Assessment of Model Forecast Temperature Bias During Cold Air Damming in the Central Appalachian Mountains

Suzanna Lindeman

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of
Master of Science
In
Geography

Andrew Ellis
Stephen Keighton
David Carroll

April 2018
Blacksburg, VA

Keywords: cold-air damming, Model Output Statistics, temperature bias, central Appalachian Mountains

Assessment of Model Forecast Temperature Bias During Cold Air Damming in the Central Appalacian Mountains

Suzanna Lindeman

Abstract

Cold-air damming (CAD) is a prevalent Mid-Atlantic United States weather phenomenon that occurs when cold, dense air is dammed alongside the eastern slopes of the Appalachian Mountains. Lower-than-normal maximum temperatures, increased and prolonged cloud cover, and precipitation that produces hazardous impacts are common features of this weather event, which are well known for presenting difficulties to both human forecasters and weather prediction models. This study explores CAD events between 2007 and 2016 archived in a Blacksburg National Weather Service 'bust' database – instances when forecasters erred by at least 8°F (4.4°C) on either maximum or minimum daily air temperature. The database includes the temperature error within Model Output Statistics (MOS) guidance in association with these forecast 'busts.' During the 10-year study period, MOS guidance produced warm-biased maximum temperatures and cold-biased minimum temperatures for most of the problematic CAD events, suggesting MOS guidance tended to underestimate the strength of CAD in these cases, seeming to struggle with weaker CAD events. During CAD erosion, MOS tended to prematurely erode CAD scenarios at night and predicted them to persist for too long during the day. Hourly surface meteorological and synoptic atmosphere composites during these 'busted' CAD events failed to reveal obvious differences from what is expected for central Appalachian CAD. However, a comparison to well-forecast classic cold-season CAD events suggest that busted cases of this same type of CAD may be drier than is typical. As the atmospheric patterns associated with busted CAD events are typical of the phenomenon, but a bit weaker or more marginal, forecast errors appear to stem from subtle model errors rather than forecaster error. It is possible that the models may inadequately characterize low-level moisture, but further research is needed to isolate the source of model forecast error. Nonetheless, the results of this research serve as guidance for operational forecasters as they consider model guidance during weak CAD events.

Assessment of Model Forecast Temperature Bias During Cold Air Damming in the Central Appalacian Mountains

Suzanna Lindeman

General Audience Abstract

Cold-air damming (CAD) is a common weather pattern that affects the Blue Ridge Mountain region of the eastern United States, in which cold air at the atmosphere's surface is directed from the Northeast and is dammed against the eastern Appalachian Mountains. This weather event causes lower-than-normal temperatures over the region and is often characterized by prolonged cloudy skies and precipitation. CAD is very difficult for forecasters to accurately predict, as they rely on weather forecast models that often simulate these situations poorly. CAD also strains emergency managers who rely on accurate forecasts to support public safety during CAD. This study explores CAD events between 2007 and 2016 archived in a Blacksburg National Weather Service 'bust' database – instances when forecasters erred by at least 8°F (4.4°C) on either maximum or minimum daily air temperature. The database includes the temperature error within Model Output Statistics (MOS) guidance in association with these forecast 'busts.' During the 10-year study period, MOS guidance forecasted maximum temperatures too high and minimum temperatures too low for most of the problematic CAD events, suggesting MOS guidance tended to underestimate the strength of CAD in these cases, seeming to struggle with weaker CAD events. During instances where CAD dissolved from the Appalachians, MOS tended to prematurely erode CAD scenarios at night and predicted them to persist for too long during the day. Hourly surface meteorological and synoptic atmosphere composites during these 'busted' CAD events failed to reveal obvious differences from what is expected for central Appalachian CAD. However, a comparison to well-forecast classic cold-season CAD events suggest that busted cases of this same type of CAD may be drier than is typical. As the atmospheric patterns associated with busted CAD events are typical of the phenomenon, but a bit weaker or more marginal, forecast errors appear to stem from subtle model errors rather than forecaster error. It is possible that the models may inadequately characterize low-level moisture, but further research is needed to isolate the source of model forecast error. Nonetheless, the results of this research serve as guidance for operational forecasters as they consider model guidance during weak CAD events.

Acknowledgements

Foremost, I express my immense gratitude towards my advisor Prof. Andrew Ellis for his continued support throughout both my undergraduate and graduate careers at Virginia Tech. His patience, expansive knowledge, and wiz programming skills were invaluable in both researching and writing this thesis. I could not have imagined a better-suited mentor for my Master's thesis and would not have succeeded without his immeasurable help.

My gratitude extends to the other members of my thesis committee, Stephen Keighton and David Carroll. Their kind words and instrumental feedback greatly guided and motivated me throughout this research process.

I also spread thanks to the National Weather Service Forecast Office in Blacksburg, Virginia for their constant love and support in helping me through my Master's degree. Feedback from expert operational meteorologists has been massively helpful in this research, and I especially thank Robert Stonefield for his kind heart and extremely valuable help as a cold-air damming expert.

Lastly, I thank my family and friends. Without them, none of this would have been possible.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

In order of appearance:

CAD	Cold-air Damming
NAM	North American Mesoscale Forecast System
RAP	Rapid Refresh Model
GFS	Global Forecast System Model
ECMWF	European Medium Range Weather Forecast Model
NWSFO	National Weather Service Forecast Office
CWA	County Warning Area
MOS	Model Output Statistics
MAV	GFS MOS Short-range Text Product
MET	NAM MOS Text Product
QPF	Quantitative Precipitation Forecast
TAF	Terminal Aerodrome Forecast
NCEI	National Centers for Environmental Information
NCEP	National Centers for Environmental Prediction
NARR	North American Regional Reanalysis
NOAA	National Oceanic and Atmospheric Administration
ESRL	Earth Systems Research Laboratory
SSC	Spatial Synoptic Classification
MP	Moist Polar
GIS	Geospatial Information System
LWB	Lewisburg
BLF	Bluefield
BCB	Blacksburg
LYH	Lynchburg
ROA	Roanoke
DAN	Danville

Chapter 1. Introduction

Mesoscale meteorology generally refers to weather phenomena that occur on a spatial scale between 10 and 1,000 kilometers. Some of the most impactful weather events are classified as mesoscale, including squall line thunderstorms, straight-line winds, and cold-air damming. Mesoscale meteorology is highly influenced by topographical features that can dictate atmospheric forces that drive airflow and precipitation patterns, introducing complexity that makes forecasting difficult.

Meteorologists rely on conceptual knowledge, forecasting experience, real-time observations, and weather prediction models collectively to produce forecasts. Forecasting mesoscale meteorological phenomena has become more successful in recent decades as a result of the introduction and improvement of mesoscale numerical weather prediction models. Such models, including the North American Mesoscale Forecast System (NAM) and the Rapid Refresh Model (RAP), have fine vertical and horizontal grid resolutions. The increased resolution better represents the details of the atmosphere and the underlying physical geography than do coarser resolution global models like the Global Forecast System Model (GFS) and the European Medium Range Weather Forecast Model (ECMWF); however, despite modern technological advancements, forecasting mesoscale weather events in mountainous regions can prove to be a difficult task. Complex topographical variations influence mountain weather patterns, generating small-scale atmospheric vagaries that weather prediction models do not always accurately represent.

The structure of mountain landscapes is known worldwide to have complex interactions with atmospheric flow, influencing both small and large-scale weather patterns (Stauffer & Warner, 1987; Rackley & Knox, 2016). One example is that of stable surface air circulating around a parent anticyclone and becoming entrenched along the slopes of mountain ranges that act as orographic barriers (Richwien, 1980; Bell & Bosart, 1988). When relatively cold surface air collects against an orographic barrier such as the Appalachian Mountains in the eastern United States, it is termed cold-air damming (CAD), a cold air wedge, or cold dome. This phenomenon is identifiable in sea-level isobar plots by an inverted high-pressure ridge in the shape of a 'U' (Stauffer & Warner, 1987; Bell & Bosart, 1988; Koch, 2001) (Figure 1.1).

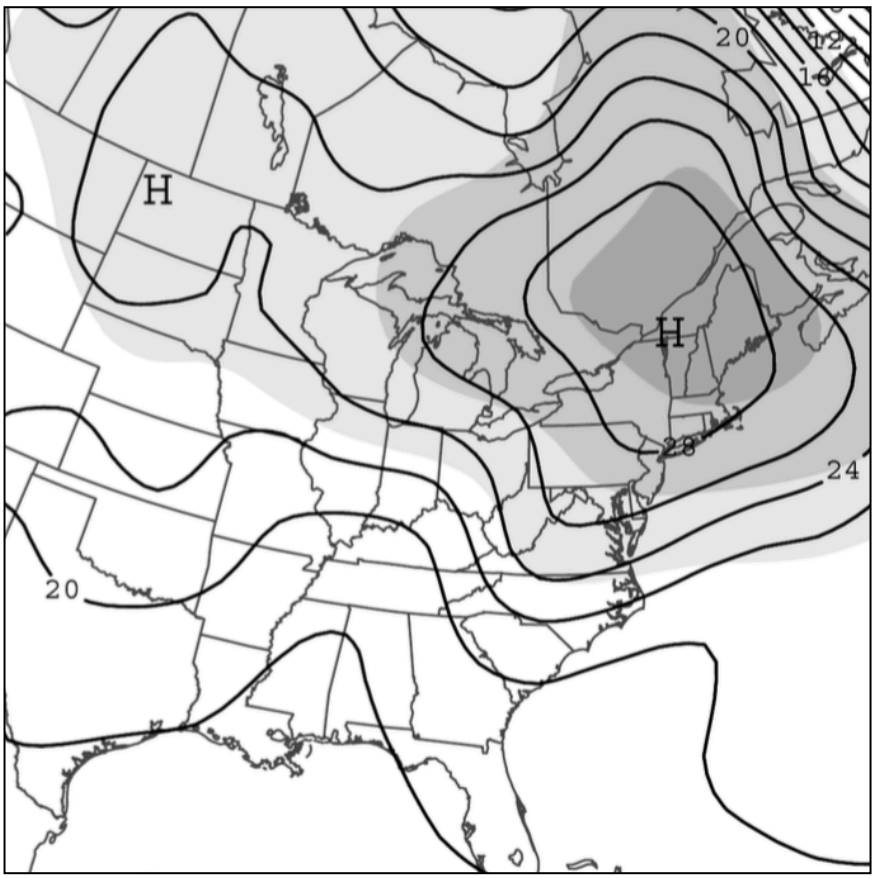


Figure 1.1. Classic cold-air damming wedge example adapted from Bailey et al. (2003).

This pressure 'wedge' results in lower temperatures within the damming region that can surpass a 20°F departure from surrounding areas (Bell & Bosart, 1988; Xu, 1990; Bailey et al., 2003). The pressure wedge is often maintained by cold-air advection at the surface and is strengthened by diabatic processes resulting from evaporative cooling and cloud cover (Bell & Bosart, 1988; Rackley & Knox, 2016). Forecasting precipitation type during CAD is well-known within the meteorological community as a difficult task; the complex topography of the Appalachians, for example, often dictates precipitation type during a damming event that can result in hazardous conditions, especially during winter, stressing the importance of these forecasts (Stauffer & Warner, 1987; Bell & Bosart, 1988; Koch, 2001). Despite modern improvements in high-resolution mesoscale models, the evolution and erosion of these events are not well represented by numerical weather prediction models (Stauffer & Warner, 1987; Bailey et al., 2003; Stanton, 2003; Mahoney, 2006, Rackley & Knox, 2016).

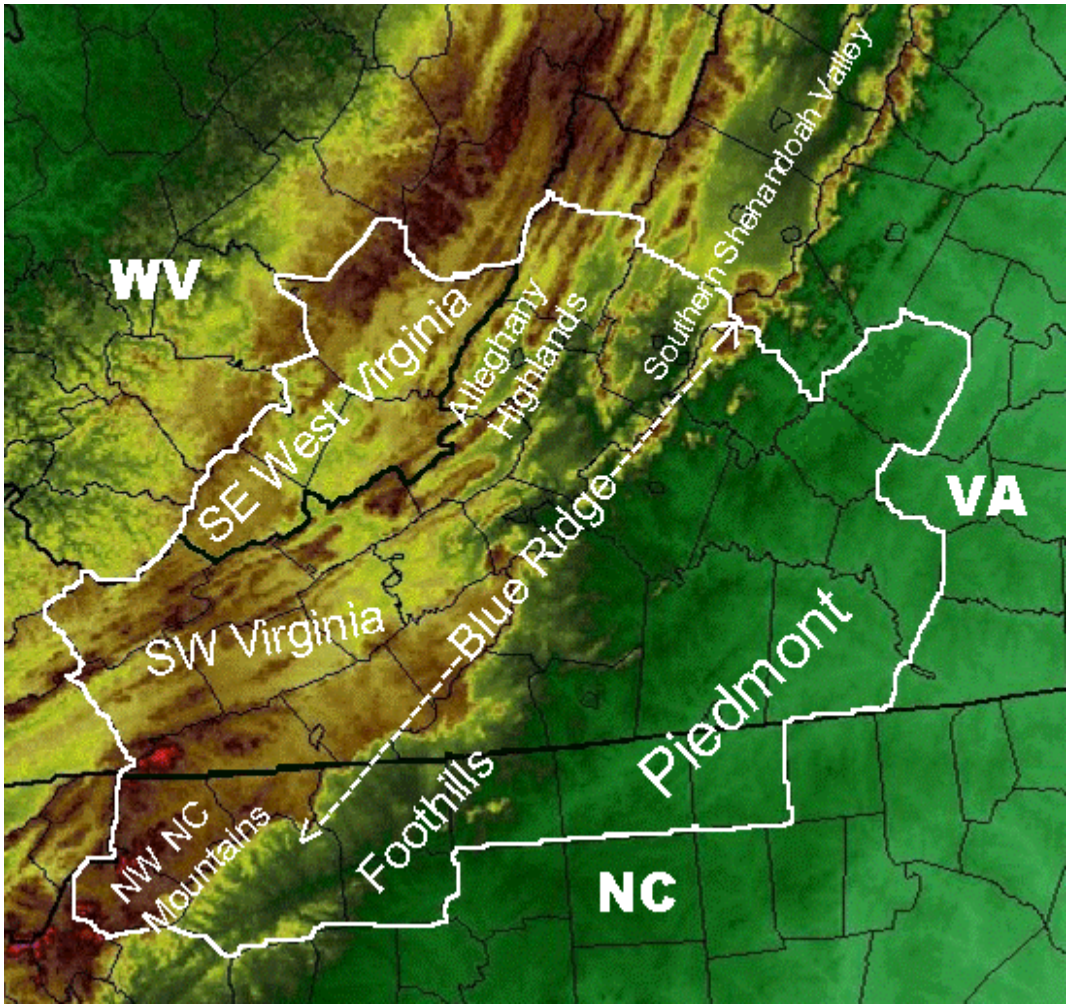


Figure 1.2. Reference map of geographical regions within the Blacksburg NWSFO CWA. Adapted from NOAA – National Weather Service Forecast Office Blacksburg (2017).

A typical Appalachian CAD wedge will settle into the Blue Ridge Mountains from northeast of the Blacksburg National Weather Service Forecast Office (NWSFO) county warning area (CWA) (See Figure 1.2), pooling southwestward against the Appalachian mountains as cold air is advected from the northeast (Figure 1.3).

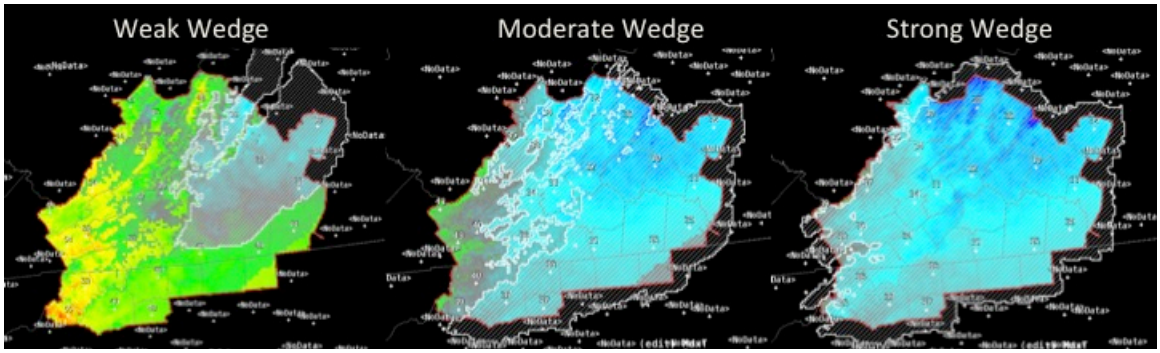


Figure 1.3. Examples of weak, moderate, and strong cold air wedging during central Appalachian cold-air damming episodes; hatched lines represent cold air wedge coverage over the CWA according to wedge strength. Adapted from NOAA (2017).

A weak cold dome, often characterized by light winds and diabatic cooling from precipitation, is wedged into the region along the eastern Appalachian Mountain slopes, settling below the Allegheny Highlands into the Shenandoah Valley and nearby foothills. A moderate wedge, generally driven by both precipitative cooling and cold-air advection, will fill the Piedmont and Blue Ridge Mountain valleys to the Eastern Continental Divide, and the strongest CAD scenarios will engulf the majority of the NWSFO CWA up through its highest elevations (about 2,500 feet for the cold air to overcome the Continental Divide). These strong CAD cases are driven primarily by cold-air advection synoptically forced by a strong parent anticyclone to the northeast. Erosion processes, commonly spurred by an approaching cold front or radiational heating, are usually reversed, in which the wedge dissolves from southwest to northeast. Subsequently, the last places to rid of the cold dome are generally the first to welcome CAD onset.

Relatively low temperatures, increased cloud cover, and varying precipitation intensity and type are characteristic of a typical CAD event. Strong damming cases are capable of producing hazardous conditions, most notably during

the winter, which can bring Appalachian communities to a halt due to low cloud cover, reduced visibility, and frozen precipitation (Forbes et al., 1987). The uncertainty of forecasting the type and amount of precipitation during these events not only increases risk to the general public, but it also pressures emergency managers who rely on this information to mitigate the impacts of snow and ice accumulation (Forbes et al., 1987; Mahoney, 2006). As Forbes et al. (1987) noted within their work, albeit 30 years ago, "a study of extended outages by six utility companies revealed that only one of six storms was predicted sufficiently in advance for crews to be placed on stand-by" (p. 564). Meteorologists erred significantly on 70 percent of these CAD temperature and precipitation forecasts 24 hours in advance (Forbes et al., 1987), and road crews heavily rely on these forecasts for treating roads in preparation for CAD storms. While the work of Forbes et al. was conducted over 30 years ago, many of the forecasting challenges presented by CAD remain today.

1.1 Problem Statement

Research during the latter half of the 20th century yielded an extensive literature focused on CAD; however, the subject has been somewhat absent from the body of literature of the past two decades. Many thorough studies have been conducted for Appalachian CAD, yet most focus on the entire extent of the Southeastern United States, and some even stem as far south as Florida while giving little focused attention to the northern extent of this phenomenon. The Blue Ridge Mountains region of Virginia is noted for persistent CAD episodes, making the

Blacksburg NWSFO CWA (see Figure 1.4) an attractive study domain for CAD research.

Figure 1.4. Blacksburg National Weather Service Forecast Office county warning area.

Additionally, this narrowed spatial domain for CAD has just recently begun to gain interest throughout literature, notably by Ellis et al. (2017). Investigating the accuracy of model-derived air temperature forecasts in this area during CAD events, which has not been previously published, in the form of MAV (GFS MOS Short-range Text Product) and MET (NAM MOS Text Product)-based MOS (Model Output Statistics) will aid forecasters. Furthermore, as Mahoney (2006) explains,

> "If a forecaster recognizes that a model forecast of a specific meteorological feature has been strongly influenced by a given physical parameterization package in a given model, he or she is in a position to better evaluate the uncertainty associated with that aspect of the forecast… this type of knowledge allows

> forecasters to best determine when they are capable of adding value to a model forecast" (p. 466).

An improved understanding of MOS guidance minimum and maximum temperature biases during CAD events in the Blue Ridge Mountains is important because operational meteorologists will be able to add value to their forecasting decisions when interpreting weather model forecasts in the form of MAV and MET output statistics. Forecasters can use this information to produce improved forecasts and, in turn, support public safety during CAD scenarios. In light of the basic and applied research needs outlined here, the proposed thesis research will engage the following research questions:

1. How inaccurate are MAV and MET-derived MOS forecasts for air temperature during central Appalachian Mountain cold-air damming events identified as problematic by forecasters between 2007 and 2016?
2. Have these models exhibited temperature biases in the forecasts, and why might have this occurred?
3. How might forecasters introduce an improved understanding of any MAV and MET guidance temperature bias during cold-air damming scenarios in the central Appalachian Mountains to improve forecasts?

Chapter 2. Literature Review

2.1 Synoptic Drivers of Cold-air Damming

A unique interaction exists between the Appalachian Mountains, the Atlantic Ocean, and atmospheric circulation at both mesoscale and synoptic scales (Keeter et al., 1995). When an anticyclone propagates over the Northeast region of the United States or Eastern Canada, cold and stable air at the surface advects toward the southwest and is dammed along eastern Appalachian Mountain slopes (Stauffer & Warner, 1987; Keeter et al., 1995); ageostrophic northerly flow funnels more air southward into the damming region through cold air advection, thus strengthening the wedge (Stauffer & Warner, 1987) (Figure 2.1).

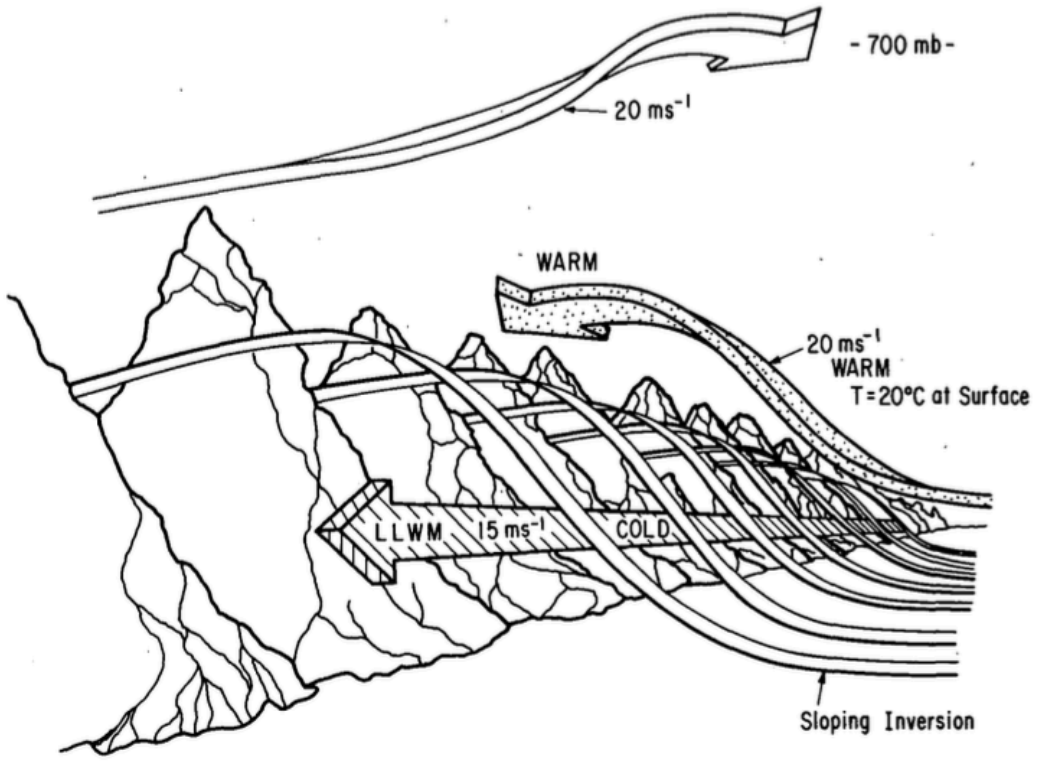


Figure 2.1. Atmospheric flow during Appalachian cold-air damming adapted from Bailey et al. (2003).

