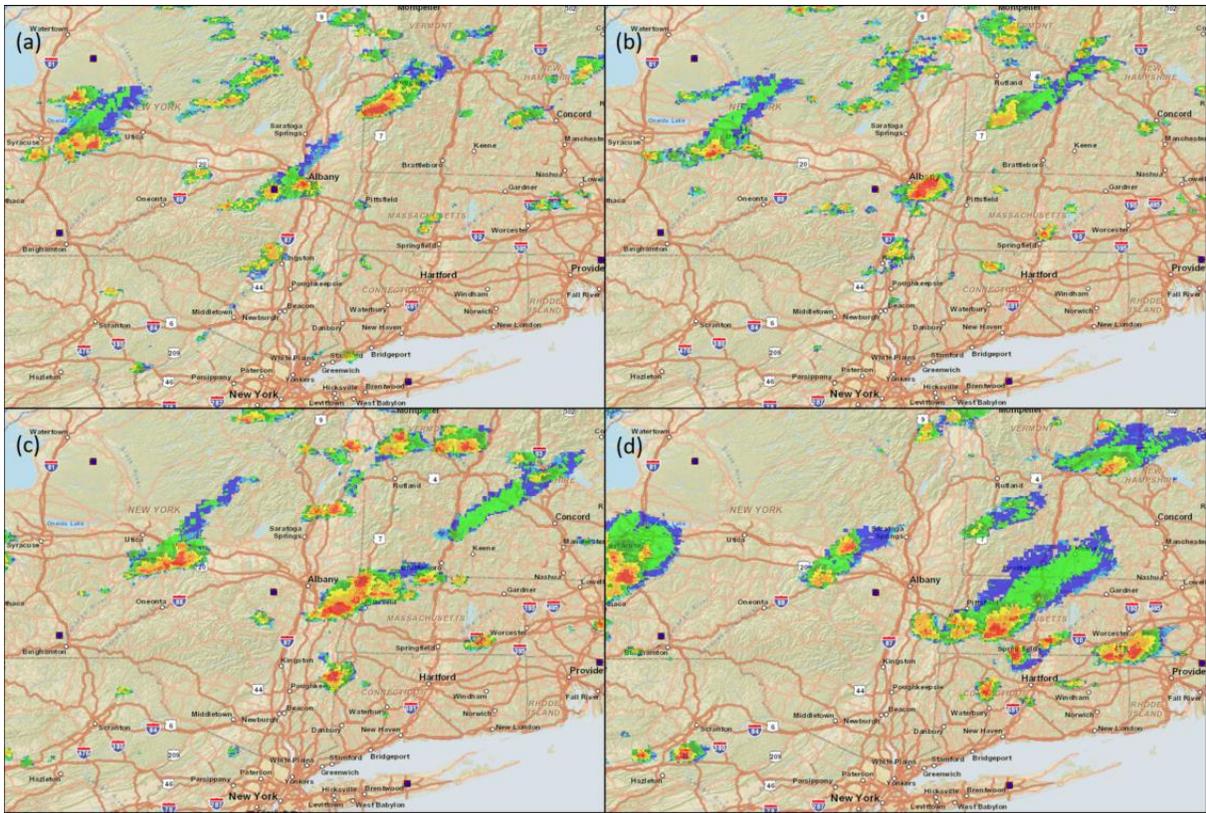


Fig. 5.1. Composite radar imagery depicting a back-building MCS in the Hudson Valley from 29 June 2021 at (a) 2005 UTC, (b) 2040 UTC, (c) 2125 UTC and (d) 2235 UTC.

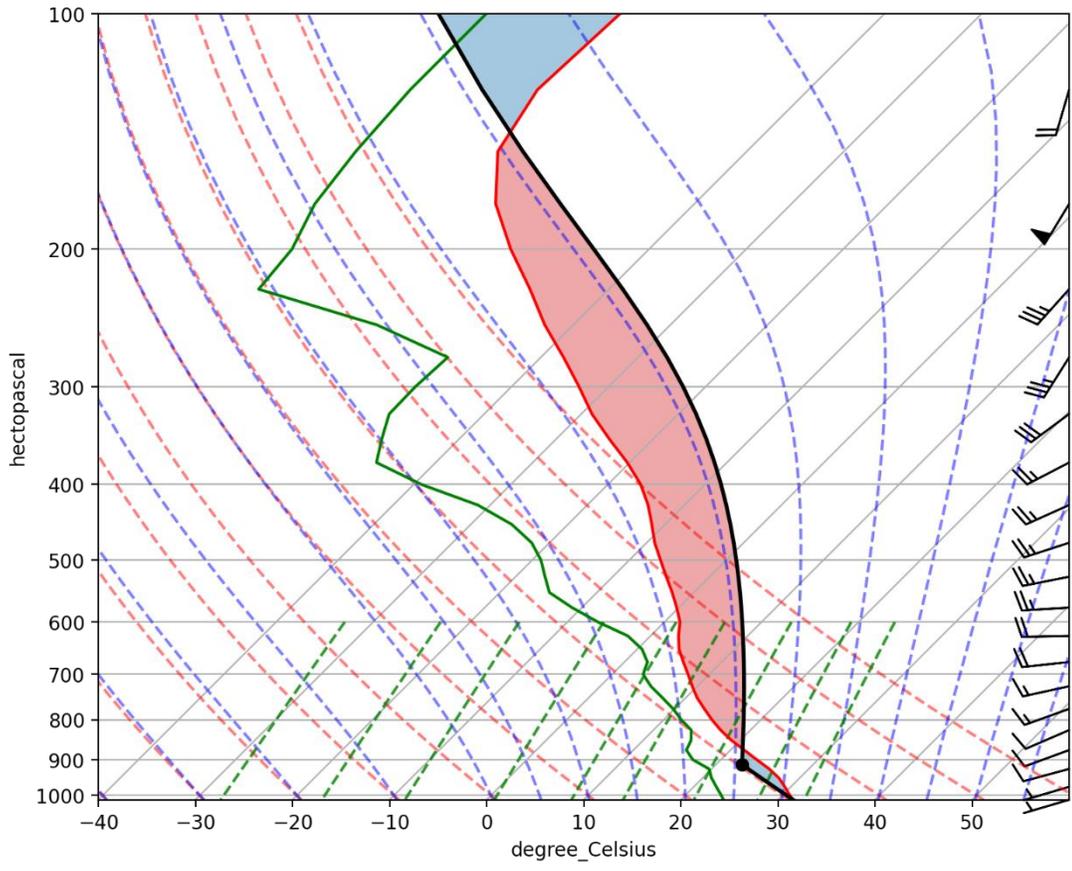


Fig. 5.2. Area-averaged Hudson Valley sounding valid at 1800 UTC 29 June 2021. Taken from HRRR 0-hr analysis.