A nearly geostrophic balance between Coriolis and pressure gradient force within the cold dome, aided by friction at the surface, directs flow normal to the Appalachians (Xu, 1990) and deepens the spatial extent of the damming region (Bell & Bosart, 1988). The highest pressure within a wedge results where surface air is at its coolest and deepest, wherein lies a correlation to the event's inversion heights (Stauffer & Warner, 1987). Overrunning warm air aloft reinforces a subsidence inversion that remains for the entirety of the event's lifecycle (Bell & Bosart, 1988), which is fortified by evaporative cooling and adiabatic cooling that causes air alongside the mountain barrier to cool and descend (Stauffer & Warner, 1987; Koch, 2001). The cold surface conditions relative to outside the wedge are enhanced by the blocking of insolation from clouds forming within the upper portion of the inversion (Stanton, 2003).

2.2 Climatology of Cold-air Damming

CAD is a prevalent year-round mesoscale phenomenon that occurs east of the Appalachians on average for over 50 days a year (Rackley & Knox, 2016) with roughly 3 to 5 strong damming scenarios each cold season (Bell & Bosart, 1988). As Rackley and Knox (2016, p. 431) state, "fifty days per year is nontrivial for a phenomenon with the potential to significantly impact forecasts of sensible weather." Stauffer and Warner (1987) found damming scenarios to last typically around 30 hours, while Bell and Bosart (1988) found weaker events to last about a day. Bailey et al. (2003) identified these CAD cases using a detection algorithm based on wedge characteristics, while Ellis et al. (2017) more recently used the

spatial distribution of synoptic weather types over central Appalachia to identify days in which CAD was present historically.

An event is considered strong if it persists for over 36 hours within the record of surface observations (Stanton, 2003). Although the phenomenon occurs year-round, CAD is strongest and most prevalent in the winter months (Stauffer & Warner, 1987; Bell & Bosart, 1988; Rackley & Knox, 2016). Strong spring and summer CAD events can also occur, potentially providing a trigger for convection along the wedge boundary during erosion (Rackley & Knox, 2016).

2.3 Classifications of Cold-air Damming

CAD varies by size, intensity, and duration, and can be classified into three broad categories based on formation mechanisms: classic, hybrid, and in-situ (Bailey et al., 2003). Although Bailey et al. (2003) developed an expansive classification scheme, this research uses a simplified spectrum of damming types that is more intuitive to operational forecasters. The primary factors in classification are the nature of the parent high-pressure system as a reflection of synoptic forcing and diabatic processes in terms of precipitation enhancement (Bailey et al., 2003; Stanton, 2003). The most common, and typically strongest, type of situation is classic CAD, in which the central high-pressure center is positioned over the Northeast or eastern Canada, with a pressure of greater than or equal to 1030 millibars (Bailey et al., 2003). These events must exceed 24 hours to rule-out the typical progression of a surface high, and although precipitation can intensify the onset of a classical wedge, most are not enhanced diabatically (Bailey et al., 2003). In contrast, hybrid CAD involves a parent high to the northeast that is weaker than

1030 millibars and is aided by diabatic processes to strengthen the cold dome (Bailey et al., 2003). Lastly, in-situ events are formed from a more easterly high-pressure center toward the Atlantic that induces lesser synoptic forcing (Bailey et al., 2003). These situations often result from a passing cold air mass that leaves behind cold conditions (Bailey et al., 2003), and are dependent on evaporative cooling and cloud processes over their duration (Bailey et al., 2003; Koch, 2001).

2.4 Sensible Weather Associated with Cold-air Damming

Strong CAD events are commonly associated with below normal temperatures, increased cloud cover, and precipitation. In the climatology generated by Bailey et al. (2003), roughly 4/5 of all classic events dramatically impacted sensible weather. Evaporative cooling, fueled by clouds and precipitation, is significant in terms of strengthening a CAD event. Stability of the surface layer, wedge strength, and mountain blocking all increase with evaporative processes (Stanton, 2003), which can contribute to nearly one-third of the cooling that occurs within the cold dome (Bell & Bosart, 1988). Evaporative cooling plays the largest role in dome strengthening during the onset of an event, typically waning as low-level air becomes saturated (Stauffer & Warner, 1987); however, nearly saturated air that is orographically blocked descends and warms adiabatically, allowing evaporative processes to continue that aid in the persistence of the cold wedge (Bell & Bosart, 1988). Relative surface temperature differences inside and outside of the damming region can heavily influence precipitation type (Stanton, 2003), and a sub-freezing layer at the surface enhanced by evaporative cooling can convert liquid precipitation into freezing rain. CAD situations can sometimes result in multiple

transitions between frozen and liquid precipitation (Mahoney, 2006), and the uncertainty in forecasting these types of events points to a need for further research.

2.5 Cold-air Damming Erosion

The erosion of CAD is not well understood and is often premature within model-driven forecasts. According to a study by Stanton (2003), the primary mechanisms that can contribute to erosion include differential thermal advection, solar heating, lower-troposphere divergence, shear-induced mixing, and the advancement of a coastal warm front on the eastern cold dome edge. Stanton (2003) found that CAD onset mechanisms are independent of erosion type. After developing synoptic composites, Stanton (2003) also determined coastal cyclones, cold-front passages, residual cold pools affected by solar heating, and a northwestern low that produces a cold front which spurs shear-induced mixing at the top of the wedge boundary are the primary synoptic patterns under which erosion occurs. No additional studies have focused solely on the drivers that end CAD events, and although it is well known that numerical weather prediction poorly handles erosion, the accuracy of forecast model predictions during the demise of CAD events has not been explicitly examined.

2.6 Vertical Soundings in Cold-air Damming Analyses

Vertical atmospheric soundings are an invaluable tool meteorologists use during daily forecasting; however, most CAD research lacks extensive sounding analyses when investigating the nature of this event, primarily due to the sparse

distribution of observation sites throughout the country. As Keeter et al. (1995) describe,

> "The temporal and spatial limitations of the current upper-air observational network, consisting of soundings taken only at 12-h intervals, at stations scattered hundreds of kilometers apart, impacts the accuracy and resolution to which events such as cold-air damming can be depicted by dynamical models" (p. 46).

Furthermore, small-scale and weak events often progress undetected by sounding data resulting from the network's spatial and temporal limitations (Bailey et al., 2003). Investigating CAD within the Blacksburg NWSFO CWA offers abundant upper air data centered within the study area, highlighting these soundings as a valuable tool for studying the vertical atmospheric profiles of CAD scenarios.

2.7 Forecasting Cold-air Damming

Over the past half-century, the ability of forecasters to predict the weather has improved significantly, yet CAD remains one of the most difficult forecasting challenges east of the Appalachians (Koch, 2001; Bosart, 2003; Rackley & Knox, 2016). The shallow nature of the wedge, typically capped by an inversion at around the 850 millibar pressure level of the atmosphere, vastly limits the tools available to forecasters during these events (Bell & Bosart, 1988; Koch, 2001). Surface observations, though critical during the evolution of a damming scenario, provide spatially-limited environmental information to interpret, forcing meteorologists to rely on their conceptual knowledge of CAD to resolve the progression of the cold dome (Stauffer & Warner, 1987; Koch, 2001). Predicting precipitation type can be one of the most difficult tasks of forecasters during CAD, as evaporative cooling can

cause quick changes in hydrometer type that are difficult to detect (Forbes et al., 1987; Keeter et al., 1995; Bailey et al., 2003). Additionally, quantitative precipitation forecasting (QPF) and maximum temperature forecasting are also challenging to resolve as a wedge persists, even when precipitation has little to no influence (Keeter et al., 1995; Mahoney, 2006). In winter months, both precipitation type and accumulation forecasts are critical, as they impact how emergency managers prepare for these events (Forbes et al., 1987). On the other hand, during the warm season, CAD wedges can initiate convective events along the boundary of cold air, resulting in difficult-to-predict severe weather (Rackley & Knox, 2016). While operational meteorologists are becoming increasingly dependent on weather models (Bosart, 2003), the weakness of models in forecasting the characteristics of CAD still presents a hurdle for forecasters (Mahoney, 2006).

2.8 Numerical Weather Prediction of Cold-air Damming

Mesoscale weather models prove to be more useful than global models at resolving the spatial and temporal nature of CAD (Keeter et al., 1995) due to their ability to generate higher resolution solutions (Koch, 2001); however, mesoscale models still tend to underestimate wedge duration and impacts despite modern computing power (Rackley & Knox, 2016) and advancements in both horizontal and vertical resolutions (Bell & Bosart, 1988; Koch, 2001). Small-scale and weak CAD setups can often progress undetected even by mesoscale weather models (Bailey et al., 2003). The tendency of models to prematurely erode the cold dome often results in a warm bias, whereas over-predicted precipitation during onset can result in a cold bias (Bailey et al., 2003). Diabatically-enhanced events are not accurately

handled when onset is not primarily driven by precipitation (Bailey et al., 2003), and the parameterization of heavy precipitation and organized convection are common sources of forecast error for weather prediction models (Bailey et al., 2003; Mahoney, 2006). Boundary layer dynamics, wind forecasts, and surface pressure all exhibit substantial model error within the wedge, and increased vertical resolution is necessary to better resolve these parameters (Stauffer & Warner, 1987). Forecast uncertainty increases when numerical weather prediction models show disagreement through different solutions for events, which commonly occurs during CAD.

Parameterizing the erosion of CAD is typically viewed as the most difficult aspect of forecasting this phenomenon. Models are unreliable during the demise of CAD events and tend to prematurely erode the wedge (Koch, 2001; Stanton, 2003). Smoothing of terrain and inadequate vertical resolution contribute to poor modeling that causes failures during erosion forecasts (Stauffer & Warner, 1987; Xu, 1990; Koch, 2001). Models also exhibit an overabundance of surface heating due to problematic interpretation of interactions between solar radiation and shallow cloud cover, and when combined with premature advancement of cold air advection aloft, low temperatures, precipitation, and cloud cover are often prematurely terminated (Stanton, 2003).

2.9 Model Output Statistics Forecasts of Cold-air Damming

According to Glahn & Lowry (1972),

> "Model Output Statistics is an objective weather forecasting technique which consisists of determining a statistical relationship between a predictand and variables forecast by a numerical model… The MOS

> technique is, in effect, the determination of 'weather-related' statistics of a numerical model" (2018).

Model Output Statistics (MOS) represent the location-specific numerical output from raw numerical forecast model output including the GFS and NAM (Glahn & Lowry, 1972). MOS is driven by statistical-dynamical multiple linear regression analysis (Glahn & Lowry, 1972; NOAA – National Weather Service, 2018) that uses variables from raw model output including temperature at the surface, saturation defecit, and 500 mb height (Glahn & Lowry, 1972) to yield predictions of minimum and maximum temperatures every 12 hours in the form of statistical guidance. MOS also estimates surface parameters like 3-hourly temperatures, quantitative precipitation forecasts, cloud cover, and many more (Glahn & Lowry, 1972; Bell & Bosart, 1988; Carter et al., 1989). Each site-specific forecast has a unique equation to determine MOS minimum and maximum forecast guidance in order to reduce bias across sites (Glahn & Lowry, 1972). Its diagnostic equations are modified seasonally based on post-accuracy assessment of all previous seasons (Carter et al., 1989; Keeter et al., 1995), resulting in MOS temperature and precipitation forecasts that rival those of experienced meteorologists (Bosart, 2003). However, numerical prediction model accuracy directly affects the success of MOS guidance (Carter et al. 1989). While MOS generally corrects for model biases, an accurate assessment of sensible weather variables from parent models is necessary for MOS to produce successful statistical guidance (Carter et al. 1989).

During CAD situations, MOS temperature forecasts commonly initially predict temperatures that are too high within the damming region and too low toward the coast outside of the wedge (Forbes et al., 1987; Bell & Bosart, 1988).

Historically, such systematic errors are often overcorrected in the next forecast cycle and MOS predicts conditions much too cold within the wedge the next day (Forbes et al., 1987). Insufficient modeling of vertical resolution and physical processes by parent models render MOS guidance of little value during cold dome forecasting (Forbes et al., 1987; Bell & Bosart, 1988), whereas this tool stands out in modern forecasts of most other weather situations as being highly valuable.

2.10 Summary

Appalachian CAD is a synoptically driven phenomenon affected by local-scale physical processes that contribute to the complexity of this weather event. When wedged into the Blue Ridge, CAD can cause societal issues, particularly in winter months when emergency management depends on CAD forecasts to support public safety. CAD is notoriously problematic for forecasters, especially given the difficulty with which models represent the cold dome and associated characteristics. No study has been conducted whereby model output during poorly forecast CAD events has been analyzed, and this research seeks to assess potential reasons as to why biases in the GFS-based MOS (MAV) and NAM-based MOS (MET) output may skew operational meteorologists' forecasts during particularly problematic instances of this weather event.

Chapter 3. Data and Methods

3.1 National Weather Service Cold-air Damming Bust Database (2007-2016)

Since 2007, forecasters at the NWSFO in Blacksburg, Virginia have archived CAD events that resulted in what they consider to be 'busted' (in forecaster vernacular) maximum or minimum temperature forecasts – those consisting of errors of 8°F (4.4°C) or higher – in an internal database in attempts to capture synoptically driven cases of this phenomenon (Robert Stonefield – NWSFO Blacksburg, Personal communication, September 2017). Each archived event spans 12 hours, capturing observed minimum temperatures between 00Z and 12Z or maximum temperatures occurring between 12Z and 00Z, depending on the forecast time period. Each archived event includes the date and location(s) of the erroneous forecast from among the six terminal aerodrome forecast (TAF) sites within the CWA for which MOS guidance data are available (Blacksburg, VA; Roanoke, VA; Lynchburg, VA; Danville, VA; Bluefield, WV; and Lewisburg, WV) (See Figure 3.1).

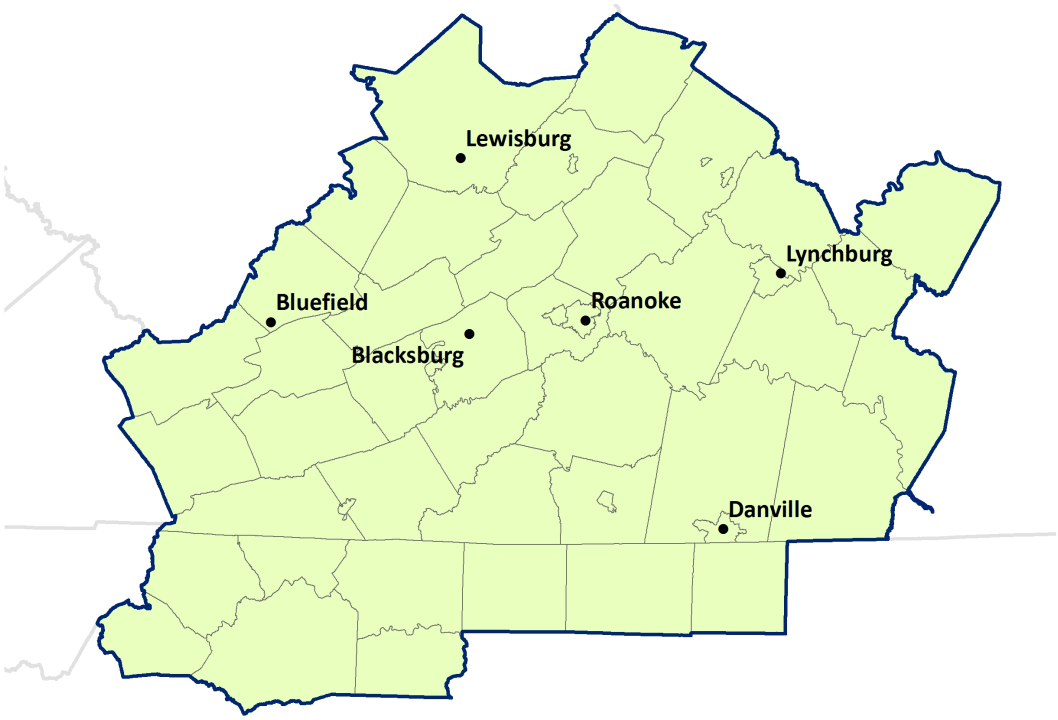


Figure 3.1. The Blacksburg NWSFO county warning area (CWA; green shading) and the six TAF sites for which temperature forecasts are analyzed.

In each instance, the official NWS forecast was at least 8°F (4.4°C) warmer or cooler than the observed temperature, signaling that these CAD cases were particularly challenging. It is the case that MOS forecasts may have proven to be more accurate than NWS forecasts, but it is more typical that they are less accurate. For each site at which the temperature forecast 'busted' during the 12-hour forecast period, the database contains numerical forecasts from both the NAM-MOS (MET) and GFS-MOS (MAV) and Blacksburg NWS Office official forecasts 12 to 60 hours in advance, stratified into five 12-hour forecast cycles. Area forecast discussions and forecaster comments are also included in the database for each event. The database contains 225 cases spanning 10 years between January 2007 and December 2016, segregated by the type of damming event (classic, hybrid, and in-situ) and/or whether the error

occurred at the onset or erosion of the event. There are likely many more problematic CAD events that impacted the central Appalachian Mountains than recorded in this database; these archived cases were particularly difficult for forecasters, but there are many more CAD events that occurred during this time period that were either well-forecasted or were below the 8°F bust threshold. An emphasis on forecasting these events using MOS guidance makes this official NWS bust database an ideal starting point for this research.

The CAD busted forecast database provided by the NWS offers an array of potential events to analyze that proved difficult for forecasters to accurately predict. Each event has been visually verified by this investigator and re-classified using the Weather Prediction Center's Surface Analysis Archive (http://wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive.php). Criteria for verifying each case included a clearly defined pressure 'wedge', easterly or northeasterly surface winds over the Blue Ridge region of Virginia, relatively low air and dew point temperatures compared to surrounding areas, and cloud cover or precipitation over the CWA. The strongest events were those re-classified as classic, exhibiting a parent high pressure anticyclone of greater than 1030 millibars. Hybrid CAD cases exhibited similar characteristics but with central pressures less than 1030 millibars for the anticyclone generally positioned slightly to the south over the Northeast United States. In-situ cases were less well-defined in terms of surface pressure wedging and showed evidence of diabatic enhancement. These cases were typically situated more easterly towards the Mid-Atlantic coast with more southeasterly winds. A lack of a clearly-defined wedge or the displacement of

synoptic high pressure passing directly over the area rather than to the north as with a typical wedge event caused 110 erroneous forecast events to be discarded from the final database. Many of these discarded events that appeared to be classic or hybrid were dismissed as "backdoor fronts" – cold fronts passing through the area from the north/northeast – that gave the false appearance of a CAD dome. The 5 archived cases with missing spatial and numerical forecast data were also discarded from this study.

Once this investigator visually verified these events, forecasters at the National Weather Service also visually verified the cases chosen from the bust database to reduce the subjectivity of these scenarios. Using their expertise, the dataset was further narrowed to 110 events to use for the final bias and composite analyses.

Lastly, forecaster comments in the NWSFO Blacksburg CAD bust database were reviewed and are discussed to set the tone for the underlying complexity of these cases.

3.2 MOS Guidance Archived Data Retrieval

Of the 110 cases of CAD characterized by poor temperature forecasts, 13 CAD busts were missing MAV and MET guidance forecasts from the official NWS bust database across all forecast cycles. These missing data were obtained using the Iowa State University Environmental Mesonet NWS MOS Download Interface (https://mesonet.agron.iastate.edu/mos/fe.phtml) and were integrated into the bust database when performing bias calculations to ensure no cases were missing any forecast data.

3.3 Abbreviated Climatology Compilation

Using the final list of busted CAD events, an abbreviated (10-year) climatology was compiled for the cases. These data are characterized in terms of frequency per year and month, maximum vs. minimum temperature forecasts, and spatial location based on the TAF sites used in the bust database.

3.4 Assessing MAV and MET Guidance Bias

Model bias statistics were calculated for each classification using Fortran programming. For problematic maximum and minimum air temperature forecasts, mean error, absolute mean error, and bias (forecast temperature divided by observed temperature) were computed to characterize the nature of the error in MAV and MET temperature forecasts. Box and whisker plots generated in Microsoft Excel are used to graphically represent any biases for each model within the CAD classifications. A two-sampe t-test using data from MET and MAV forecasts throughout all busted CAD cases was run for both maximum and minimum temperature forecasts to test whether one model produced significantly more accurate forecasts than the other. The above statistics were computed for each of the six TAF sites to account for the spatial distribution of forecast errors, as long as there were enough data points at each TAF site to produce meaningful statistics.

3.5 Assessing Model Biases Over Time

Annual boxplots of both minimum and maximum temperature MAV and MET biases were plotted through time in Microsoft Excel to examine how MAV and MET guidance biases changed through the 10-year period. The purpose is to ascertain whether seasonally updated MOS guidance parameters had a systematic effect on

model improvement through time. Annual median forecast error for MAV and MET was plotted in Excel to assess guidance error covariance. This was done for each of the five maximum temperature forecast cycles (-12 hours to -60 hours). Any significant differences between MAV and MET forecast error were cross-referenced with NCEP Central Operations' Changes to NCEP Models/Implementation Dates to NOAAPORT Database (http://www.nco.ncep.noaa.gov/pmb/changes/), which provides a record of MAV and MET guidance equation updates.

3.6 Hourly Surface Composites

Hourly meteorological surface variables including air temperature, relative humidity, wind direction, and cloud ceiling heights spanning between 0Z and 23Z for each of the problematic CAD forecast dates at all six TAF sites were gathered from the National Centers for Environmental Information (NCEI) Climate Data Online (https://www7.ncdc.noaa.gov/CDO/cdoselect.cmd). These data were then used to generate hourly composites of atmospheric variables during busted CAD forecast events at each TAF site. Further, daily data were downloaded for all days over the 10-year study period for the Blacksburg TAF site to construct period means of hourly data for the cold season (November to April) and warm season (May to October) at Blacksburg. The composites were computed using Fortran and plotted in Microsoft Excel.

3.7 Synoptic Composites

High-resolution (32km, or 0.3 degrees latitude/longitude) synoptic atmospheric composites were generated from the dates of the erroneous forecasts within each CAD classification to characterize the dominant large-scale atmospheric

circulation associated with these events. Composites segregated by CAD classification were created from the Daily Average NCEP NARR Composites (https://www.esrl.noaa.gov/psd/cgi-bin/data/narr/plotday.pl) available through NOAA's Earth System Research Laboratory. Each composite represents mean synoptic atmospheric conditions on days within the NWS database of busted forecasts of CAD. The composites display sea level pressure and geopotential heights at 925 hectopascals (hPa), 850 hPa, and 500 hPa; air temperature at the surface, 925 hPa, 850 hPa, and 500 hPa; vector winds at 925 hPa, 850 hPa, and 500 hPa; and 2-meter dew points and relative humidity values. The spatial domain of the composites include the eastern the extent of the United States centered above the central Appalachian study region. The synoptic composites for busted events are compared to synoptic composites available within the published work of Bell and Bosart (1988), Bailey et al. (2003), and Ellis et al. (2017) to identify any major differences thay might lend to forecasting difficulties. Similarly, composites displaying surface conditions one to three days in advance of each CAD classification were also created, but they are not included within this document.

Additional surface composites of coarser resolution (2.5 degrees latitude/longitude) showing pressure, air temperature, and vector wind differences of each classification based on classic CAD scenarios were provided by the NOAA/ESRL Physical Sciences Division, Boulder Colorado from their Daily Mean Composites page (https://www.esrl.noaa.gov/psd/data/composites/day/). The spatial domain of these composites is consistent with the NARR daily composites.

3.8 Upper Air Sounding Composites

Vertical atmospheric profile data for each of the CAD forecast periods examined in this study were retrieved from the University of Wyoming's Soundings Archive (http://weather.uwyo.edu/upperair/sounding.html) for the Upper Air station located in Blacksburg, VA (KRNK). These data were used to create a simple mean composite from the problematic forecast periods for each CAD classification in both the warm and cold season using Fortran programming and Microsoft Excel. Atmospheric sounding composites characterize the mean condition of the atmosphere through a vertical cross-section, providing insight into the vertical profile of the atmosphere associated with busted forecasts.

3.9 Classic Cold-air Damming Climatology

Ellis et al. (2017) identified the presence of CAD events in the central Appalachian Mountains using a spatial synoptic classification (SSC) approach. Identifying days in which moist polar (MP) air was in place to the east of the mountains and non-MP air to the west produced a cold-season climatology of 219 classic CAD events in the region over a 35-year period. Presumably, CAD days within the database of Ellis et al. that do not appear within the NWS database of problematic CAD days were well-forecast wedge events. A comparison of vertical atmospheric composites from the two databases may help to reveal sources of forecast difficulty associated with the busted CAD forecasts.

Chapter 4. Results and Discussion

4.1 Review of 'Bust' Database Forecaster Comments

As forecasters at the Blacksburg NWSFO developed the archive of busted CAD forecasts, they often added personal comments and included reasons as to what may have caused each individual forecast to bust. These reasons generally include sky cover, precipitation, and low-level mixing, among others. For the problematic forecasts archived in the bust database, nearly all of them (72 of 110) were further complicated by sky cover and a large number (46 of 110) by precipitation. Only 10 of the 110 cases had the sole reason that a CAD setup impacted forecast error. These challenging scenarios alone are difficult to accurately forecast, and the high frequency of complicating factors imply these cases may be more complex than typical CAD scenarios for the region.

4.2 Climatology of Cold-air Damming 'Busts'

Due to the limited temporal range of the NWSFO-Blacksburg's database of problematic CAD forecasts, an abbreviated (10-year) descriptive climatology of CAD forecast busts is constructed. A total of 110 'busted,' or unsuccessfully forecast, CAD events are stratified into the following classifications based on synoptic characteristics: 15 classic, 18 hybrid, 12 in-situ, 18 onset, and 47 erosion (Table 4.1). Erroneous maximum air temperatures forecast at 0Z UTC (7 or 8 pm local) for the subsequent 12Z to 0Z period are referenced here as "Hi," as temperature forecasts within this period typically reflect daily maximum values. Erroneous minimum air temperature forecasts at 12Z (7 or 8 am local) for the subsequent 0Z to 12Z period

in which temperature forecasts in this period typically reflect daily minimum values are referenced as "Lo."

Table 4.1. Occurrences of problematic cold-air damming forecasts by classification type. Dates are listed as month/day/year (last 2 digits). The erroneous temperature forecast is indicated parenthetically as either maximum (Hi) or minimum (Lo) daily temperature.

Classifications				
Classic	**Hybrid**	**In-situ**	**Onset**	**Erosion**
2/28/15 (Hi)	10/6/16 (Hi)	6/29/14 (Hi)	8/8/16 (Hi)	12/12/16 (Hi)
12/8/14 (Hi)	8/4/16 (Hi)	7/31/13 (Hi)	2/22/16 (Hi)	3/12/16 (Hi)
3/18/13 (Hi)	4/30/16 (Hi)	4/15/13 (Hi)	3/24/15 (Hi)	2/24/16 (Hi)
9/18/11 (Hi)	2/2/16 (Hi)	7/16/11 (Hi)	3/23/15 (Hi)	12/29/15 (Hi)
4/22/11 (Hi)	6/4/15 (Hi)	3/15/11 (Hi)	10/11/14 (Hi)	12/1/15 (Hi)
11/29/10 (Hi)	4/19/15 (Hi)	8/2/10 (Hi)	9/23/14 (Hi)	3/25/15 (Hi)
1/27/09 (Hi)	9/8/14 (Hi)	5/11/10 (Hi)	3/16/14 (Hi)	3/20/15 (Hi)
4/6/08 (Hi)	5/5/13 (Hi)	4/13/09 (Hi)	4/4/13 (Hi)	3/3/15 (Hi)
3/30/08 (Hi)	5/4/13 (Hi)	12/28/07 (Hi)	9/6/11 (Hi)	12/24/14 (Hi)
12/21/07 (Hi)	11/6/12 (Hi)	2/25/07 (Hi)	3/14/11 (Hi)	8/10/14 (Hi)
4/23/13 (Lo)	4/18/12 (Hi)	5/11/10 (Lo)	2/22/11 (Hi)	5/7/14 (Hi)
11/19/12 (Lo)	10/12/09 (Hi)	4/11/07 (Lo)	9/26/10 (Hi)	4/30/14 (Hi)
11/18/12 (Lo)	9/11/10 (Hi)		7/31/10 (Hi)	4/7/14 (Hi)
9/17/11 (Lo)	8/14/10 (Hi)		3/19/15 (Lo)	3/19/14 (Hi)
11/19/09 (Lo)	8/1/10 (Hi)		12/21/14 (Lo)	2/23/13 (Hi)
	4/10/15 (Lo)		7/9/09 (Lo)	12/20/12 (Hi)
	10/13/11 (Lo)		11/29/08 (Lo)	7/13/12 (Hi)
	8/22/08 (Lo)		3/21/07 (Lo)	1/23/12 (Hi)
	8/21/08 (Lo)			12/27/11 (Hi)
				5/13/10 (Hi)
				4/14/10 (Hi)
				1/24/10 (Hi)
				12/13/09 (Hi)
				12/9/09 (Hi)
				11/19/09 (Hi)
				3/10/09 (Hi)
				1/28/09 (Hi)
				12/9/08 (Hi)
				5/11/08 (Hi)
				3/31/08 (Hi)
				2/17/08 (Hi)
				2/12/08 (Hi)
				2/1/08 (Hi)
				12/23/07 (Hi)
				4/11/07 (Hi)
				3/21/07 (Hi)
				3/28/16 (Lo)
				11/10/15 (Lo)
				3/19/13 (Lo)
				1/12/13 (Lo)
				4/23/11 (Lo)
				5/12/10 (Lo)
				3/11/09 (Lo)

				4/4/08 (Lo) 2/13/08 (Lo) 10/27/07 (Lo) 4/12/07 (Lo)

The CAD forecast busts are rather evenly distributed across the classifications, with the exception of erred forecasts associated with CAD erosion. The scenario of in-situ accounts for the lowest proportion at 11%, whereas erosion makes up the most at 43%. The proportionally high number of erred forecasts during CAD erosion is consistent with the convention that erosion processes present the largest forecasting challenge during this phenomenon, as erosion mechanisms are not fully understood and models traditionally end the wedge prematurely (Stanton, 2003).

4.2.1 Annual Distribution

The 110 busted CAD forecasts are fairly evenly distributed across the 10 years of the database (2007-2016) (Figure 4.1); however, the proportions of each classification type are highly variable from year-to-year (Figure 4.2).

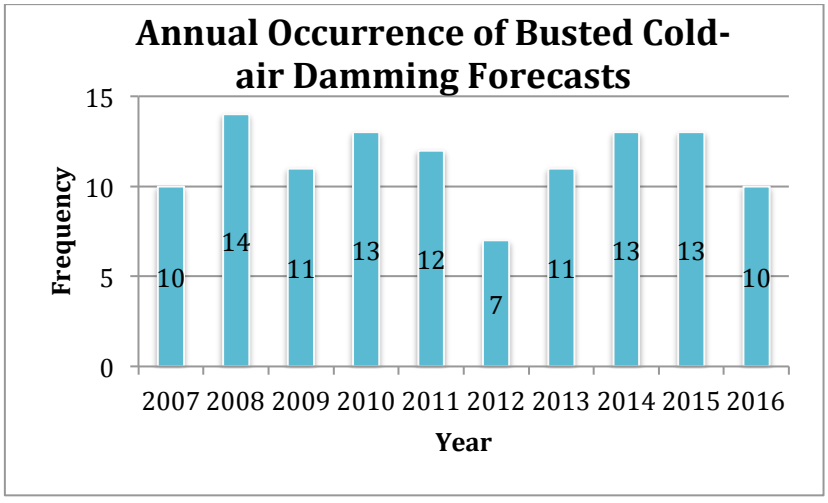


Figure 4.1. The annual occurrence of busted cold-air damming forecasts, 2007 to 2016.

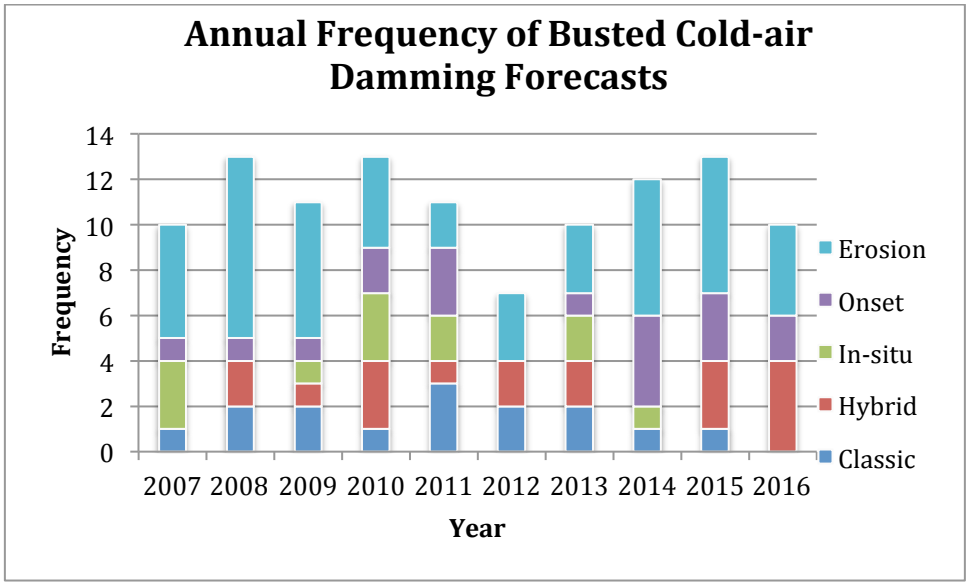


Figure 4.2. The annual frequency of busted cold-air damming forecasts stratified by classification type, 2007 to 2016.

Forecast busts associated with CAD erosion events were frequent in 2008 and 2009, appear much less frequently thereafter, and increase again in 2014 and 2015. This likely speaks to the frequency with which a particular type of difficult-to-forecast form of CAD erosion occurred in these years. Additionally, forecast busts associated with CAD onset became increasingly prevalent toward the end of the study period with a peak in 2014. Overall, there does not appear to be a significant trend in the distribution of busted CAD forecasts by classification over the 10-year study period. This suggests that there has not been a systematic change in the forecast difficulties presented by the different classes of CAD and its onset and demise.

4.2.2 Monthly Distribution

The distribution of busted CAD forecasts by month suggests that March CAD events present a particular forecast challenge (Figure 4.3); however, as CAD climatologically occurs more often during the month of March (Bell & Bosart, 1988),

this could simply be a reflection of the frequency with which CAD occurs. June is the month with the fewest busted CAD forecasts over the 10-year period of the database (Figure 6.1.1), but June is also characterized by relatively few CAD events climatologically (Bell & Bosart, 1988).

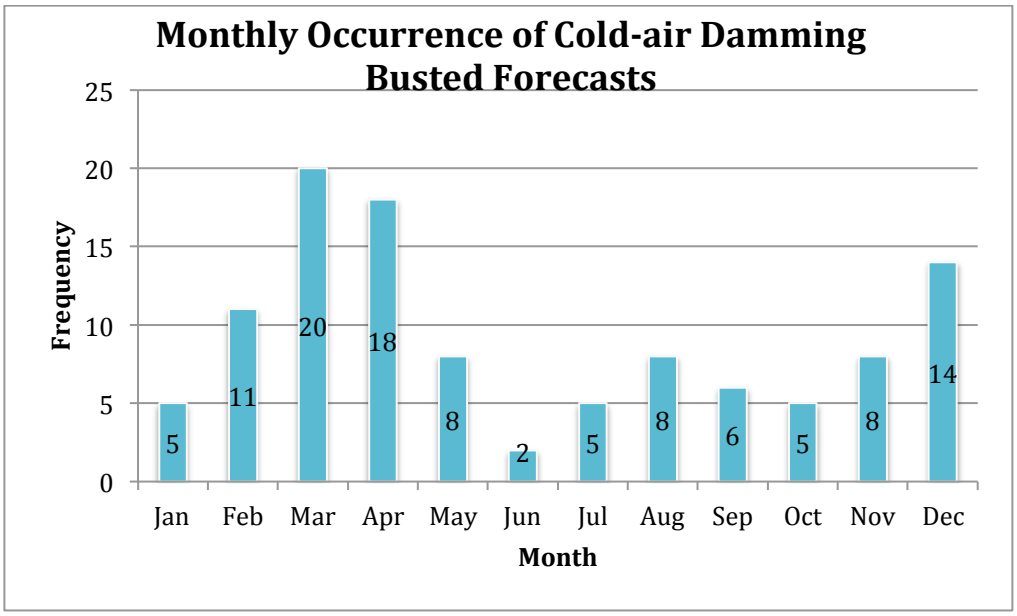


Figure 4.3. The frequency of busted cold-air damming forecasts by month, 2007 to 2016.

The distribution of busted CAD forecasts segregated by classification is highly variable by month (Figure 4.4), much like that for the annual distribution (Figure 4.1). Busted CAD forecasts associated with erosion events occur mostly in March and December, while those for the other CAD classifications are fairly consistent throughout the cold season and dwindle during summer.

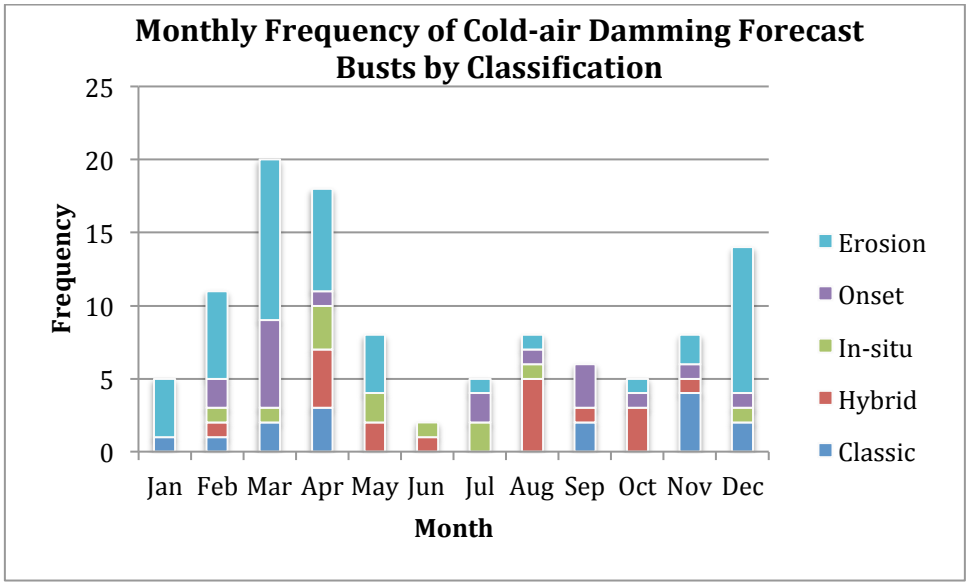


Figure 4.4. The monthly frequency of busted cold-air damming forecasts stratified by type, 2007 to 2016.

The only trend in monthly occurrences of busted CAD forecasts segregated by classification stems from erosion cases that occur predominantly in the cold season (November to April), which may be a product of enhanced warm air advection and mixing over the shallow surface layer of cold, stable air. The majority of all problematic CAD cases are found in the cold season (November to April), rather than the warm season (May to October); this may be attributed to stronger seasonal pressure and temperature gradients that amplify potential errors in model and forecaster judgement, or simply that CAD occurs more often in winter months. Warm season CAD busts account for 37 of the 110 cases, while 73 occurred in the cold season.

4.2.3 Maximum vs. Minimum Temperatures

A majority of the busted CAD forecasts occurred during daytime hours (maximum temperature), or between 12Z and 0Z, accounting for 83 of the 110 cases

(75%). Far fewer (27) occurred at night (minimum temperature), or between 0Z and 12Z UTC (Figure 4.5).

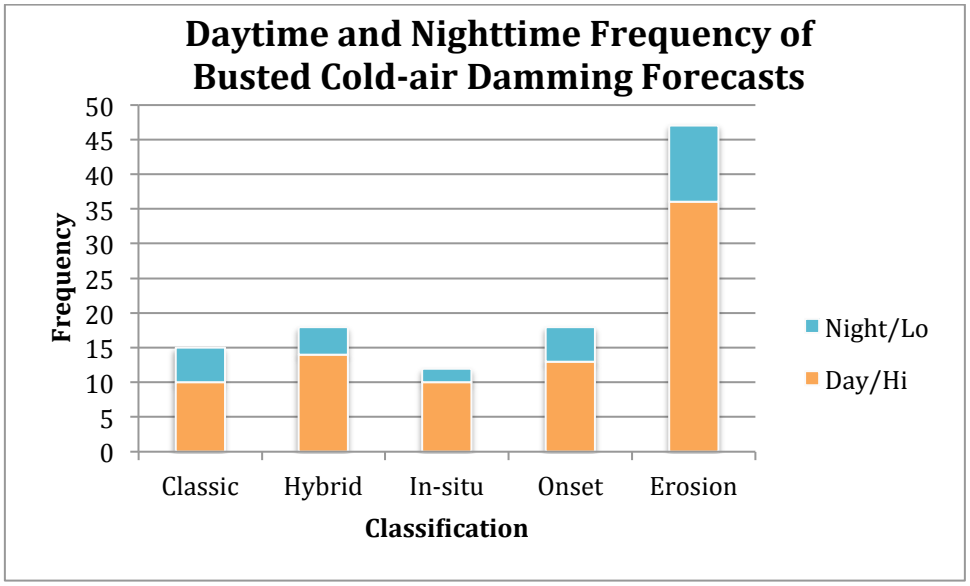


Figure 4.5. The daytime and nighttime frequency of busted cold-air damming forecasts stratified by classification, 2007 to 2016.

Three out of every four busted CAD forecasts were daytime high temperature forecasts, and this is consistent across all CAD classifications. This suggests that daily maximum air temperature presents a consistently greater forecast challenge. Frequent maximum temperature errors during CAD erosion may be enhanced during daytime erosion if a wedge scenario erodes but is forecast to persist, causing temperatures to rise rapidly and amplifying forecast error.

4.2.4 Spatial Distribution

A majority of the busted CAD forecasts were for only one of the six TAF sites in the Blacksburg CWA, with very few instances of poor forecasts at more than three of the TAF sites simultaneously during the 12-hour forecast period (Figure 4.6).

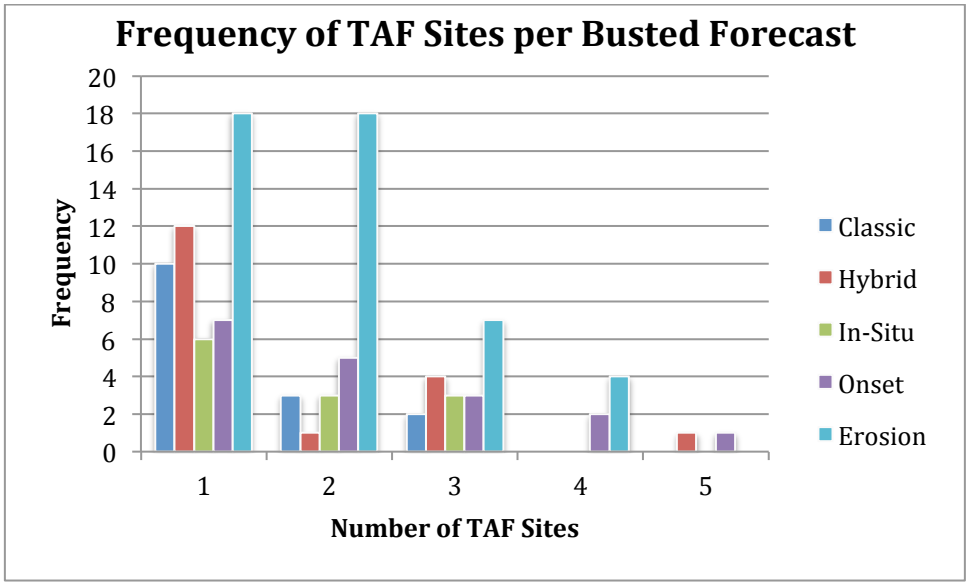


Figure 4.6. The frequency with which a busted cold-air damming forecast was evident at a single TAF site or at multiple TAF sites simultaneously, 2007 to 2016.