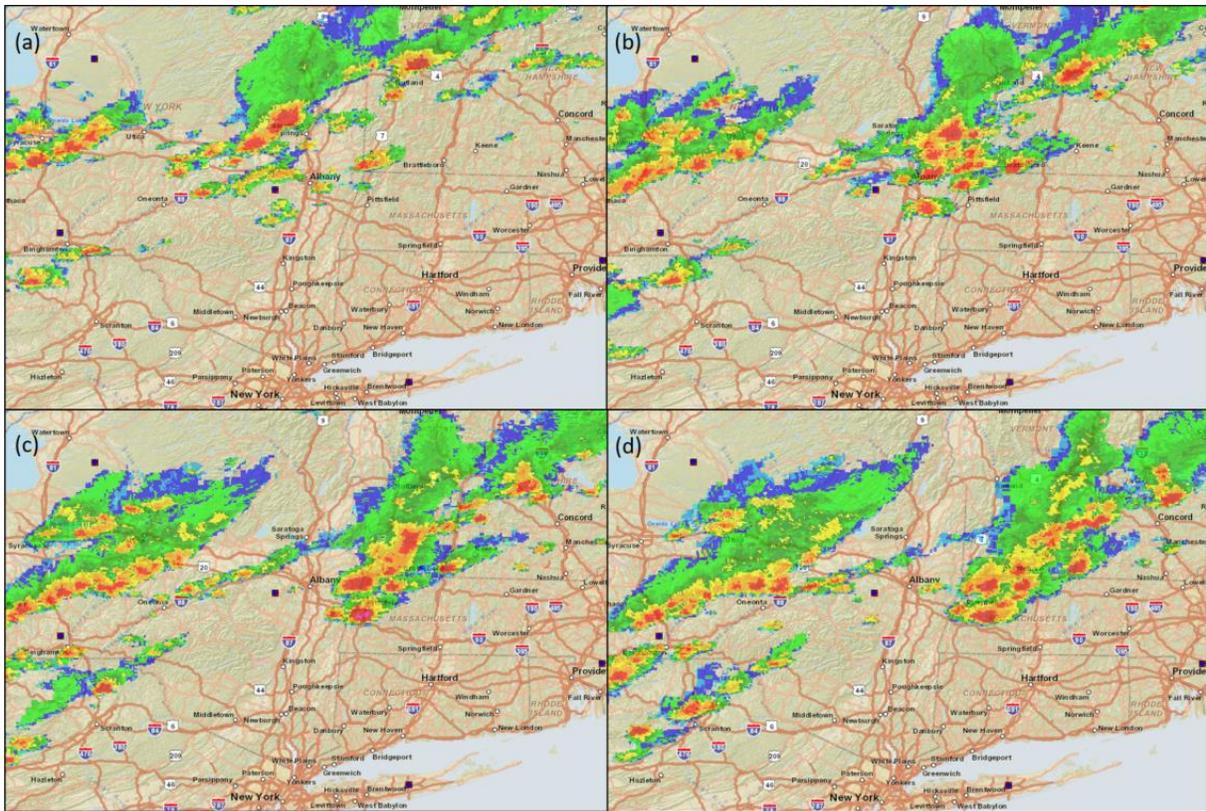


Fig. 5.3. Composite radar imagery showing convection in the Hudson Valley from 30 June 2021 at (a) 1835 UTC, (b) 1920 UTC, (c) 1950 UTC and (d) 2015 UTC.