Of the 110 total busted CAD forecasts within the database, there were 53 instances for which the temperature forecast busted at only one of the six TAF sites, 30 at two of the sites, 19 at three sites, 6 at four sites, and two at five sites (Figure 4.6). No CAD cases busted at all six TAF sites, which may indicate that erroneous forecasts tend to be localized rather than across the entire CWA; this may result from mesoscale idiosyncrasies produced by CAD that are further complicated by local physiography-induced microclimates. Overall, busted temperature forecasts occurred 42 times in Lewisburg (LWB), 36 in Bluefield (BLF), 34 in Blacksburg (BCB), 31 in Lynchburg (LYH), 31 in Roanoke (ROA), and 30 in Danville (DAN) (Figure 4.7).

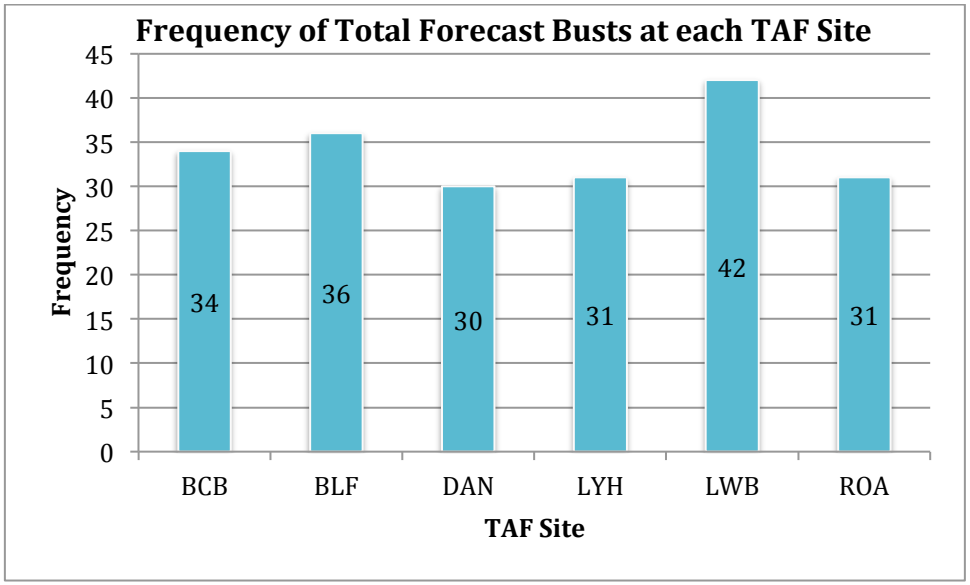


Figure 4.7. The frequency of busted cold-air damming forecasts stratified by TAF site location, 2007 to 2016.

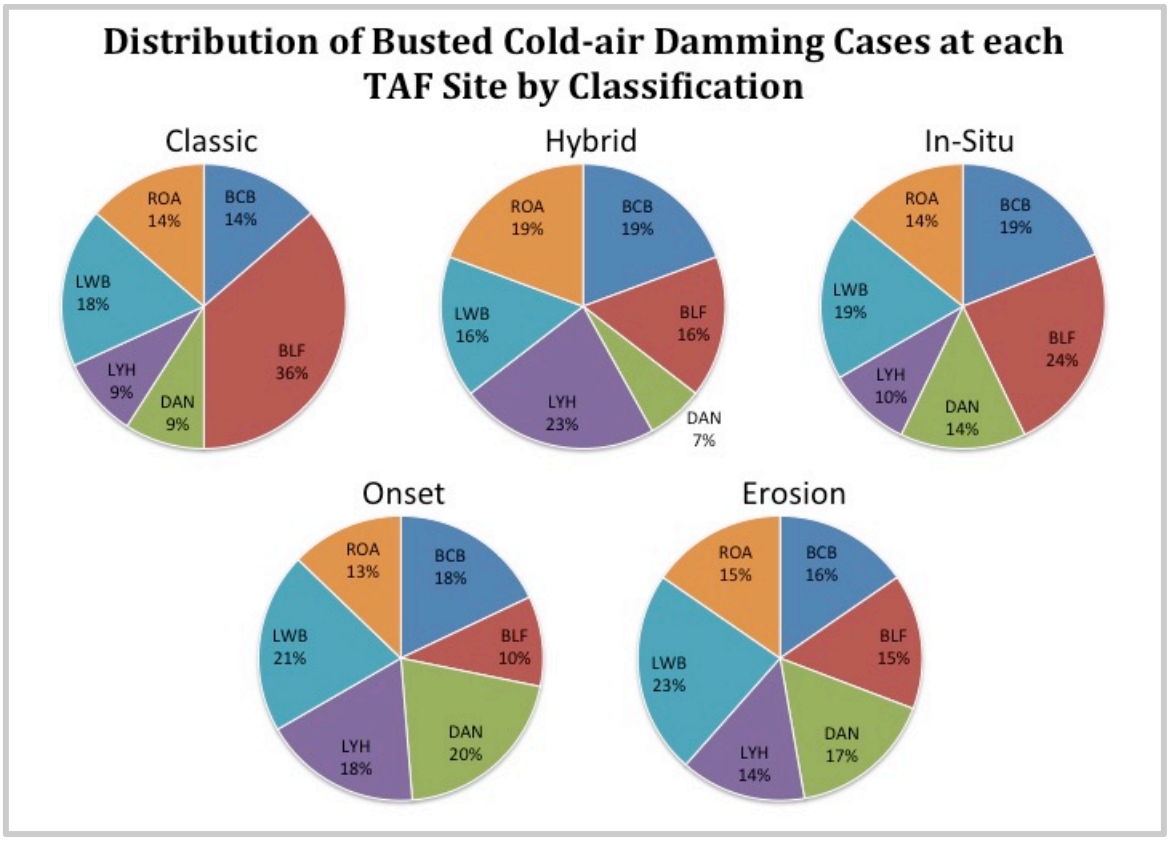


Figure 4.8. The distribution of busted cold-air damming cases at each TAF site segregated by classification type, 2007 to 2016.

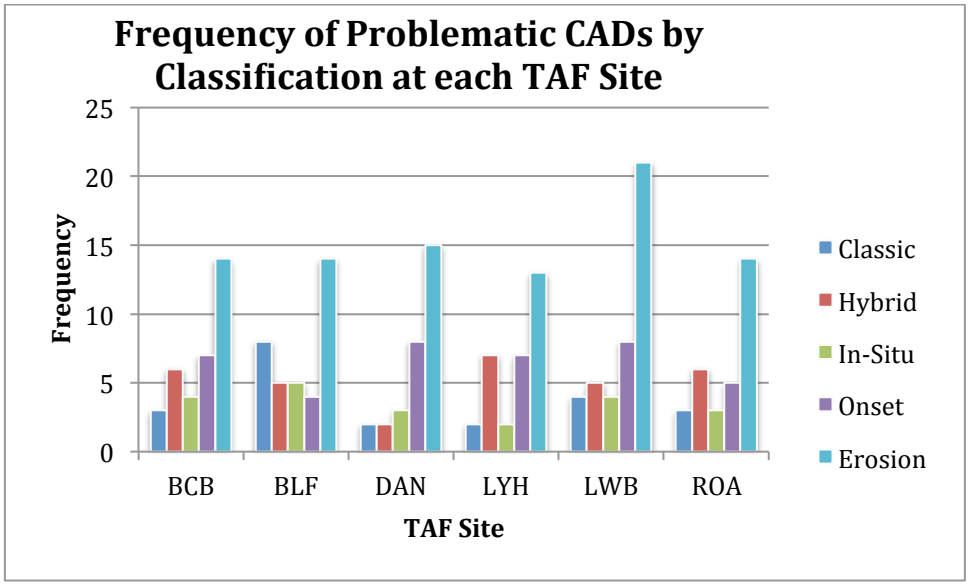


Figure 4.9. The frequency of problematic cold-air damming forecasts stratified by TAF site location and cold-air damming classification, 2007 to 2016.

Busted CAD forecasts seem to be rather well distributed across the six TAF sites, despite a proportionally high percentage of classic busted forecasts at Bluefield and a proportionally low number of hybrid busts at Danville (Figure 4.8). This may be a product of BLF and DAN situated on the edge of the cold dome boundary during these cases, producing difficult forecast scenarios. The number of occurrences of busts associated with CAD erosion is comparable for the TAF sites except for Lewisburg, which experiences more frequent busts (Figure 4.9). Forecast busts are also more frequent at higher-elevated sites like LWB, BLF, and BCB. DAN and LYH experience the fewest number of mis-forecasted CAD events, both with relatively few incidences of busted forecasts associated with classic and in-situ CAD scenarios; terrain influence may affect these cases less, and since these sites are generally some of the last to see the wedge erode from the region, they may see more

consistent conditions created by CAD that are perhaps easier to predict. There does not appear to be a correlation between classification type and TAF site distribution.

4.3 MAV and MET Guidance Temperature Bias

The accuracy of MAV and MET maximum and minimum temperature forecasts during CAD bust events between 2007 and 2016 are evaluated out to 60 hours lead-time to identify any temperature biases the GFS and NAM models may exhibit in the form of MOS guidance statistics. Temperature biases for each CAD type regardless of TAF site consider the mean error, absolute error, percent error, and bias statistic (forecast divided by actual) for MET and MAV erroneous high and low temperature forecasts out to five forecast cycles (-60 hours). Mean error is used to assess how temperature forecasts deviate from actual temperatures, identifying a warm or cool bias based on the sign of the error. Absolute error reveals the average departure from observed temperatures by calculating the magnitude of error from model-driven forecasts regardless of the sign of the error. Bias highlights precision within the average departure from observed temperatures and gauges forecast accuracy regardless of spatial location with a value of one representing an ideal forecast (NOAA – WPC Verification Threat Score and Bias Computation, 2013). These statistics, when combined, will reveal any model biases that may be present in MET and MAV-derived temperature guidance of problematic CAD events between 2007 and 2016.

The problematic forecasts in this study are influenced by underlying factors that caused forecasters to produce poor forecasts using MOS guidance input to help shape their forecasts. Typical CAD events yield depressed maximum temperatures

coupled with elevated minimum temperatures for a relatively small diurnal temperature range as discussed in the literature review. When a CAD scenario remains undetected by MOS guidance, MAV and MET solutions will produce warm-biased maximum temperature forecasts and cool-biased minimum temperature forecasts. Alternately, if CAD does not occur but is forecast to occur by GFS and NAM models and is reflected in statistically-derived MOS output, MAV and MET solutions will be cool-biased in maximum temperature forecasts and warm-biased in minimum temperature forecasts. Additionally, MOS guidance tends towards climatology as time increases ahead of the forecast period, influencing MOS temperature biases across all forecast cycles, but more heavily during later cycles (Carter et al. 1989). This often produces less accurate forecasts with greater lead time (Stephen Keighton – NWSFO Blacksburg, Personal communication).

4.3.1 Temperature Bias Results by Cold-air Damming Classification

MAV and MET maximum temperature forecasts for problematic classic (Table 4.2) and in-situ (Table 4.4) events exhibit warm biases across high temperature forecasts. The strength of the bias signal increases ahead of the forecast period, as does the inaccuracy of model-derived forecasts with greater lead-time. MET high temperature forecasts during problematic hybrid CAD events follow the same trend, whereas MAV exhibits a consistent warm bias in advance of the forecast period throughout all hybrid high temperature busts (Table 4.3). MET high temperature forecasts during problematic classic CAD reflect a small peak in warm bias error during the second forecast cycle (24 hours ahead of the forecast period), producing greater inaccuracy than do the surrounding forecast cycles.

Table 4.2. Error statistics for forecasts of problematic classic cold-air damming events by forecast cycle, 2007 to 2016. Statistics are presented for high and low temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
		MAV, High			MET, High	
1 (-12h), n=13	2.9	9.2	1.07	1.3	5.9	1.04
2 (-24h), n=13	3.2	9.1	1.08	2.2	6.5	1.05
3 (-36h), n=13	3.6	10.4	1.09	1.7	5.5	1.04
4 (-48h), n=13	4.3	10.5	1.11	1.9	5.9	1.05
5 (-60h), n=13	4.6	11.5	1.12	2.7	6.1	1.08
		MAV, Low			MET, Low	
1 (-12h), n=9	-8.7	8.7	0.82	-7.9	7.9	0.84
2 (-24h), n=9	-9.0	9.0	0.81	-6.0	6.0	0.87
3 (-36h), n=9	-8.4	8.4	0.82	-6.9	6.9	0.85
4 (-48h), n=9	-8.6	8.6	0.82	-7.8	7.8	0.84
5 (-60h), n=9	-8.7	8.7	0.82	-8.0	8.0	0.84

Table 4.3. As in Table 4.2, except for problematic hybrid CAD events.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
		MAV, High			MET, High	
1 (-12h), n=24	5.8	6.4	1.11	5.5	5.8	1.11
2 (-24h), n=24	6.2	8.1	1.12	5.8	7.4	1.11
3 (-36h), n=24	7.5	9.0	1.14	7.4	7.9	1.14
4 (-48h), n=24	8.5	9.4	1.16	7.0	8.5	1.14
5 (-60h), n=24	9.6	11.1	1.17	9.9	10.0	1.18
		MAV, Low			MET, Low	
1 (-12h), n=7	0.9	7.1	1.03	-0.7	6.4	1.00
2 (-24h), n=7	1.3	7.6	1.04	0.3	8.6	1.02
3 (-36h), n=7	0.3	9.1	1.02	0.0	7.4	1.02
4 (-48h), n=7	-1.3	9.9	1.00	-1.0	11.0	1.01
5 (-60h), n=7	-0.3	8.9	1.01	-1.3	9.9	1.00

Table 4.4. As in Table 4.2, except for problematic in-situ CAD events and only for maximum (high) temperature.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
		MAV, High			MET, High	
1 (-12h), n=20	6.3	7.6	1.11	4.7	5.7	1.08
2 (-24h), n=20	6.0	9.6	1.11	4.3	7.3	1.07
3 (-36h), n=20	7.5	8.8	1.13	5.3	7.2	1.09
4 (-48h), n=20	7.0	8.0	1.12	7.8	9.1	1.14
5 (-60h), n=20	8.3	9.7	1.14	7.9	8.3	1.14

Both MAV and MET guidance forecasts of minimum temperature produce values too low for all classic CAD forecast cycles (Table 4.2), as well as throughout

the single nighttime in-situ event (not shown). The strength of the cold-biased signal increases ahead of the forecast period for both CAD types, as does model guidance error. Alternately, MAV low temperature forecasts during problematic hybrid CAD situations exhibit an initial warm bias that transitions into a cool bias farther ahead of the forecast period as the accuracy of MAV low temperature forecasts rather consistently decreases as lead-time increases (Table 4.3). MET low temperature forecasts for hybrid busts illustrate an inconsistent temperature bias through the forecast cycles, trending from a cool to warm bias and back to a cool bias moving back in time ahead of the forecast period. In the third forecast cycle (36 hours ahead of the forecast period), the mean MET bias for low temperature in problematic hybrid CAD cases is zero; this could result from the limited sample size of low temperature forecasts associated with problematic hybrid CAD cases.

MET temperature forecasts are more accurate for both high and low temperatures in all forecast cycles compared to the MAV for classic and in-situ CAD scenarios. Both model statistics forecasted greater-than-actual high temperatures and lesser-than-actual low temperatures in these scenarios. While the above is true for high temperature forecasts during problematic hybrid CAD events, neither MAV nor MET guidance portray a systematic warm or cold bias in low temperature forecasts.

Table 4.5. Error statistics for forecasts of all classic, hybrid, and in-situ problematic cold-air damming events by forecast cycle, 2007 to 2016. Statistics are presented for high and low temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=57	5.3	7.4	1.10	4.3	5.8	1.08

2 (-24h), n=57	5.5	8.9	1.11	4.4	7.1	1.09
3 (-36h), n=57	6.6	9.2	1.13	5.4	7.1	1.10
4 (-48h), n=57	7.0	9.1	1.13	6.1	8.1	1.12
5 (-60h), n=57	8.0	10.7	1.15	7.6	8.5	1.14
		MAV, Low			**MET, Low**	
1 (-12h), n=17	-4.6	7.9	0.91	-5.0	7.4	0.90
2 (-24h), n=17	-4.5	8.2	0.91	-3.6	7.2	0.93
3 (-36h), n=17	-4.7	8.5	0.91	-4.0	7.1	0.92
4 (-48h), n=17	-5.2	8.7	0.90	-4.8	8.9	0.91
5 (-60h), n=17	-4.9	8.4	0.91	-5.0	8.5	0.91

Both MAV and MET high temperature forecasts for all three CAD classifications (classic, hybrid, and in-situ) combined exhibit an overall warm bias that consistently increases farther ahead of the forecast period (Table 4.5) (Figure 4.10a, 4.10b); meanwhile, forecast accuracy for both MAV and MET decrease through the forecast cycles as lead-time increases.

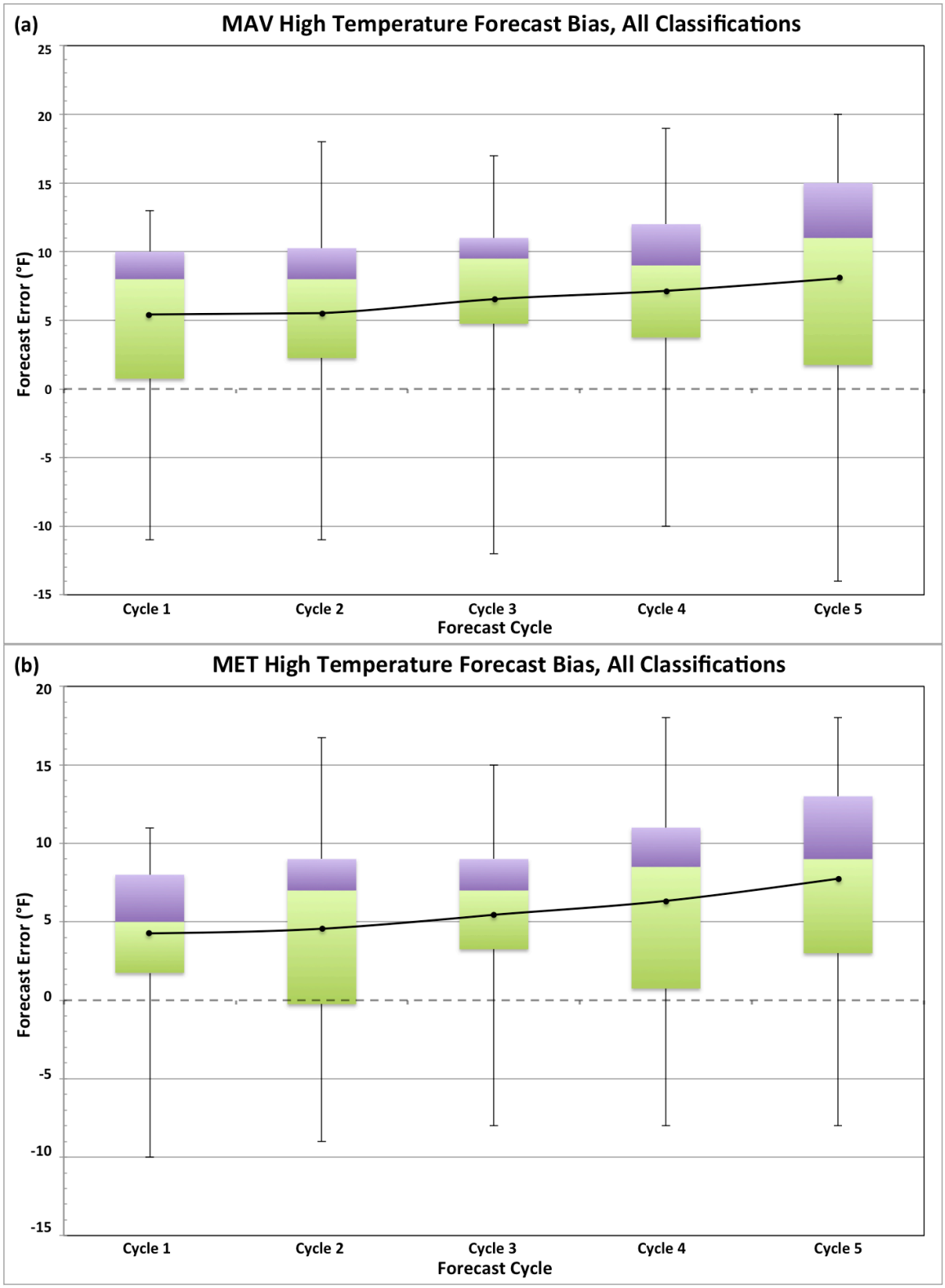


Figure 4.10a-b. Box plots of (a) MAV and (b) MET maximum temperature guidance forecast error of all classified (classic, hybrid, in-situ) problematic CAD events by forecast cycle (-12 hours to -60 hours), 2007 to 2016. The tops of the purple and

bottoms of the green boxes represent the upper and lower quartiles of forecast error, respectively, for each period, and whisker bars indicate minimum and maximum data values across each period's data range. Between the two boxes lies the median of each forecast cycle's error, and mean error values are marked with a black circle and connected with a smoothed trend line.

Alternately, MAV and MET low temperature forecasts exhibit a consistent overall cold bias and a slight increase in accuracy in the second forecast cycle (24 hours ahead of the forecast period) (Table 4.5) (Figure 4.11a, 4.11b). MET forecast temperatures are generally more accurate for both high and low temperatures in all classified CAD forecast cycles when compared to the MAV. Both model statistics forecast greater-than-actual high temperatures and lower-than-actual low temperatures in these scenarios, suggesting the problematic CAD cases were rather marginal.