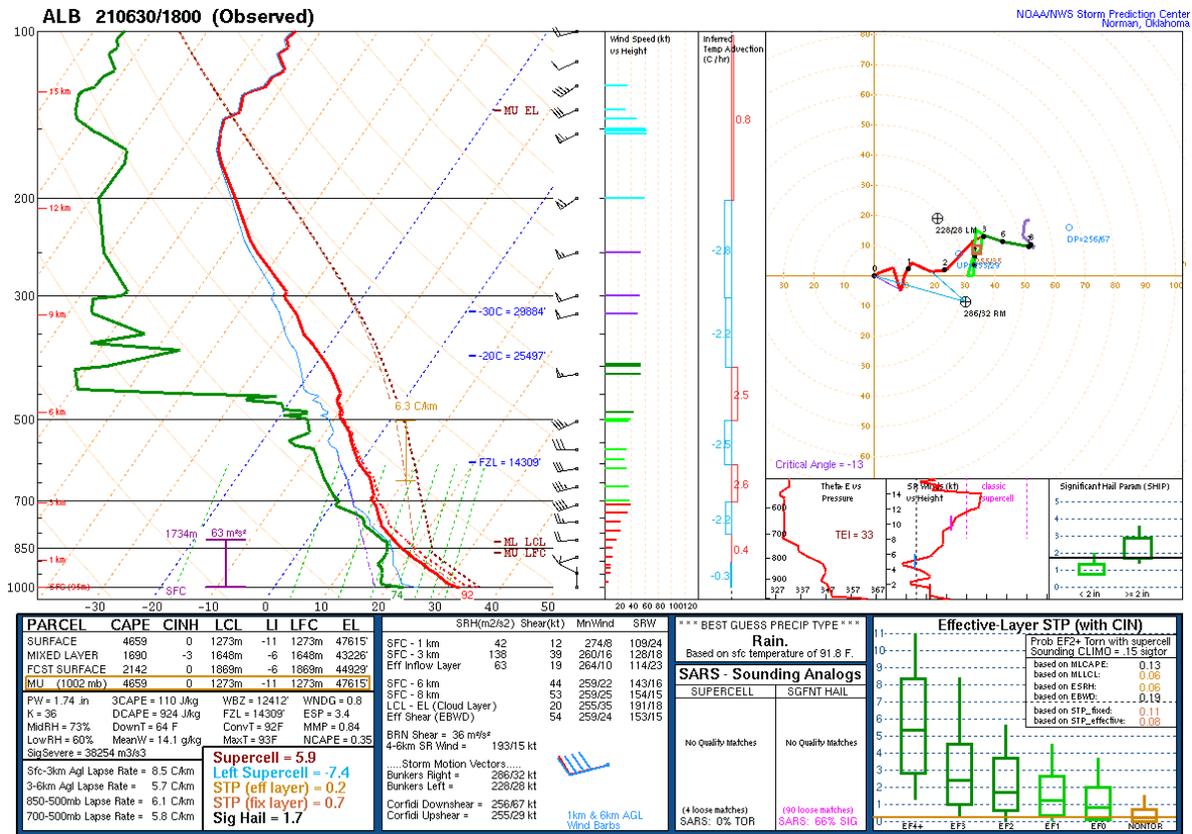


Fig. 5.4. Sounding launched from KALY at 1800 UTC 30 June 2021. [Reproduced from SPC Observed Sounding Archive <https://www.spc.noaa.gov/exper/soundings/>]

6. Conclusions and for Future Work

6.1 Conclusions

The goal of this research was to improve the understanding and prediction of back-building MCSs in the Hudson Valley of Upstate New York. Though the structure (Parker 2000), maintenance (Schumacher 2005) and propagation (Corfidi 1996; 2003) of back-building MCSs has been studied extensively in the Great Plains, comparatively little research has focused on back-building MCSs in the Northeast. WRF simulations of a back-building MCS case from 21 August 2019 were used to study how interactions between the terrain, background flow and convective cold pool lead to the upscale growth of convection. HRRR 0-hr analyses were collected for convective days from June, July and August of 2015–2020 and classified by whether or not back building occurred in the Hudson Valley. Decision trees were then constructed to identify parameters that made the pre-convective environment in the Hudson Valley conducive to back building. Two cases from June 2021 were compared to illustrate how the decision tree results could be used to assess the likelihood of back-building convection.

6.1.1 Physical Mechanisms

The physical mechanism responsible for back-building convection in the Hudson Valley is very similar to that found in the Great Plains, as described in Schumacher et al. (2005). Once an outflow boundary is created through evaporatively cooled air collecting beneath the convection, a positive feedback loop begins. High θ_e air is lifted over the gust front, initiating a new convective cell. That cell matures as it passes over the cold pool, reinforcing it with evaporatively cooled air. The evaporatively cooled air falls to the surface and expands outward,

progressing the gust front down the Hudson Valley, and initiating the next convective cell.

Ultimately, the positive feedback loop is broken when the inflow is disrupted or becomes too weak to initiate new cells.

Back-building MCSs in the Hudson Valley tend to be smaller and more short-lived than those in the Great Plains. The back-building MCSs tend to be oriented nearly perpendicular to the Hudson Valley as to maximize the surface convergence between the northerly winds in the cold pool and southerly winds upstream. The upper Hudson Valley is ~40 km wide which provides an upper limit for the typical length of the back-building MCS, as opposed to the ~150 km length scale in the Great Plains (Schumacher 2005). The back-building MCSs studied in this research often persisted for 2–3 hr before dissipating or merging with another convective line. In contrast, past studies set in the Great Plains have used a minimum time scale of 3 hr to qualify a convective line as an MCS (Parker 2000).