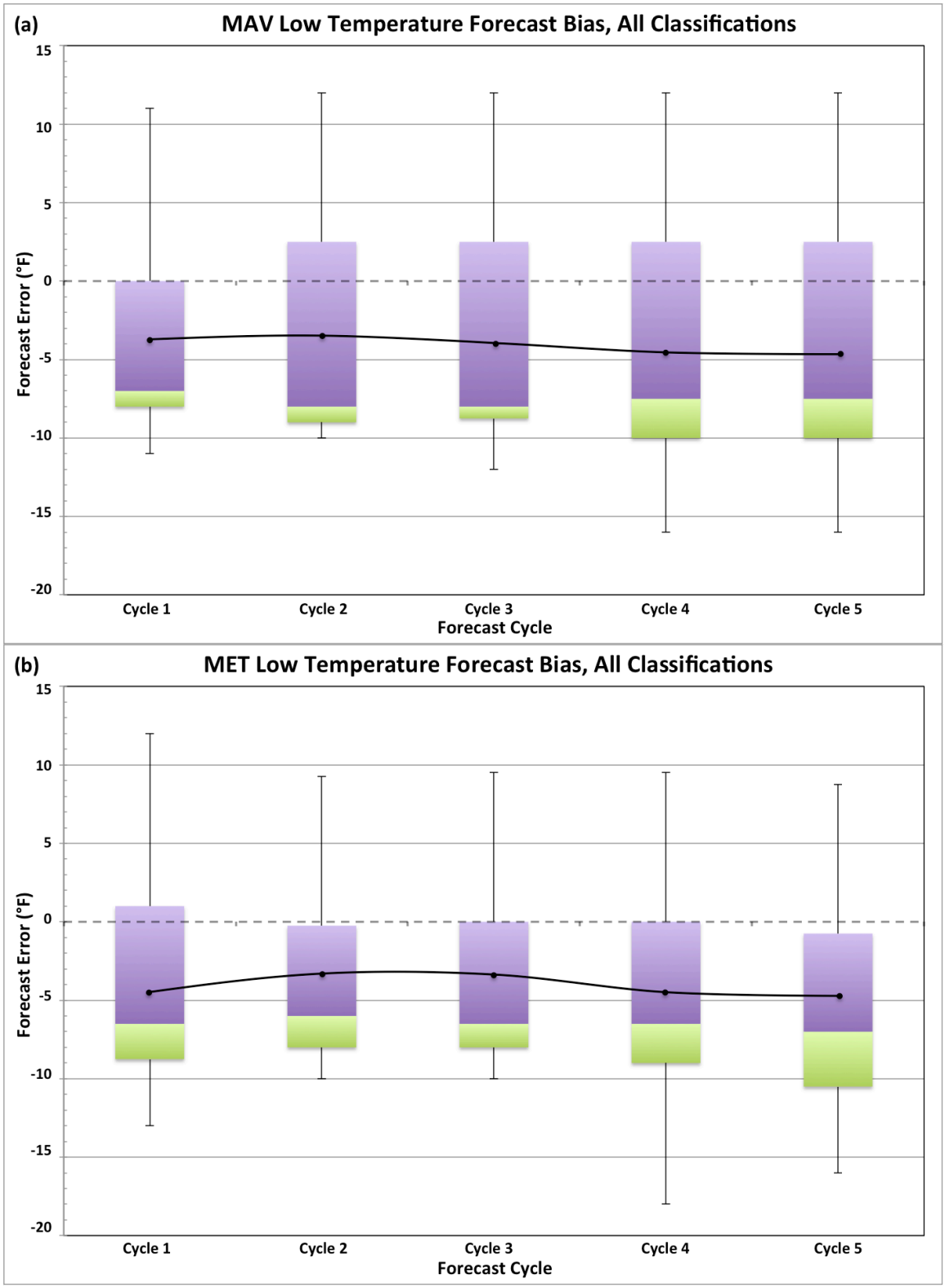


Figure 4.11a-b. As in Figure 4.11a, b, except for (a) MAV and (b) MET minimum temperature guidance error.

4.3.2 Temperature Bias Results During Cold-air Damming Onset

MAV high temperature forecasts for CAD events characterized by problematic onset exhibit a warm bias that increases ahead of the forecast period (Table 4.6). MET high temperature forecasts follow a similar warm bias, but the bias is not monotonic through the forecast cycles as for the MET, which is generally more accurate than the MAV.

Table 4.6. Error statistics for forecasts of cold-air damming events characterized by problematic onset by forecast cycle, 2007 to 2016. Statistics are presented for high and low temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=26	3.3	8.4	1.07	3.0	7.1	1.06
2 (-24h), n=26	3.4	9.9	1.07	2.5	8.1	1.06
3 (-36h), n=26	4.8	11.3	1.11	3.0	10.0	1.07
4 (-48h), n=26	5.0	11.2	1.11	3.4	9.0	1.08
5 (-60h), n=26	5.0	11.4	1.11	4.0	9.8	1.09
	MAV, Low			MET, Low		
1 (-12h), n=13	-5.4	8.2	0.86	-5.2	9.1	0.86
2 (-24h), n=13	-6.1	8.9	0.84	-4.9	8.1	0.87
3 (-36h), n=13	-6.8	9.7	0.82	-5.2	8.6	0.86
4 (-48h), n=13	-6.8	9.1	0.83	-5.9	8.5	0.85
5 (-60h), n=13	-5.5	8.2	0.86	-5.3	10.1	0.85

MAV low temperature forecasts for problematic onset of CAD events reveal an inconsistent cold bias that increases through the first three forecast cycles and lessens at 60 hours ahead of the forecast period (Table 4.6). The same is true for MET forecast accuracy, which exhibits an overall cold bias that varies in degree but remains consistent in sign throughout all five forecast cycles. The accuracy of MET forecasts generally decreases ahead of the forecast period aside from relatively poor forecasts for the cycle just prior to the forecast period.

MET solutions are overall more accurate for both high and low temperature forecasts in all forecast cycles during the onset of problematic CAD events. Both MAV and MET exhibit a warm bias in high temperature forecasts, while low temperature forecasts are cool-biased.

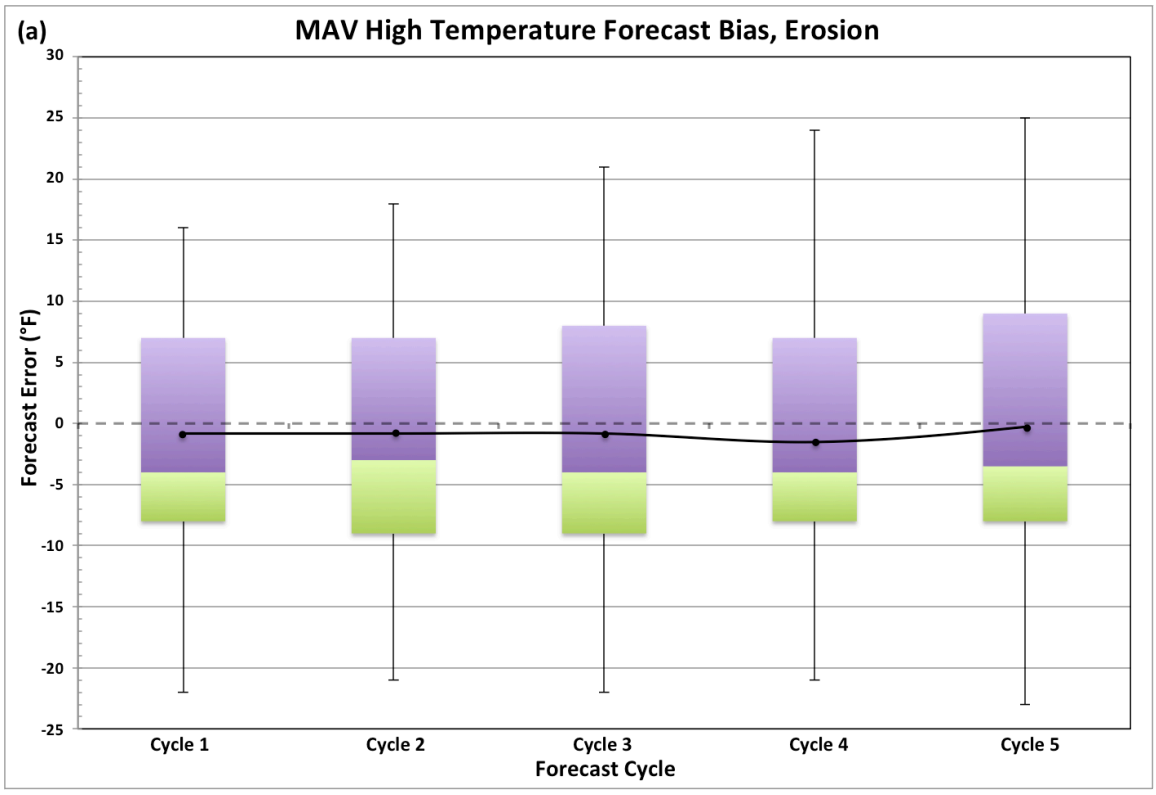
4.3.3 Temperature Bias Results During Cold-air Damming Erosion

MAV high temperature forecasts during the erosion of problematic CAD events exhibit a weak cold bias that increases ahead of the forecast period that is opposite of the other CAD types discussed above (Table 4.7) (Figure 4.12a); however, in the fifth forecast cycle (60 hours ahead of the forecast period) it is evident that warm and cold biases are of equal magnitude, as the raw error value is near zero while the absolute error is the largest for the five forecast cycles. MET high temperature forecasts also exhibit a cold bias that increases ahead of the forecast period, save for a large error in the second forecast cycle that may result from models forecasting the timing of wedge erosion too late (Table 4.7) (Figure 4.12b). A slight cool bias in maximum temperature forecasts implies MOS guidance may have a tendency to forecast the wedge to persist more often than eroding the wedge prematurely.

Table 4.7. Error statistics for forecasts of cold-air damming events characterized by problematic erosion by forecast cycle, 2007 to 2016. Statistics are presented for high and low temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=74	-1.0	7.7	1.00	-1.3	6.6	1.00
2 (-24h), n=74	-0.7	8.2	1.01	-5.1	9.7	0.90
3 (-36h), n=74	-0.9	8.6	1.01	-3.2	8.6	0.97
4 (-48h), n=74	-1.2	8.5	1.00	-3.1	9.0	0.97
5 (-60h), n=73	-0.1	9.3	1.03	-5.9	10.6	0.92

	MAV, Low			MET, Low		
1 (-12h), n=17	1.8	7.1	1.08	3.9	8.8	1.15
2 (-24h), n=17	2.6	7.2	1.11	2.8	7.8	1.12
3 (-36h), n=17	2.4	6.9	1.09	-6.5	12.2	0.80
4 (-48h), n=17	2.4	7.7	1.10	2.2	6.7	1.09
5 (-60h), n=17	2.1	7.9	1.10	3.0	7.9	1.11



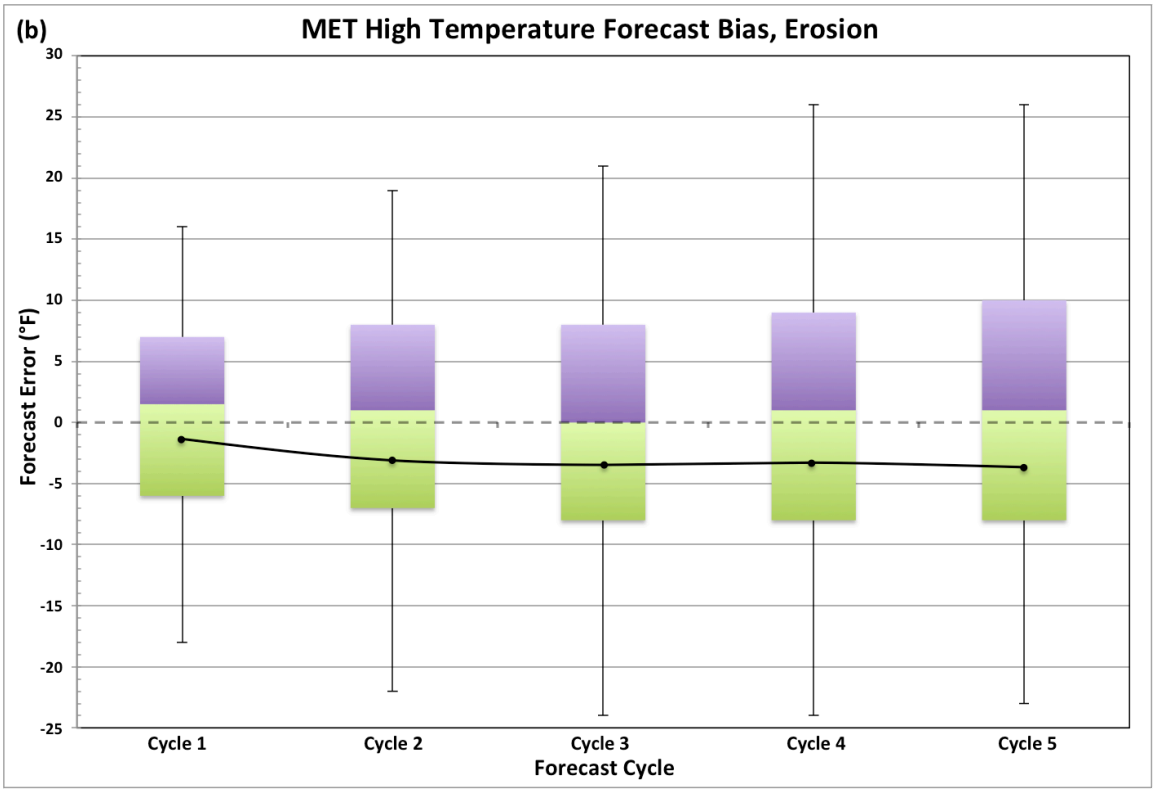
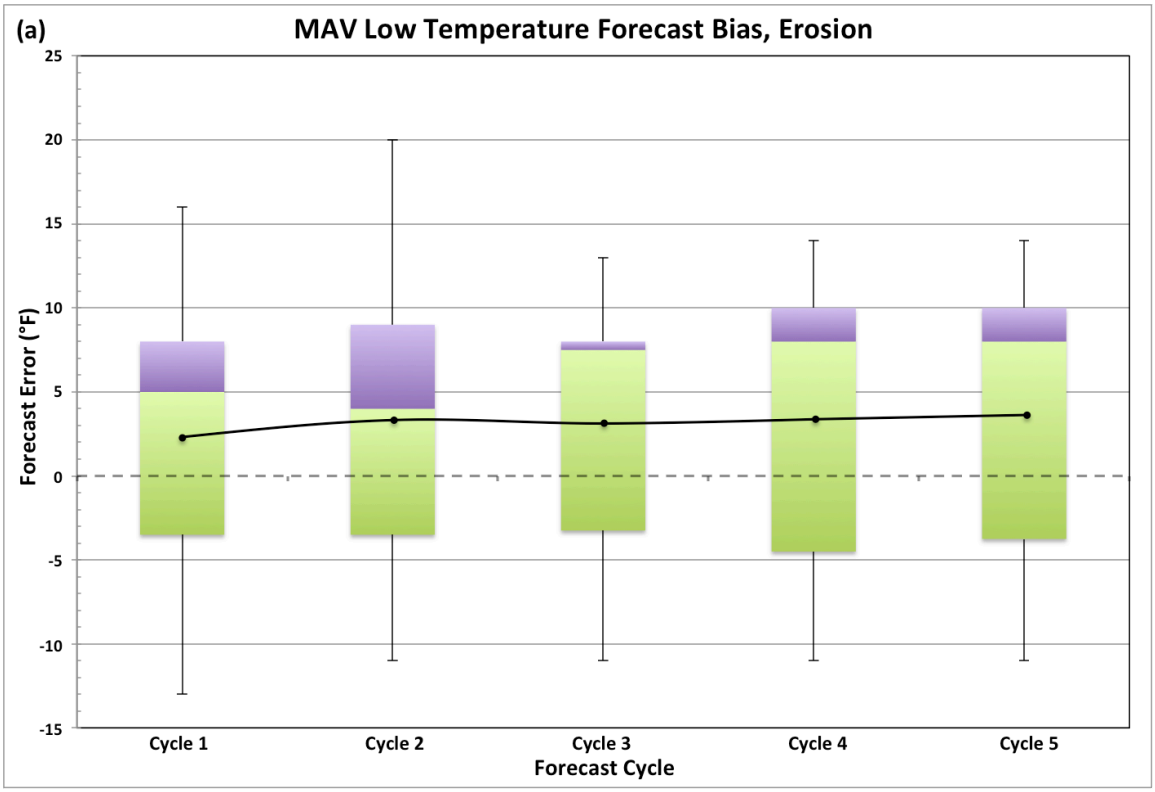


Figure 4.12a, b. Box plots of (a) MAV and (b) MET maximum temperature guidance forecast error of problematic CAD events associated with erosion by forecast cycle (-12 hours to -60 hours), 2007 to 2016. The tops of the purple and bottoms of the green boxes represent the upper and lower quartiles of forecast error, respectively, for each period, and whisker bars indicate minimum and maximum data values across each period's data range. Between the two boxes lies the median of each forecast cycle's error, and mean error values are marked with a black circle and connected with a smoothed trend line.

MAV low temperature forecasts maintain a somewhat consistent warm bias through the forecast cycles, opposite of the other CAD cases discussed above, increasing slightly ahead of the forecast period (Table 4.7) (Figure 4.13a). MAV forecast accuracy also remains fairly consistent through the forecast cycles, decreasing ahead of the forecast period. MET low temperature forecasts exhibit a warm bias that slowly lessens ahead of the forecast period and becomes more accurate with greater lead time (Table 4.7) (Figure 4.13b). However, at 36 hours prior to the forecast period, there is a large cool bias outlier in MET low

temperature forecasts, and the poorest accuracy for all MET low temperature forecasts. The slight warm bias exhibited in MAV and MET low temperature guidance suggests the NAM and GFS may more commonly tend prematurely erode the wedge at night, possibly stemming from inadequate modeling of overnight cloud cover or moisture levels.



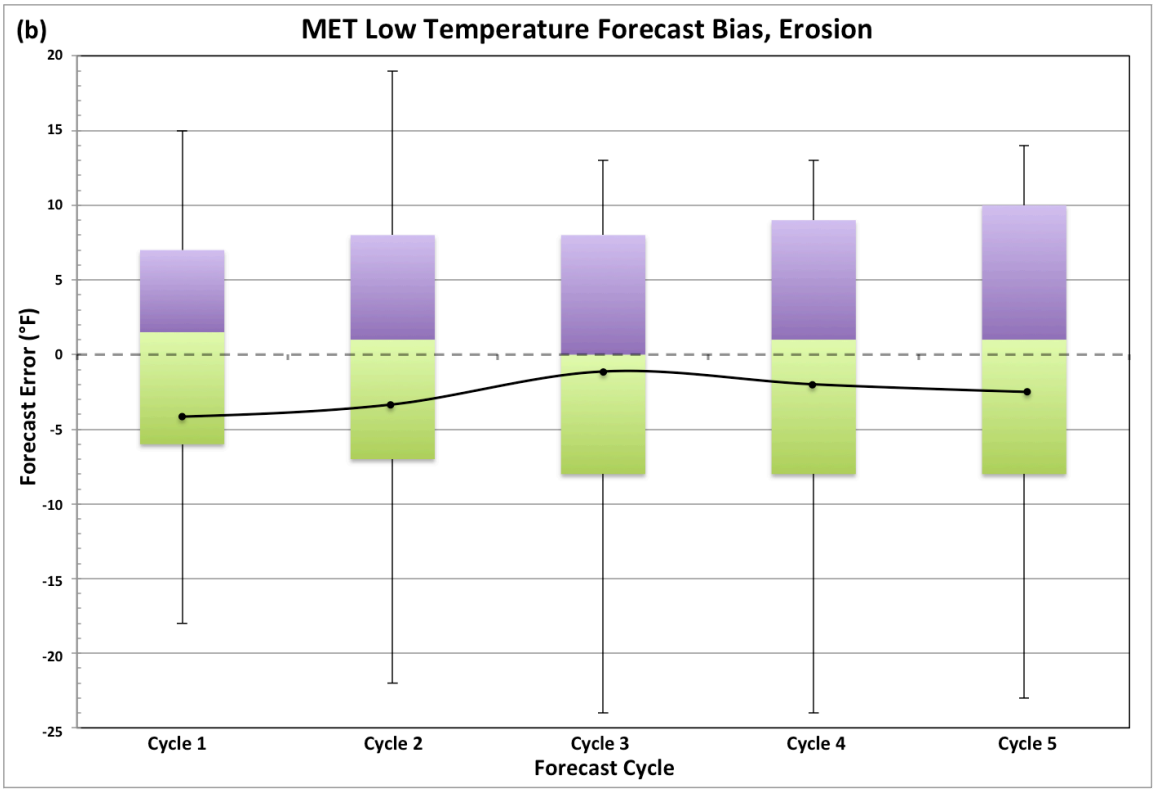


Figure 4.13a, b. As in Figure 4.12a-b, except for (a) MAV and (b) MET minimum temperature guidance error.

Erosion is the only CAD classification with a cold bias for both MET and MAV guidance during high temperature forecasts and a warm bias for both MET and MAV low temperature forecasts. This suggests that models used in MOS guidance may have forecast the wedge to persist during the day when in actuality it eroded premature of model predictions; alternately, these models tend to prematurely erode the wedge at night when it persists in actuality. Overall, MAV forecasts are more accurate than MET forecasts for both high and low temperatures, opposite of the above CAD types. Both MAV and MET produce large anomalies within the sequence of 12-hour forecast cycles for high and low temperatures that are generally larger than for other classifications, despite having the largest data population of any of the classification types.

4.3.4 Statistical Significance

Relative MAV and MET forecast guidance accuracy is assessed using two-tailed t-tests for both maximum and minimum temperature forecasts to determine whether the accuracies of the two model output statistics are significantly different when considering the problematic CAD events between 2007 and 2016 (Table 4.8).

Table 4.8. Two-Sample t-test results of significance for MAV and MET maximum and minimum forecasts over the five 12-hour forecast cycles in advance of problematic cold-air damming events, 2007 to 2016. Statistics are presented for both maximum and minimum temperature forecasts, irrespective of CAD type, for MAV and MET-based guidance through all forecast cycles. Sample size, sample mean (°F), standard deviation (°F), standard error (°F), and p-value (unitless) are presented.

	Maximum Temperature Forecast Cycle									
	Cycle 1 (-12h)		**Cycle 2 (-24h)**		**Cycle 3 (-36h)**		**Cycle 4 (-48h)**		**Cycle 5 (-60h)**	
	MAV	**MET**	**MAV**	**MET**	**MAV**	**MET**	**MAV**	**MET**	**MAV**	**MET**
N-size	156	156	153	153	156	156	156	156	153	153
Mean	1.94	1.40	2.08	0.76	2.77	0.93	2.80	1.20	3.90	2.01
StDev	8.48	7.28	9.56	8.45	9.94	9.30	10.10	10.10	10.90	9.95
SE	0.68	0.58	0.77	0.68	0.79	0.74	0.81	0.81	0.88	0.80
P-val	0.553		0.202		0.091		0.180		0.116	
	Minimum Temperature Forecast Cycle									
	Cycle 1 (-12h)		**Cycle 2 (-24h)**		**Cycle 3 (-36h)**		**Cycle 4 (-48h)**		**Cycle 5 (-60h)**	
	MAV	**MET**	**MAV**	**MET**	**MAV**	**MET**	**MAV**	**MET**	**MAV**	**MET**
N-size	47	47	47	47	44	44	47	47	46	46
Mean	-2.40	-1.74	-2.28	-1.51	-3.30	-2.91	-2.77	-2.45	-2.89	-2.37
StDev	7.98	9.06	8.53	8.28	8.49	7.58	8.84	8.58	9.02	9.19
SE	1.2	1.3	1.2	1.2	1.3	1.1	0.71	0.69	1.3	1.4
P-val	0.709		0.660		0.822		0.859		0.784	

Test results show that neither model was statistically significantly better than the other through all five forecast cycles in both maximum and minimum temperature forecast guidance. Despite a lack of significant statistical findings, operational forecasters can possibly apply the nuanced biases in MAV and MET model-derived forecasts during busted CAD events to improve forecasting decisions when combined with conceptual knowledge and model guidance during these scenarios.

4.4 Spatial MAV and MET Guidance Bias

Temperature biases at each of the six terminal aerodrome forecast (TAF) sites in the Blacksburg CWA are characterized using the measures mean error, absolute error, percent error, and the bias statistic (forecast divided by actual). Values are computed for MET and MAV high and low temperature forecasts out to five forecast cycles (-60 hours) from the forecast period. This analysis of TAF site bias is segregated by CAD classification, CAD onset, and CAD erosion. Furthermore, a lack of robust sample size at each site, particularly for low temperature forecast errors and forecasts for CAD classifications at Danville, does not allow for statistical tests of significance and many low temperature forecast statistics are not shown due to limited n-size. This discussion aims to provide guidance to forecasters as a nuance to consider when forming CAD forecasts in the CWA.

4.4.1 Spatial Temperature Bias of Cold-air Damming Classifications

Blacksburg MAV and MET forecasts for problematic CAD events exhibit a warm bias evident in all high temperature forecast cycles prior to the forecast period (Table 4.9) and a weak bias in low temperature forecasts (not shown), in which the MET is more accurate across the board.

Table 4.9. Error statistics for forecasts of problematic cold-air damming events at the Blacksburg TAF site by forecast cycle, 2007 to 2016. Statistics are presented for high and low temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=11	7.4	8.5	1.16	6.1	6.6	1.14
2 (-24h), n=11	9.9	10.6	1.21	8.1	8.8	1.17
3 (-36h), n=11	9.7	11.6	1.20	8.1	8.6	1.17
4 (-48h), n=11	10.2	11.8	1.21	9.1	9.8	1.19
5 (-60h), n=11	12.6	13.8	1.26	10.6	11.1	1.22

The problematic forecasts for the other TAF sites (Bluefield, Table 4.10; Danville (not shown); Lewisburg, Table 4.11; Lynchburg, Table 4.12; and Roanoke, Table 4.13) exhibit a warm bias for MAV and MET high temperatures across all forecast cycles. A cold bias is exhibited across Danville, Lewisburg, Lynchburg, and Roanoke minimum temperature forecasts in all forecast cycles, while Bluefield did not experience any low temperature forecast busts during classic, hybrid, and in-situ CAD events. The MET either performs consistently with the MAV or more accurately during these cases segregated by TAF site.

Table 4.10. Error statistics for forecasts of problematic cold-air damming events at the Bluefield TAF site by forecast cycle, 2007 to 2016. Statistics are presented for high and low temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=18	6.8	8.7	1.14	4.3	6.0	1.09
2 (-24h), n=18	6.5	9.2	1.13	4.3	6.3	1.09
3 (-36h), n=18	8.6	10.4	1.17	5.5	6.4	1.11
4 (-48h), n=18	8.9	10.6	1.18	6.4	7.4	1.13
5 (-60h), n=18	8.8	11.4	1.18	6.9	8.2	1.14

Table 4.11. As in Table 4.10, except for error statistics at the Lynchburg TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=6	4.2	4.5	1.07	4.3	4.3	1.08
2 (-24h), n=6	2.8	5.8	1.05	3.3	6.7	1.06
3 (-36h), n=6	3.5	5.5	1.06	4.3	6.0	1.08
4 (-48h), n=6	4.7	5.0	1.08	5.7	7.3	1.11
5 (-60h), n=6	5.8	6.8	1.10	7.3	7.3	1.13
	MAV, Low			MET, Low		
1 (-12h), n=5	-4.0	8.4	0.93	-6.4	8.0	0.89
2 (-24h), n=5	-3.8	8.6	0.93	-4.2	8.2	0.93
3 (-36h), n=5	-4.8	9.6	0.92	-5.0	8.2	0.91
4 (-48h), n=5	-5.6	10.4	0.90	-6.6	10.6	0.88
5 (-60h), n=5	-4.6	9.4	0.92	-6.6	9.8	0.89

Table 4.12. As in Table 4.10, except for error statistics at the Lewisburg TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias

	MAV, High			MET, High		
1 (-12h), n=11	1.8	6.7	1.02	1.8	5.5	1.03
2 (-24h), n=11	1.6	8.1	1.02	1.1	7.1	1.02
3 (-36h), n=11	2.6	8.5	1.04	2.4	6.9	1.04
4 (-48h), n=11	3.4	8.3	1.05	2.2	7.6	1.04
5 (-60h), n=11	4.5	9.6	1.07	6.6	6.6	1.09

Table 4.13. As in Table 4.10, except for error statistics at the Roanoke TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=8	4.7	8.3	1.09	3.6	6.7	1.07
2 (-24h), n=8	5.8	9.0	1.11	0.9	8.7	1.00
3 (-36h), n=8	5.3	9.1	1.10	3.6	9.1	1.08
4 (-48h), n=8	5.4	9.0	1.11	4.4	9.4	1.10
5 (-60h), n=8	6.8	10.6	1.13	6.0	9.7	1.12

The warm bias coincides with MOS guidance tendencies to underestimate the strength of the cold air wedge (Rackley & Knox, 2016), possibly resulting from the parent GFS and NAM models parameterizing lower-than-actual inversion heights and underestimating the role of cold-air advection in strengthening the wedge. Both models lack sufficient vertical resolution to accurately resolve surface CAD conditions that are normally confined to 1 kilometer above the ground (NOAA – National Weather Service Forecast Office Blacksburg, 2017, slide 2), causing these forecasts to trend warm.

Located less than two miles from the Greenbrier River in West Virginia, the Lewisburg TAF site at the Greenbrier Valley Airport experiences more frequent fog than the other TAF sites in the bust database (Andrew Loconto – NWSFO Blacksburg, Personal communication). The proximity of Lewisburg to a source of moisture that produces frequent fog at the observation site may slow cooling and maintain warmth at the surface during the night and early morning, possibly

causing observed minimum temperatures to remain higher than MOS guidance predictions during these problematic CAD cases.

The warm maximum temperature bias coupled with a cold minimum temperature bias reflects the tendency of MOS-derived solutions to overestimate diurnal temperature trends, evident in warmer-than-actual high temperature forecasts within the wedge and colder-than-actual low temperature forecasts. The GFS and NAM behind MAV and MET guidance inadequately model 24-hour temperature changes, exaggerating diurnal differentials between 12Z and 0Z forecasts (Forbes et al., 1987); however, diurnal trends between high and low temperatures during Appalachian CAD scenarios change very little, oftentimes remaining within a couple degrees between daytime and overnight temperatures (Robert Stonefield – NWSFO Blacksburg, Personal communication).

Despite their proximity geographically, the biases in busted low temperature forecasts for Blacksburg and Roanoke are in contrast to one another – Blacksburg experiences a consistent warm bias seen in all MOS temperature forecasts, while low temperatures forecasted for Roanoke show a cool bias. Rising nearly one thousand feet (from 1175 feet in Roanoke to 2132 feet in Blacksburg) across 26 miles, the rapid elevation change may lead MOS guidance to underestimate cold-air advection and wedge strength across the CWA. This may also lead to undetected CAD conditions at elevated areas like Blacksburg, Bluefield, and Lewisburg, potentially causing MAV and MET guidance based on raw GFS and NAM model output to forecast temperatures that are too high in these higher elevations.

4.4.2 Spatial Temperature Bias During Cold-air Damming Onset

Despite rather small sample sizes, problematic CAD forecasts at Blacksburg, Bluefield, Danville, and Roanoke (all not shown) exhibit a warm bias in high temperature guidance and cold bias in low temperature guidance. The daytime warm bias suggests MOS-based GFS and NAM models predict the onset of the wedge to initialize in the Blue Ridge too late, and the opposite goes for the nighttime cold bias; however, a limited sample size of these cases minimizes the implication of these findings.

Lynchburg also exhibits a warm maximum temperature bias (not shown), yet the TAF site's minimum temperature forecasts also reveal a slight warm bias at night. This would suggest the cold dome encroaching on Lynchburg earlier than MOS predictions during the day and at night, but the limited sample size of nighttime busts at the site (2 events) is not enough to draw a conclusion.

The most frequent CAD busts associated with onset occurred at Lewisburg, in which the 6 daytime events resulted in a cold high temperature bias (not shown). The incidence angle of air as a cold air mass pools alongside the Appalachian Mountains may play a role in this site's consistent cold maximum temperature bias. Incoming cold air advection driven by the wedge's parent anticyclone in the northeast is generalized in GFS and NAM parameterization, smoothing the cold air over much of the area's complex terrain and resulting in MOS guidance forecasts that are too cold in Lewisburg during the day. Depending on the angle of incidence of incoming easterly air from the northeast, air may residually pool into higher elevations in the CWA after cold surface air has begun to pool alongside the mountains at lower elevations (Jim Hudgins – NWSFO Blacksburg, personal

communication). By assuming the cold air reaches much of the area at a consistent time, the models may depict cold air reaching the Greenbrier Valley quicker than in actuality. As with above, despite a cold minimum temperature bias at Lewisburg (not shown), a limited n-size limits the validity of these interpretations.

4.4.3 Spatial Temperature Bias During Cold-air Damming Erosion

MAV and MET forecasts for problematic CAD erosion at Bluefield and Lewisburg exhibit a cold bias evident in all high temperature erosion forecast cycles prior to the forecast period (Table 4.14, Table 4.15), as well as a cold bias in low temperature forecasts (not shown). The MET performs with higher accuracy at Bluefield, whereas the MAV provides better predictions during erosion at Lewisburg. A consistent cold bias indicates that the cold dome erodes from the Blue Ridge before NAM and GFS predictions.

Table 4.14. Error statistics for forecasts of problematic cold-air damming events associated with erosion at the Bluefield TAF site by forecast cycle, 2007 to 2016. Statistics are presented for high temperature forecasts for MAV and MET. Mean error (°F), mean absolute error (°F), and bias statistic (unitless) are presented.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=11	-6.0	7.5	0.91	-6.6	7.0	0.90
2 (-24h), n=11	-5.9	8.1	0.91	-6.2	6.4	0.90
3 (-36h), n=11	-6.7	8.2	0.90	-7.2	8.3	0.89
4 (-48h), n=11	-7.6	8.5	0.88	-6.9	7.1	0.89
5 (-60h), n=11	-7.9	9.0	0.88	-8.3	8.3	0.87

Table 4.15. As in Table 4.14, except for error statistics at the Lewisburg TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=16	-6.2	8.1	0.90	-7.0	8.4	0.89
2 (-24h), n=16	-6.3	7.9	0.90	-7.8	9.4	0.88
3 (-36h), n=16	-6.3	8.4	0.90	-8.9	10.3	0.86
4 (-48h), n=16	-7.3	9.4	0.88	-9.6	11.1	0.84
5 (-60h), n=16	-5.6	8.9	0.92	-12.8	14.0	0.78

Danville, Lynchburg, and Roanoke reveal warm biases in MAV high temperature erosion forecasts, and MET forecasts in these events fail to exhibit a consistent trend (Tables 4.16-4.19).

Table 4.16. As in Table 4.14, except for error statistics at the Danville TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=14	1.4	9.1	1.05	2.9	7.6	1.08
2 (-24h), n=14	2.6	9.9	1.08	-2.4	12.6	0.96
3 (-36h), n=14	2.5	11.2	1.08	-0.7	8.7	1.01
4 (-48h), n=14	2.9	10.1	1.08	-0.4	10.0	1.03
5 (-60h), n=14	3.0	11.6	1.09	-5.5	13.1	0.94

Table 4.17. As in Table 4.14, except for error statistics at the Lynchburg TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=11	5.0	7.2	1.11	2.0	4.0	1.05
2 (-24h), n=11	4.1	8.6	1.10	-1.7	9.4	0.98
3 (-36h), n=11	4.7	7.8	1.11	2.0	6.4	1.05
4 (-48h), n=11	3.2	7.9	1.08	1.8	7.1	1.06
5 (-60h), n=11	5.7	8.6	1.13	0.4	4.4	1.03

Table 4.18. As in Table 4.14, except for error statistics at the Roanoke TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=12	0.9	7.6	1.04	0.6	6.4	1.03
2 (-24h), n=12	2.9	7.6	1.07	-9.3	12.7	0.84
3 (-36h), n=12	1.7	8.5	1.05	-0.5	10.2	1.03
4 (-48h), n=12	1.9	8.4	1.05	1.4	10.4	1.07
5 (-60h), n=12	4.1	10.1	1.10	1.4	8.9	1.06

Moreover, neither MAV nor MET maximum temperature forecasts for Blacksburg during problematic CAD portray a systematic bias.

Table 4.19. As in Table 4.14, except for error statistics at the Blacksburg TAF site.

Forecast Cycle	Mean Error	Absolute Error	Bias	Mean Error	Absolute Error	Bias
	MAV, High			MET, High		
1 (-12h), n=10	0.4	6.2	1.04	2.4	5.2	1.08
2 (-24h), n=10	-0.6	6.6	1.02	-7.0	10.4	0.85
3 (-36h), n=10	-0.1	6.3	1.03	-1.3	7.5	1.01
4 (-48h), n=10	1.0	5.8	1.05	-1.3	8.9	1.02
5 (-60h), n=10	2.4	7.4	1.09	-7.5	13.7	0.89

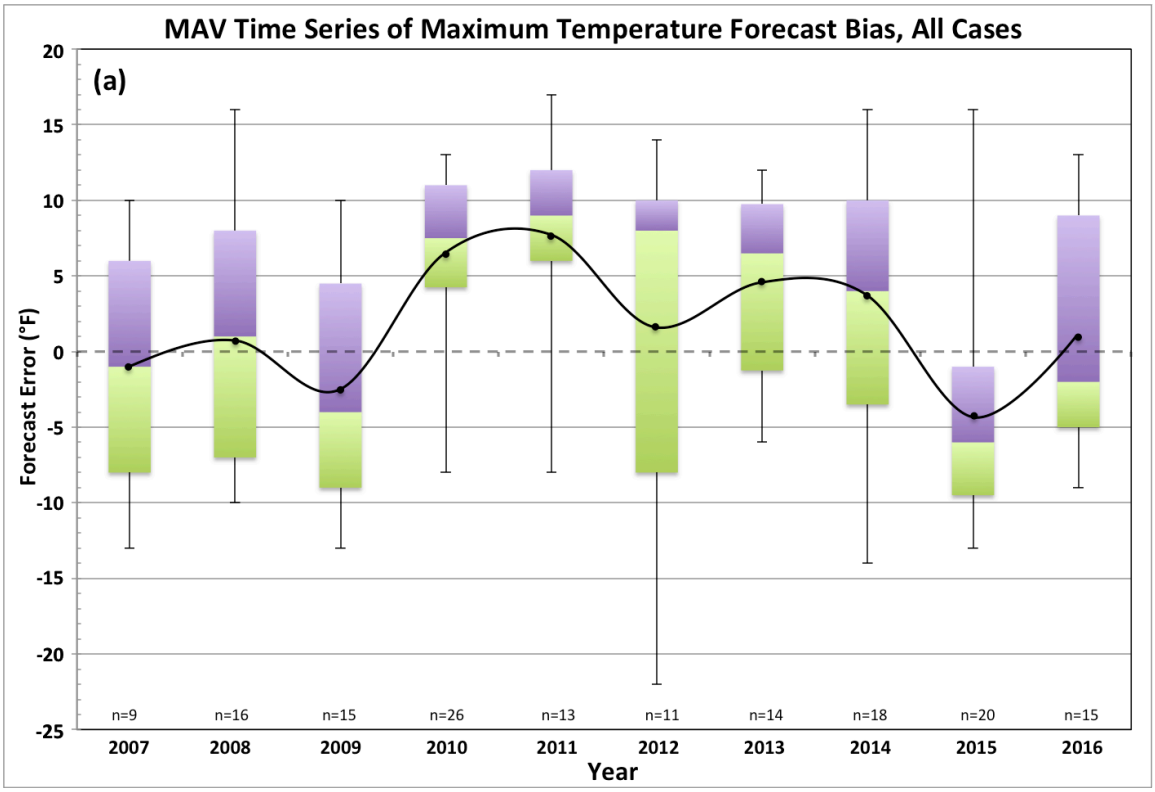
Inconsistent temperature biases during erosion indicate comparable NAM and GFS tendencies to either erode the wedge too soon or assume the wedge will persist for too long. As Blacksburg is positioned between consistent cold biases at Lewisburg and Bluefield and warm biases at Lynchburg, Danville, and Roanoke, it makes sense that this in-between site exhibits an inconclusive temperature bias. The spread of temperature biases across the CWA suggests that models tend to overestimate wedge duration at higher elevations, pointing to a potential issue of sufficiently modeling topographical affects on CAD erosion.

Though not shown due to limited sample size, minimum temperature biases during problematic CAD erosion exhibit more of the same. Low temperature forecasts for Blacksburg, Lynchburg, and Roanoke were too high, whereas minimum temperature forecasts were too low at Bluefield, Danville, and Lewisburg. These biases imply the NAM and GFS erode the wedge prematurely at night, though limited data again challenges the validity of these results.

4.5 MAV and MET Temperature Bias Over Time

Box plots of forecast error for the documented maximum temperature busted forecasts of CAD between 2007 and 2016 reveal how maximum temperature biases have varied over time. Only the first forecast cycle (-12 hours) is discussed here for brevity and for the purpose of assessing what should be the most accurate forecast. Furthermore, the sample size for minimum temperature forecasts across each year was generally not large enough to perform an analysis through time. High temperature forecast errors through the 10-year span of this study reveal similar patterns for both the MAV (Figure 4.14a) and MET (Figure 4.14b) guidance, with

MET forecast errors appearing to be slightly more conservative. Though the dispersion of error appears to be slightly larger in MAV forecasts, the two output statistics follow the same warm and cold bias patterns throughout the study period.



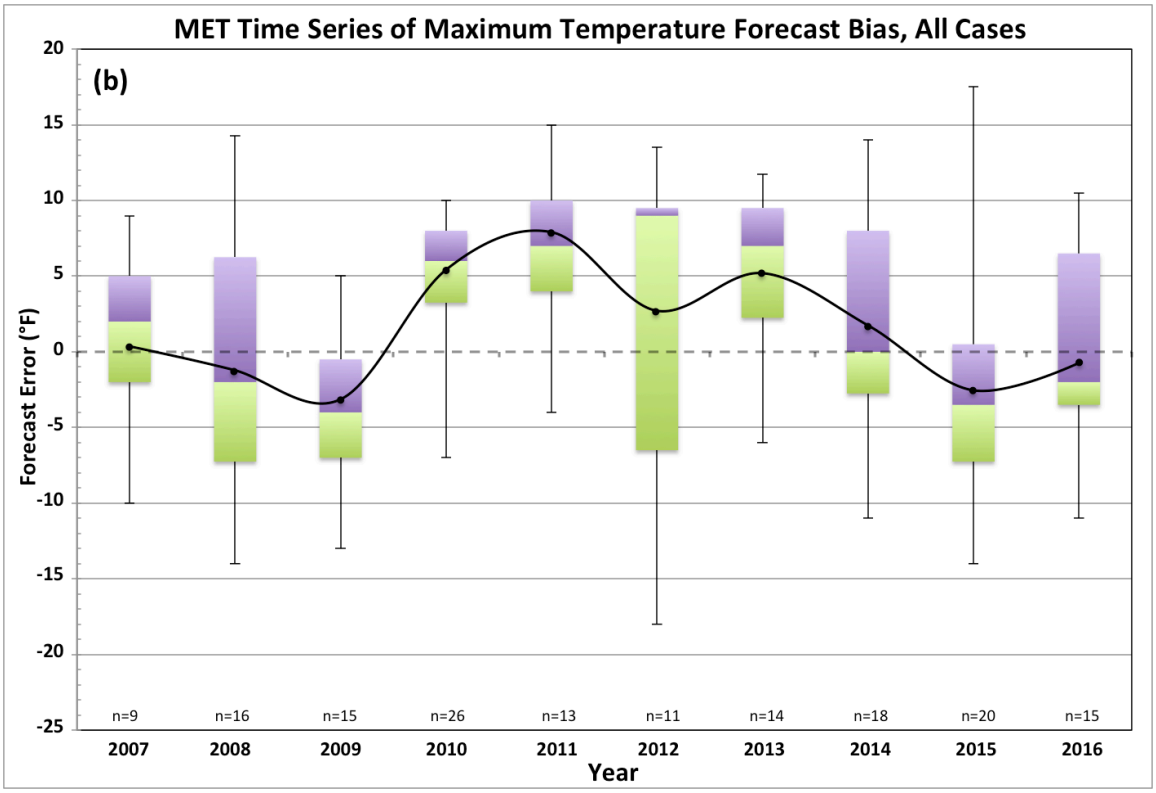


Figure 4.14a-b. Box plots of yearly MAV (a) and MET (b) guidance maximum temperature forecast errors for all problematic cold-air damming events regardless of type in the first forecast cycle (-12h) through the period 2007 to 2016. The tops of the purple and bottom of the green boxes represent the upper and lower quartiles, respectively, of forecast error for each period, and whisker bars indicate minimum and maximum data values across each year's data range. Between the two boxes lies the median of each year's forecast error, and mean error values are marked with a black circle and connected with a smoothed time series line. The sample size of each year is listed above the x-axis.

The median values of the MAV and MET errors for the first forecast cycle (-12h) for each year of the 10-year period are highly co-variable (Figure 4.15). The median of the annual error values are plotted rather than statistically assessed to avoid issues produced by outliers within the relatively small populations of data.

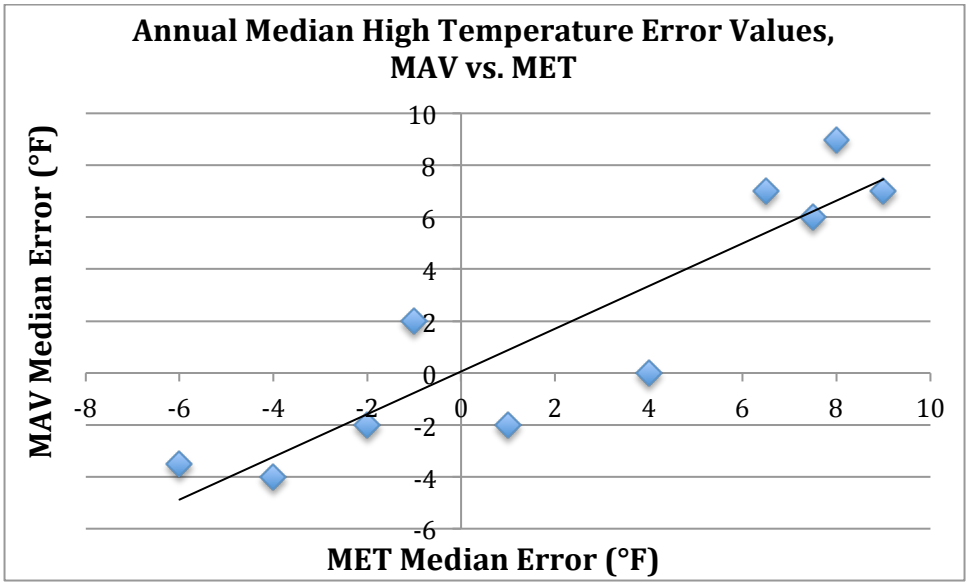


Figure 4.15. The median of annual maximum temperature forecast error values, MAV vs. MET, 12 hours in advance of the forecast period, 2007 to 2016.