The channeling of background flow by the terrain played two important roles in the development and maintenance of the simulated back-building MCS. The southerly flow within the Hudson Valley facilitated a low-level maximum in water vapor flux between ~1600 UTC and 1800 UTC. This led to an increase in 2-m T_d which made the Hudson Valley and more favorable location for convection than the higher terrain to the west, allowing a strong convective cold pool to form. Once the MCS had formed the cold pool, the terrain-channeled flow continued to flux water vapor up the valley and increased the low-level convergence along the gust front.

6.1.2 Prediction

The characteristics of the pre-convective environment in the Hudson Valley identified as conducive to back building were dependent on the variables entered into the decision tree classifier. The short list of features, consisting of surface and mid-level thermodynamic variables, surface wind variables and severe-weather indices, resulted in the 2-m DD and SBCAPE being chosen most often. An expanded list of features included more wind variables at 500, 700, and 850 hPa; more thermodynamic variables, and more severe-weather indices. LI became the most commonly chosen feature, followed by the 2-m DD, 2-m T and 850-hPa T_d .

While the decision tree classifier did not choose the same features from each list, the results were different representations of the same physical characteristic. LI and SBCAPE were both chosen by many of the decision trees, as they represent the potential for a parcel lifted from the surface to grow into a deep convective cell. Deep convection is necessary to produce enough evaporatively cooled air for strong cold pools to form and act as a lifting mechanism. The 2-m DD was chosen by from both lists of features to represent the surface thermodynamic environment within the Hudson Valley. A low 2-m DD means a parcel lifted from the surface is more likely to be positively buoyant as it will quickly reach saturation and cool at the moist adiabatic lapse rate as it rises.

The remaining features that were chosen by the decision trees, a low 850-hPa T_d and low 2-m T, do not have obvious interpretations. A low 850-hPa T_d may lead to more evaporative cool if precipitation falls through the layer and lead to stronger cold pools. However, dry air at the 850-hPa level would also be entrained into the updraft of any growing convective cells, which

would decrease the positive buoyancy and therefore the likelihood of the cell growing into deep convection. A low 2-m T may make surface air with the same moisture content become saturated faster, similar to the 2-m DD. However, a low 2-m T would also decrease the T at all levels along the parcel's path, decreasing the instability. Lastly, it's possible that these features being chosen is a consequence of the small sample size used in the decision tree analysis and do not represent a meaningful characteristic of the pre-convective environment.

The decision tree classifier was run with different weights assigned to the back-building cases in order to balance the amount of missed events with low impurities. When the back-building cases were given the same weight as cases without back building, the impurity scores were often low, meaning few false-positive cases were identified by the decision tree. However, the thresholds were highly restrictive, resulting in a large number of back-building cases not falling into the area of the phase space that was conducive to back building. As the back-building cases were given more weight in the decision tree classifier, the impurities were higher, but the number of missed events decreased.

When the correctly-identified back-building cases were compared to the false-positive cases, the mean tropospheric wind for back-building cases was 10 kt slower. The back-building cases had a mean tropospheric wind of 20 kt, which allows the individual convective cells to stay in the Hudson Valley for longer, leading to stronger convective cold pools. The false-positives had a mean tropospheric wind of 30 kt, making it more likely that the convective cells will move out of the Hudson Valley before creating enough evaporative cooling for a strong cold pool. This point was highlighted in the comparison of the back-building MCS case from 29 June 2021 and no-back-building case 30 June 2021. The surface thermodynamic environments were similar on

both days, but the magnitude of the mean tropospheric wind made the difference as to whether back building occurred.

6.2 Future Work

The portion of this research focused on predicting back-building MCSs looked at the pre-convective environment as represented in the HRRR 0-hr analyses. This perspective was chosen to provide forecasters with parameters that could be forecasted with some amount of lead time based on model data. However, this perspective does not make use of the surface observations taken by ASOS and NYSM stations in the Hudson Valley. Further investigation could focus on using surface observations to assess the pre-convective environment, including the surface moisture content and magnitude of southerly surface winds.

Another avenue of future research could focus more on the small-scale structure of the convective cold pools under back-building MCSs in the Hudson Valley. Past research by Houston et al. (2011) argued that the depth of the cold pool is an important metric for the strength and longevity of convective lines. The scanning Doppler LiDAR found at the NYSM profiler sites in Albany and Red Hook could be a useful source of observational data to investigate this claim.

Terrain channeling has been found by this research and past studies (Riley and Bosart 1983) to have a profound impact on the strength and evolution of severe convection. Past numerical studies on channeled flow in valleys have used a variety of horizontal grid spacing in their simulations. Whiteman et al. (1993) used a horizontal grid spacing of 11 km to resolve terrain-induced flows in the Tennessee Valley, which is roughly 70 km wide and has terrain

reaching 700–1700 m above the valley floor on either side. Gross et al. (1987) used a horizontal grid spacing of 1.2 km for the upper Rhine Valley, which has dimensions of 35 km wide and ~500 m high terrain, more similar to the Hudson Valley. Future research could focus on assessing how well operational convection-allowing models resolve valley flows and if better resolution of terrain induced flows would impact the strength and mode of forecasted convection.

REFERENCES

- Abdullah, A. J., 1954: The Meridional Growth of Squall Lines. *J. Meteor.*, **11**, 301–308.
- Benjamin, S. G., and P. A. Miller, 1990: An Alternative Sea Level Pressure Reduction and a Statistical Comparison of Geostrophic Wind Estimates with Observed Surface Winds. *Mon. Wea. Rev.*, **118**, 2099–2116.
- Bluestein, H. B., and M.H. Jain, 1985: Formation of Mesoscale Lines of Precipitation: Severe Squall Lines in Oklahoma during the Spring. *J. Atmos. Sci.*, **42**, 1711–1732.
- Carrera, M. L., J. R. Gyakum, and C. A. Lin, 2009: Observational Study of Wind Channeling with the St. Lawrence River Valley. *J. Appl. Meteor. Climatol.*, **48**, 2341–2361.
- Corfidi, S. F., J. H. Merritt, and J. M. Fritsch, 1996: Predicting the Movement of Mesoscale Convective Complexes. *Wea. Forecasting*, **11**, 41–46.
- , 2003: Cold Pools and MCS Propagation: Forecasting the Motion of Downwind-Developing MCSs. *Wea. Forecasting*, **18**, 997–1017.
- Freedman, J. M., and D. R. Fitzjarrald, 2017: Mechanisms Responsible for the Observed Thermodynamic Structure in a Convective Boundary Layer over the Hudson Valley of New York State. *Bound.-Layer Meteor.*, **164**, 89–106.
- Gross, G., and F. Wippermann, 1987: Channeling and Countercurrent in the Upper Rhine Valley: Numerical Simulations. *J. Appl. Meteor. Climatol.*, **26**, 1293–1304.
- Henz, J. F., 1973: Characteristics of Severe Convective Storms on Colorado's High Plains. *Preprints, Eighth Conf. on Severe Local Storms*, Denver, CO, Amer. Meteor. Soc., **105**, 96–103.
- Hurlbut, M. M., and A. E. Cohen, 2014: Environments of Northeast U.S. Severe Thunderstorm Events from 1999 to 2009. *Wea. Forecasting*, **29**, 3–22.
- Iowa Environmental Mesonet, 2001: NWS Processed Datasets, Accessed 8 November 2019, <https://mesonet.agron.iastate.edu/nws/>.
- Jessup, S. M., and S. J. Colucci, 2011: Organization of Flash-Flood-Producing Precipitation in the Northeast United States. *Wea. Forecasting*, **27**, 345–361.
- Katona, B., P. Markowski, C. Alexander, and S. Benjamin, 2016: The Influence of Topography on Convective Storm Environments in the Eastern United States as Deduced from the HRRR. *Wea. Forecasting*, **31**, 1481–1490.