Based on median maximum temperature forecast error, neither model significantly outperformed the other throughout the ten-year period. Annual median values of MAV and MET high temperature forecast error for the second through fifth forecast cycles (-24h through -60h) were also plotted and reveal the same degree of co-variability.

Taking the difference between MAV and MET (MAV minus MET) annual median error for each maximum temperature forecast cycle reveals an increase for forecast cycles of greater lead time (Figure 4.16); it is not surprising that the models tend to converge closer to the forecast period.

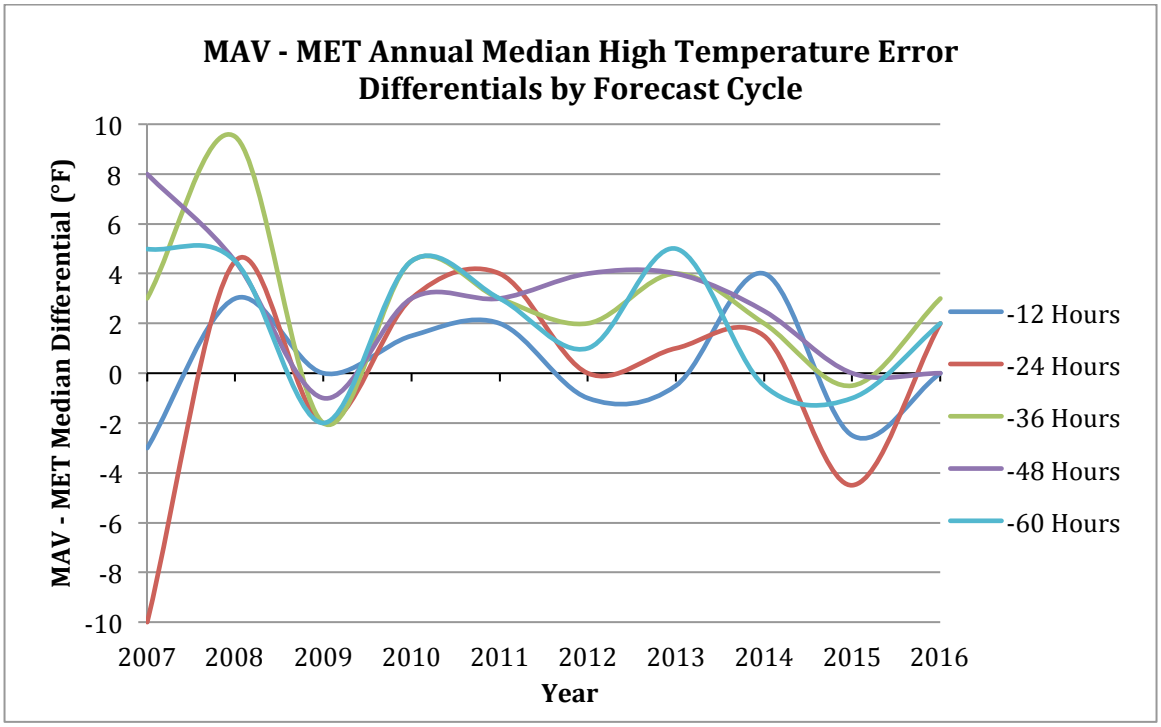


Figure 4.16. The annual median maximum temperature error differentials (MAV - MET) by forecast cycle (-12 hours to -60 hours in advance of the forecast period) between MAV and MET guidance forecast error, 2007 to 2016.

The first forecast cycle shows the smallest difference between median error of MAV and MET guidance forecasts, illustrating how the models performed similarly with the least amount of lead time (-12 hours). Though differences in median error generally fluctuate by several degrees over time, forecast cycles farther out from the forecast period reveal greater variance between MET and MAV median forecast error. This may potentially stem from MOS tendencies to weigh climatologically normal temperatures more heavily in predictions farther out from the forecast period (Stephen Keighton – NWSFO Blacksburg, Personal communication).

MAV forecasts between 2010 and 2015 consistently predicted temperatures higher than MET forecasts, most notably in forecast cycles of greater lead time; this may stem from many more warm-season CAD cases than cool-season in 2010, and

lesser resolution in GFS-based MAV guidance may have led to a high proportion of weaker hybrid and in-situ cases to progress undetected while higher resolutions in NAM-based MET guidance may have detected these events and forecasted accordingly. Additionally, on March 3rd, 2010, the GFS-based MOS forecast equations underwent updates to "produce improved forecast guidance at short-range and long-range projections from the 0Z and 12Z model runs of daytime maximum and nighttime minimum temperature" (NOAA – NCEP Central Operations, 2018). It should be noted that these updates were made for the contiguous United States, and while these changes in near- and long-term forecasts may have improved forecast accuracy across much of the country, these updates could have potentially introduced new problems to MAV guidance during Appalachian CAD. Similar updates were made on April 1st, 2015, in which the GFS-based warm and cool season equations were refreshed (NOAA – NCEP Central Operations, 2018). This equation update coincides with a shift in guidance bias, in which the MET began consistently predicting temperatures lower than the MAV. The latter half of the study period lacks an overall pattern of relative model performance based on MAV and MET median error differentials.

Overall, both guidance statistics poorly forecast the problematic CAD events to a similar degree, indicating that neither model should be used preferentially over the other when formulating CAD forecasts. This also suggests that incorporating both forms of guidance when producing forecasts may provide more useful information when considering MOS guidance temperature bias.

4.6 Hourly Surface Composites

Hourly surface atmospheric composites of observed air temperature, relative humidity, wind direction, and cloud ceiling height are plotted using all problematic CAD cases at each TAF site regardless of CAD classification. Surface conditions at each of the six TAF sites are plotted for each date within the bust database regardless of which specific sites among the six resulted in busted forecasts; data for any one site may be skewed if a CAD wedge partially eroded or was not strong enough to encompass all sites. Anomalous conditions may simply result from a lack of presence of CAD at a specific site rather than forecast error, as many sites included within the composites were likely well-forecast despite other sites within the CWA being flagged as a forecast bust during that specific CAD scenario. Cloud ceiling heights are not reported over 3,657 meters (12,000 feet), so despite the possibility of higher clouds being present at the time of hourly observations, only recorded values below this height are considered in composite calculations. Wind speeds are not shown but remain between 3 and 11 knots throughout the 24-hour period in both seasons, denoting weak flow during these CAD episodes. The composites are divided into warm season (May through October) and cold season (November through April) and provide visual insight as to whether the mean surface atmosphere during the problematic CAD events is unusual relative to what is typical for a CAD event.

4.6.1 Warm Season Hourly Surface Composites

Of the 37 CAD events with busted forecasts that occurred during the warm season (May to October) between 2007 and 2016, the diurnal temperature range did not exceed 10°F across all TAF sites (Figure 4.17).

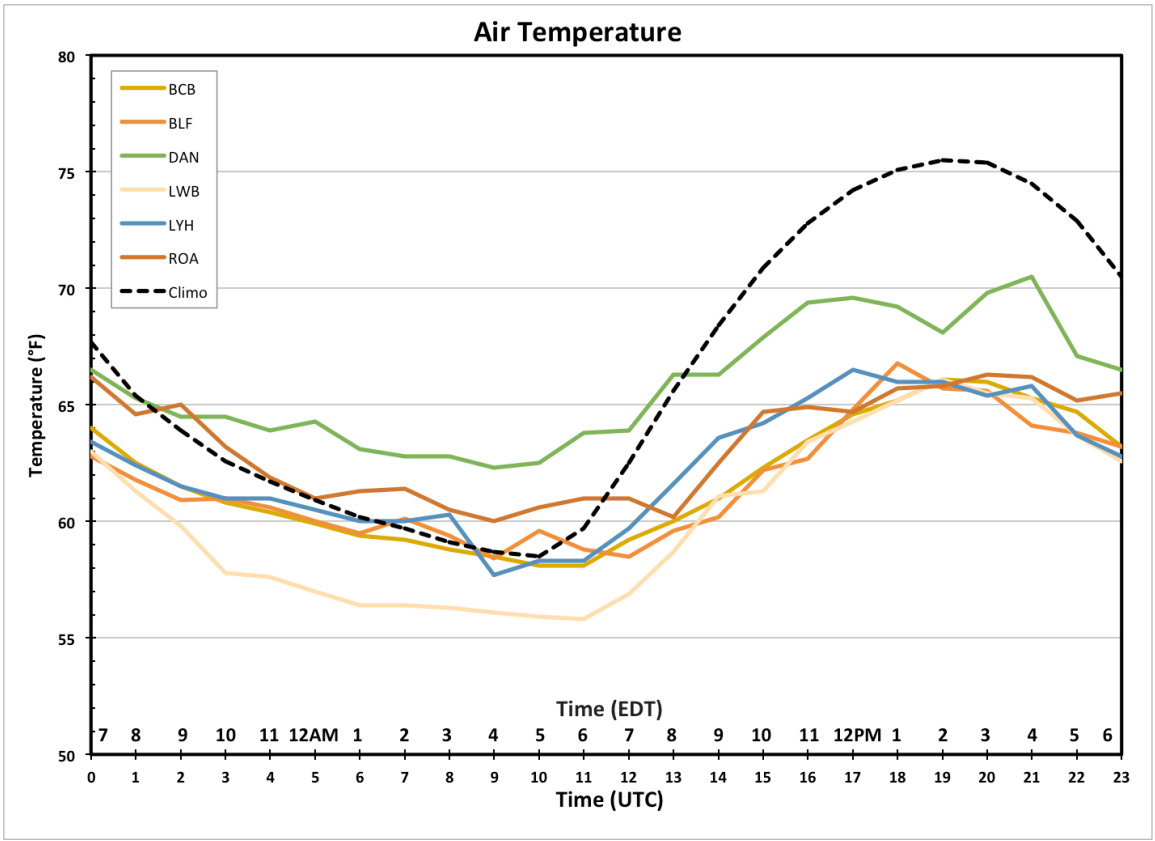


Figure 4.17. Hourly surface composites of air temperature (°F) during problematic warm season cold-air damming events at each of the six TAF sites (Blacksburg (BCB), Bluefield (BLF), Danville (DAN), Lewisburg (LWB), Lynchburg (LYH), and Roanoke (ROA)). For reference, the climatological normal temperature between May and October for the ten-year period at Blacksburg is plotted as a dashed line.

Climatological minimum and maximum temperatures at Blacksburg range 18°F over the course of 24 hours, whereas most TAF site diurnal temperatures remain within an 8°F range. Lewisburg is an exception and remains coldest at night by several degrees, but stays within a 10°F diurnal range. The compressed diurnal temperature trend during problematic CAD events shows a narrower range than climatologically

normal at Blacksburg for the region coupled with depressed daytime temperatures roughly 10°F lower relative to normal at Blacksburg than at most sites, typical of CAD. Danville's relatively high temperatures may be a result of its southern location, or the site's tendency to be on the outside of the wedge boundary – more notably during summer months when the cold dome is often weaker (Robert Stonefield – NWSFO Blacksburg, Personal communication).

Warm season composite relative humidity values for problematic CAD events reflect the normal diurnal pattern, but conserved relative to the range in climatological values for Blacksburg (Figure 4.18).

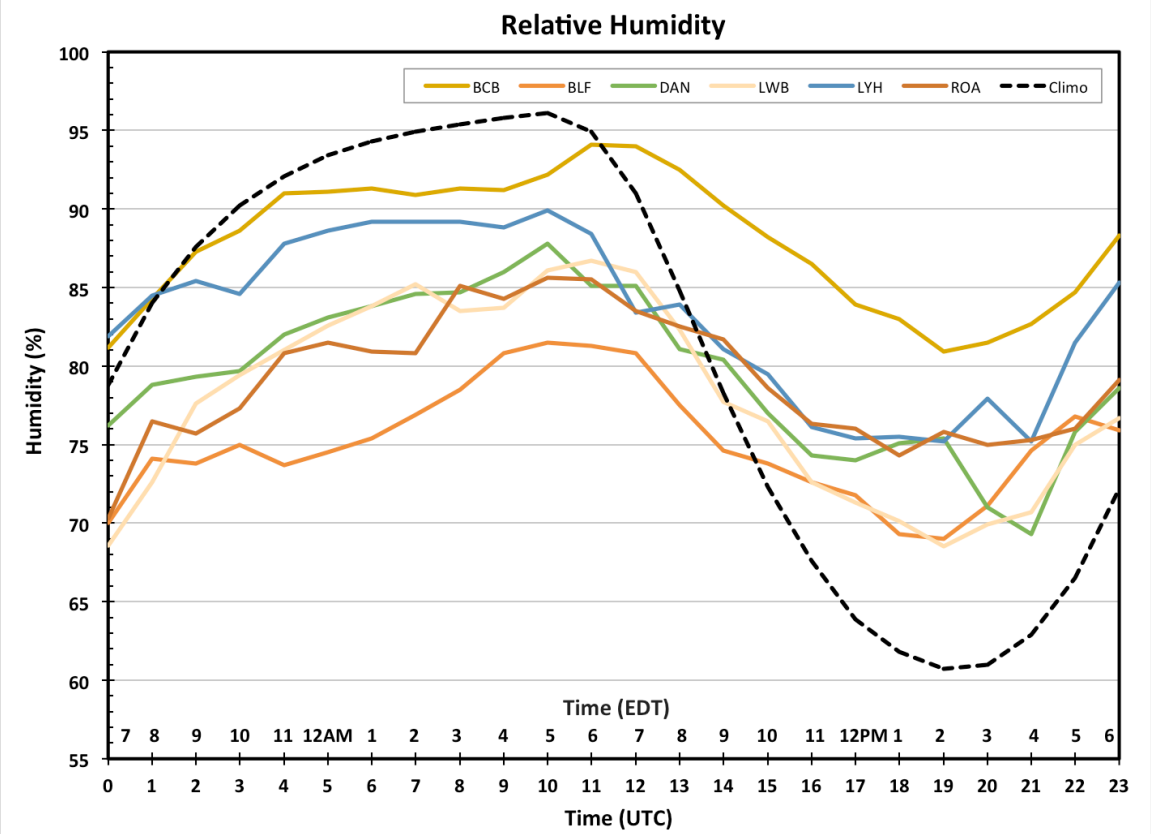


Figure 4.18. As as Figure 4.17, except for relative humidity (%).

Relative humidity levels at each TAF site are slightly lower relative to Blacksburg's climatologically normal values at night. The opposite occurs during the day, indicating a moist atmosphere that is near saturation, as is characteristic of CAD. Relatively lower humidity values at Bluefield could possibly be attributed to erosion processes ensuing at this site, reducing humidity values as a potential result of rising temperatures and dry air entraining into the area as the wedge erodes (Robert Stonefield – NWSFO Blacksburg, Personal communication). To this point, the majority of sites show a dominant easterly component to wind direction during problematic CAD events, with an exception at Bluefield, where consistent southwesterly flow is evident (Figure 4.19).

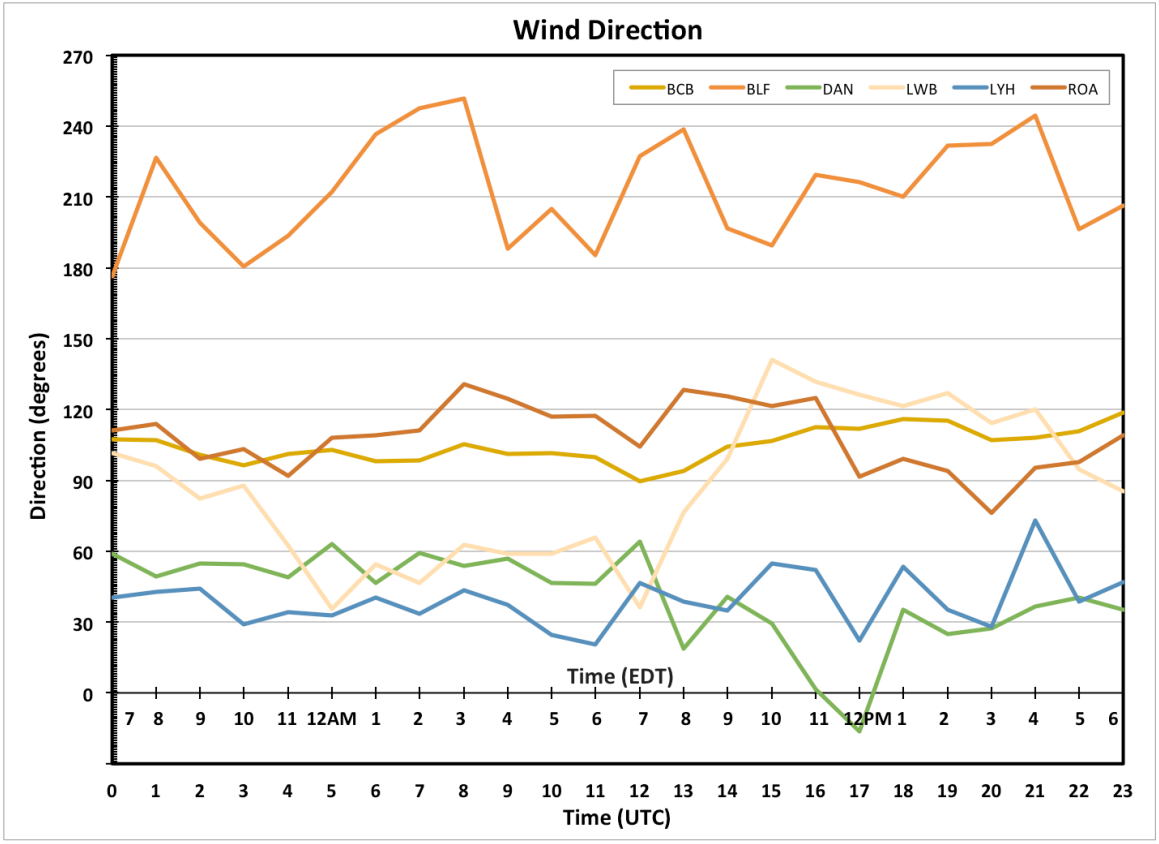


Figure 4.19. As in Figure 4.17, except for wind direction (degrees).

Danville and Lynchburg experience northeasterly flow, oftentimes as a result of a northeasterly cross-barrier jet that funnels dry air into the Blue Ridge (Robert Stonefield – NWSFO Blacksburg, Personal communication) typical of CAD scenarios. Lewisburg experiences a strong directional diurnal wind shift, indicative of anabatic/katabatic flow within the Greenbrier Valley. This is typical when particularly cold air is present (Jim Hudgins – NWSFO Blacksburg, Personal communication). Bluefield exhibits predominantly southwesterly surface winds; however, only 9 of the 37 warm season CAD busts occurred at this site with 3 transpiring during erosion processes. Winds at Bluefield may well have been from the typical easterly direction during the nine CAD busts, with the mean southwesterly wind reflecting conditions during the other 28 CAD busts within the CWA when Bluefield may have been outside the wedge.

Though cloud ceiling heights are variable throughout the day, warm season cloud ceilings descend by between 15 and 20 kilometers overnight and rise back during the day, hovering on average around 30 to 35 km agl. Fairly steady ceilings are produced by a narrow diurnal temperature range that results in limited change in atmospheric thickness throughout the day (Figure 4.20).

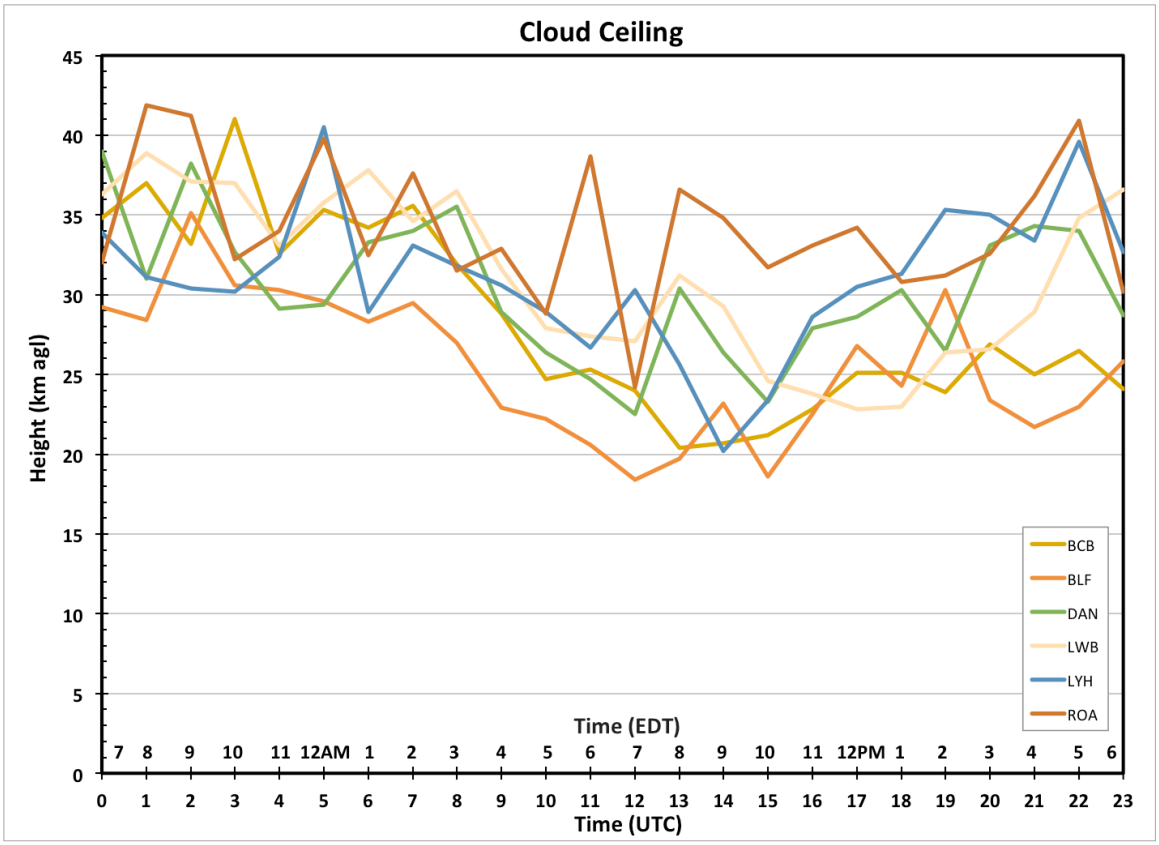


Figure 4.20. As in Figure 4.17, except for cloud ceiling heights (km agl).

Atmospheric conditions during warm season CAD events that produced busted forecasts appear conventional, though relatively higher temperatures and lower relative humidity values at Bluefield suggest these problematic cases are weak scenarios that do not influence all six TAF sites in the CWA. Atypical wind direction at Bluefield again suggests this to be the case. Though models struggle with summer wedge setups in general and overestimate wedge strength during the warm season, these cases may have been particularly weak, resulting in acute conditions that MAV and MET guidance failed to handle successfully.