Muccilli, M.: Using the Froude Number to Improve Westerly Flow Upslope Snow Forecasts in the Green Mountains of Vermont. National Weather Service, Accessed 24 January 2020, <https://www.weather.gov/btv/froude>.

National Oceanic and Atmospheric Administration, 2016: High Resolution Rapid Refresh (HRRR). Google Cloud, Accessed 24 February 2021, <https://console.cloud.google.com/marketplace/product/noaa-public/hrrr?project=python-232920&pli=1>.

National Centers for Environmental Information, 1999: Storm Events Database, Accessed 8 November 2019, <https://www.ncdc.noaa.gov/stormevents/>

_____, #####: Stage IV Quantitative Precipitation Estimate (QPE). Research Data Archive, Accessed 23 September 2020.

Newton, C. W., 1963: Dynamics of Severe Convective Storms. *Meteor. Monogr. Amer. Meteor. Soc.* **27**, 33–58.

_____, and J. C. Fankhauser, 1964: On the Movements of Convective Storms, with Emphasis on Size Discrimination in Relation to Water Budget Requirements. *J. Appl. Meteor.*, **3**, 651–668.

_____, and S. Katz, 1958: Movements of Large Convective Rainstorms in Relation to Winds Aloft. *Bull. Amer. Meteor. Soc.*, **39**, 129–136.

Parker, M. D., and R. H. Johnson, 2000: Organizational Modes of Midlatitude Mesoscale Convective Systems. *Wea. Forecasting*, **128**, 3413–3436.

Razy, A., S. M. Milrad, E. H. Atallah, and J. R. Gyakum, 2012: Synoptic-Scale Environments Conducive to Orographic Impacts on Cold-Season Surface Wind Regimes at Montreal, Quebec. *J. Appl. Meteor. Climatol.*, **51**, 598–616.

Riley, G. T., and L. F. Bosart, 1983: The Windsor Locks, Connecticut tornado of 3 October 1979: An analysis of an intermittent severe weather event. *Mon. Wea. Rev.*, **115**, 1655–1677.

Sakai, R. K., D. R. Fitzjarrald, M. Czikowsky, and J. M. Freedman, 2004: Mechanisms of Wind Channeling in the Hudson Valley, NY. *11th Conf. on Mountain Meteor.*, Bartlett, NH, Amer. Meteor. Soc., P4.5.

Schumacher, R. S., and R. H. Johnson, 2005: Organization and Environmental Properties of Extreme-Rain-Producing Mesoscale Convective Systems. *Mon. Wea. Rev.*, **133**, 961–976.

Stout, G. E., and H. W. Hiser, 1955: Radar Scope Interpretations of Wind, Hail, and Heavy Rainstorms between 27 May and 8 June 1954. *Bull. Amer. Meteor. Soc.*, **36**, 519–527.

- Thompson, R. L., and J. S. Grams, 2008: Synoptic Environments and Convective Modes Associated with Significant Tornadoes in the Contiguous United States. *24th Conf on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 16A.3.
- Vaughan, M. T., B. H. Tang, and L. F. Bosart, 2017: Climatology and Analysis of High-Impact, Low Predictive Skill Severe Weather Events in the Northeast United States. *Wea. Forecasting*, **32**, 1903–1919.
- Wasula, A. C., L. F. Bosart, and K. D. LaPenta, 2002: The Influence of Terrain on the Severe Weather Distribution across Interior Eastern New York and Western New England. *Wea. Forecasting*, **17**, 1277–1289.
- Whiteman, D. C., 1990: Observations of Thermally Developed Wind Systems in Mountainous Terrain. *Meteorological Monographs*,. *American Meteorological Society*, **23**, 5–42.
- , and J. C. Doran, 1993: The Relationship between Overlaying Synoptic-Scale Flows and Wind within a Valley. *J. Appl. Meteo.*, **32**, 1669–1682.