4.6.2 Cold Season Hourly Surface Composites

Hourly composites generated from the 73 cases of problematic CAD during the cold season (November to April) between 2007 and 2016 exhibit a diurnal temperature range of 11°F or less across all TAF sites (Figure 4.21), less than the climatologically normal range at Blacksburg of 14°F.

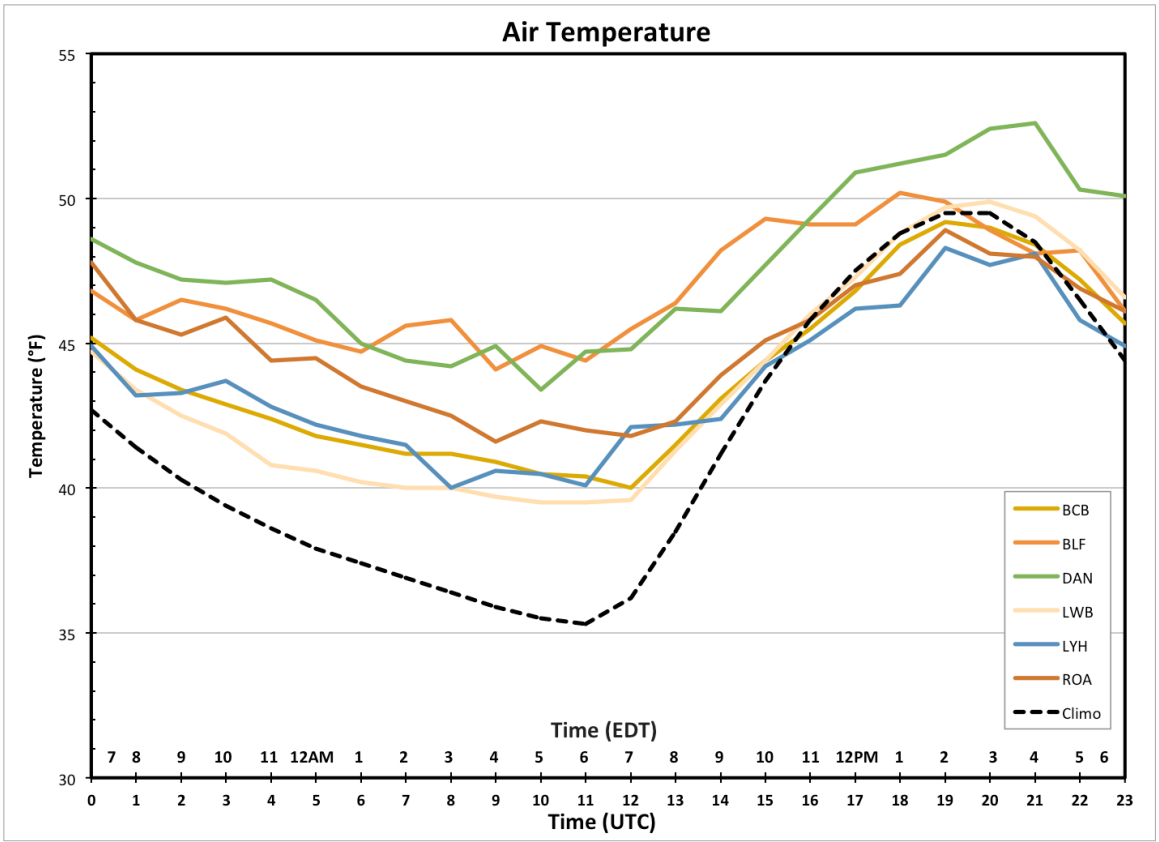


Figure 4.21. Hourly surface composites of air temperature (°F) during problematic cold season cold-air damming events at each of the six TAF sites (Blacksburg (BCB), Bluefield (BLF), Danville (DAN), Lewisburg (LWB), Lynchburg (LYH), and Roanoke (ROA)). For reference, the climatological normal temperature between November and April for the ten-year period at Blacksburg is shown as a dashed line.

Overnight temperatures at the TAF sites remain higher relative to normal values for Blacksburg by between 4°F and 9°F, likely as a product of increased cloud cover and humidity, while daytime temperatures stay near Blacksburg's climatological average at most TAF sites with an exception of higher temperatures at Danville. Bluefield

and Danville are often the last two sites in the CWA to experience the onset of an evolving wedge of cold air and the first to see it erode; both remain warmer than the other sites, once again suggesting the two locations are not always within the cold dome based on wedge strength.

Much like warm season relative humidity, values during cold season CAD busts increase overnight and fall during the day (Figure 4.22), following the typical diurnal pattern but in a conserved way.

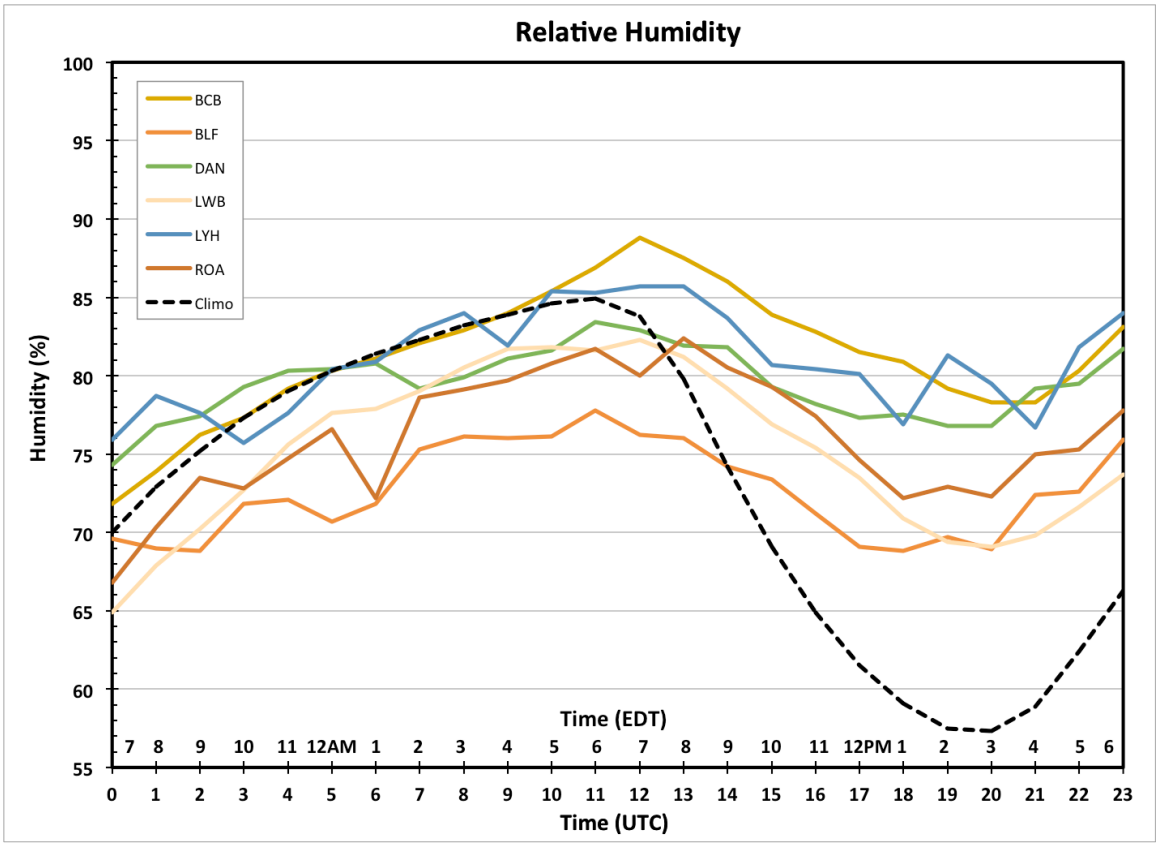


Figure 4.22. As in Figure 4.21, except for relative humidity (%).

Once again, relative humidity levels at each TAF site are similar to climatologically normal values at night and remain higher during the day, reflecting a moist, near-saturated atmosphere characteristic of Appalachian CAD.

Cool season wind direction at the six TAF sites is nearly identical to wind direction during the warm season with most exhibiting a dominant easterly component (Figure 4.23).

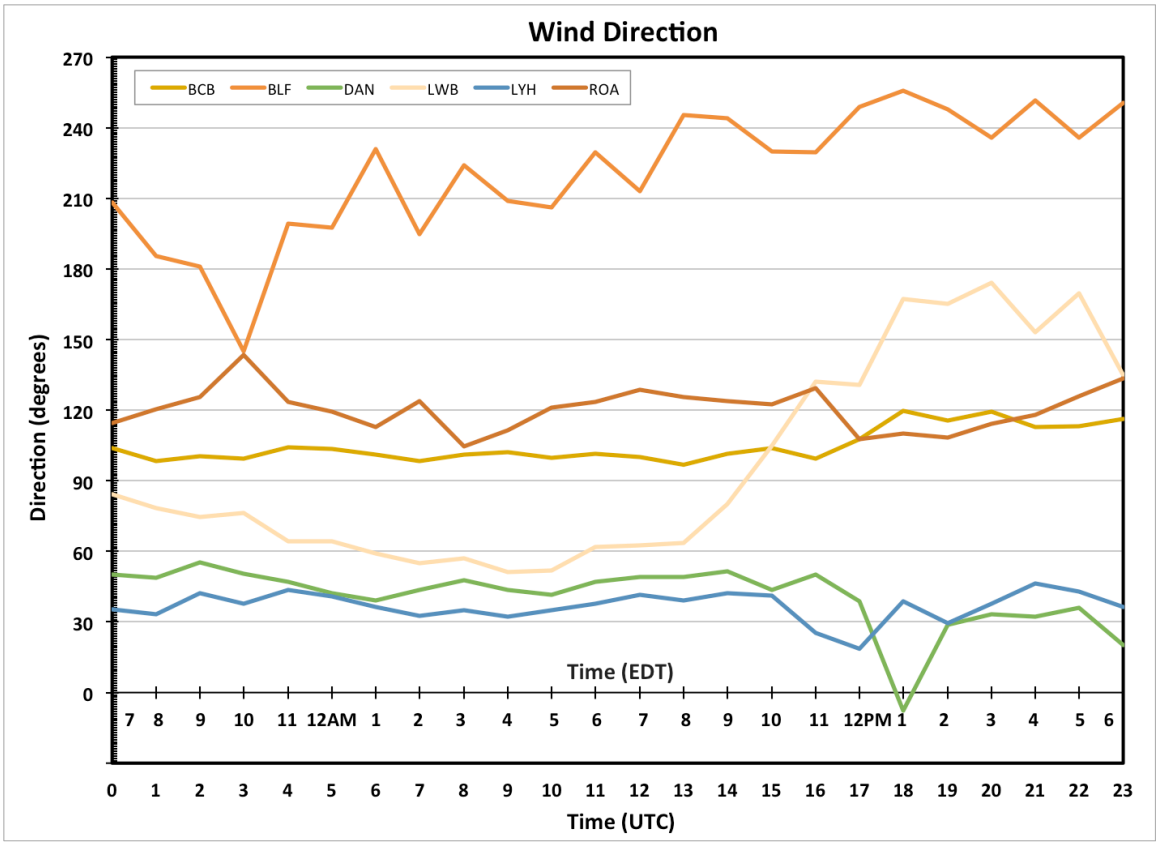


Figure 4.23. As in Figure 4.21, except for wind direction (degrees).

As with the warm season, Bluefield is characterized by a composite southwesterly wind flow, in which 27 of the 73 cold season CAD busts occurred at this site with 11 transpiring during erosion. Wind direction at Bluefield may potentially be skewed by the high frequency of erosion cases in the cold season, whereby Bluefield erodes out of the wedge before other locations, and by weak CAD occurrences over the region that may not encompass Bluefield. Winds at Lewisburg depict an

anabatic/katabatic flow regime, as in summer, indicative of a cold dome in place at that site.

Cold season cloud ceilings slowly drop by between 10 and 15 kilometers in the morning and rise again at night, remaining fairly steady around 25 km agl on average over the 24-hour period (Figure 4.24).

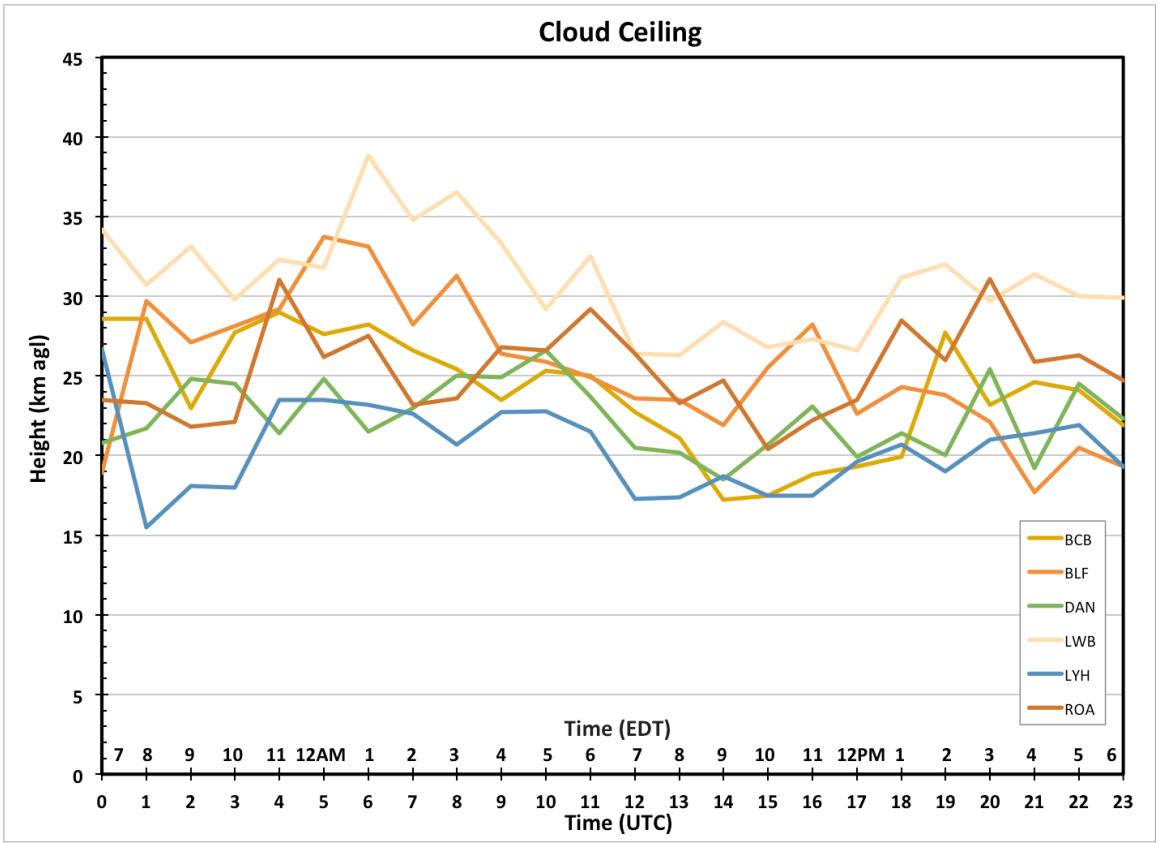


Figure 4.24. As in Figure 4.21, except for cloud ceiling heights (km agl).

Cold season ceilings are lower than during the warm season, resulting from a shallower boundary layer during colder months. Like the warm season, limited ceiling height change reflects the narrow range of diurnal temperatures typical of CAD cases while lower temperatures create a thinner atmosphere drawing cloud celilings closer to the ground.

As for the warm season, mean surface atmospheric conditions during cold season CAD events characterized by busted forecasts appear conventional, with relatively low temperature and high humidity values coupled with low ceilings and easterly winds at most TAF sites. Winds at Bluefield flow southwesterly, suggesting these problematic cases are weaker scenarios that do not encompass all six sites in the CWA. Weaker CAD situations are notably problematic for MOS guidance, and may pose difficulties causing forecast error.

4.7 Synoptic Composites

The large-scale atmospheric characteristics of the problematic CAD events are examined during both warm and cold season scenarios using composites of the synoptic atmosphere. Composited variables include mean sea level pressure, air temperature, relative humidity, and dew point temperature (not shown) at the surface; geopotential heights, air temperature, and wind speed and direction at the 925 and 850 millibar pressure levels; and geopotential height at the 500 millibar pressure level (not shown). Composites from the days in which warm season classic and cold season in-situ CAD forecasts busted are not shown due to the limited number of dates within each season. Examining the nuances within synoptic composites of the different types of problematic CAD cases may provide insight as to why these cases were difficult to successfully forecast.

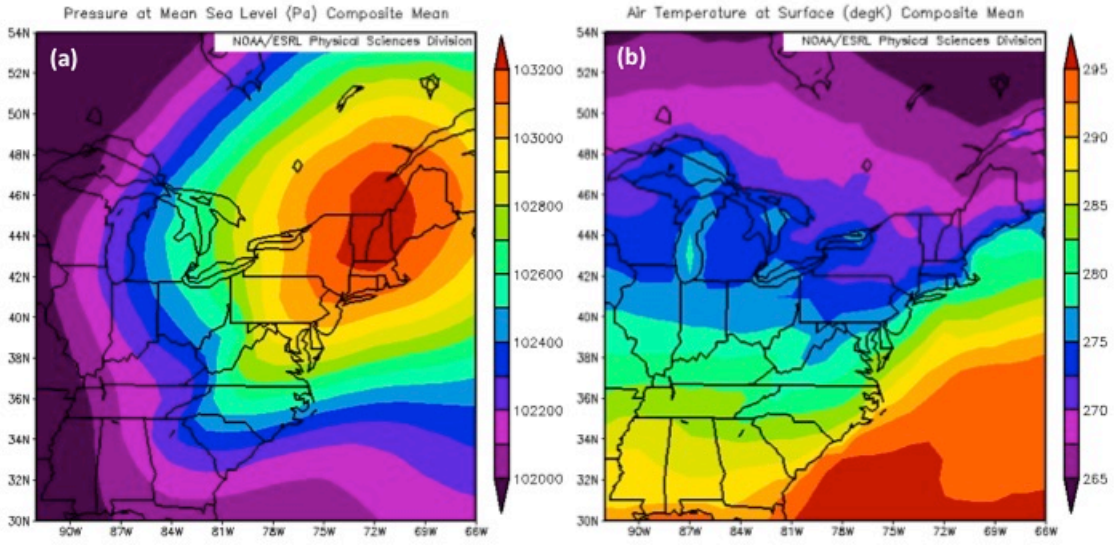
Next, synoptic composites of problematic classic CAD are then used as a standard, as they are generally the purest CAD scenarios synoptically, to generate difference composites between classic and other wedge classifications. By subtracting atmospheric composite values of a specific variable for each

classification from classic composites, disparities between the two scenarios are highlighted in subsequent difference maps of surface temperature (K) and sea level pressure (mb).

Lastly, composites for days 1 through 3 in advance of poorly forecast CAD days (-24 hours to -72 hours) are generated to assess whether these events initialized rather suddenly, leaving models little time to predict the approaching cold dome, or if models generally had sufficient time to anticipate the approaching wedge episodes.

4.7.1 Synoptic Composites of Cold-air Damming Classifications

For the 14 days of cold season classic CAD on which forecasts busted, the composite sea level pressure field reveals wedging along the Appalachians (Figure 4.25a), as a strong (1032 millibar) parent anticyclone positioned in the nearly-saturated (Figure 4.25c) central Northeast drives cold air southward along the eastern side of the mountains (Figure 4.25b).



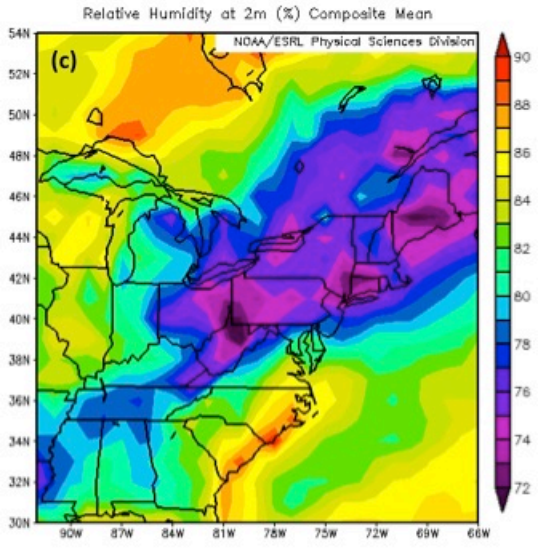
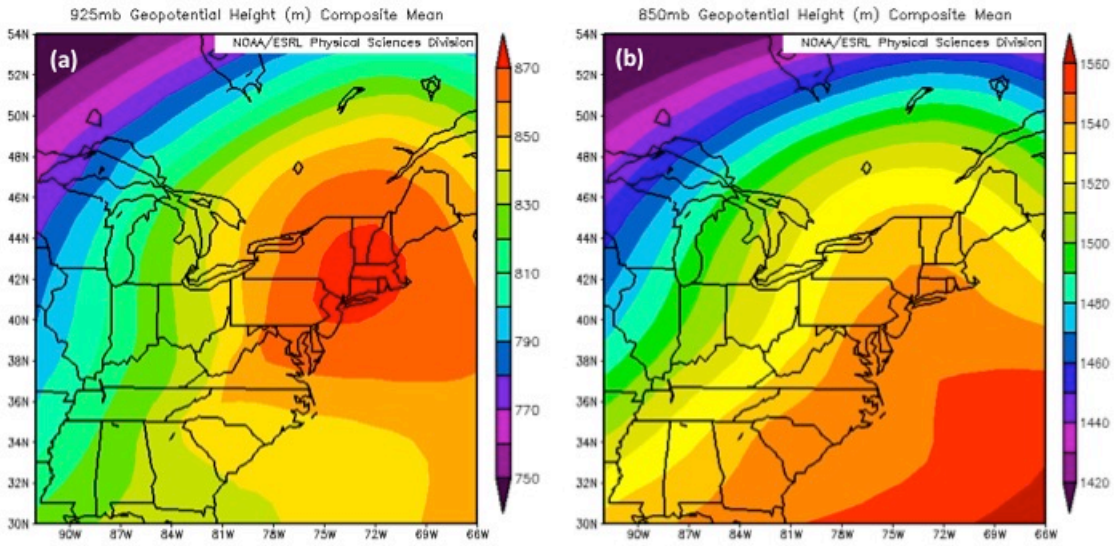


Figure 4.25a-c. Synoptic composites from the 14 days of cold season classic CAD that resulted in busted forecasts. Shown are (a) sea level pressure (Pa), (b) surface air temperature (K), and (c) 2-meter relative humidity (%).

At the 925 millibar pressure level, winds veer easterly (Figure 4.26e) and the evidence of wedging weakens along the mountains (Figure 4.26a, Figure 4.26c).



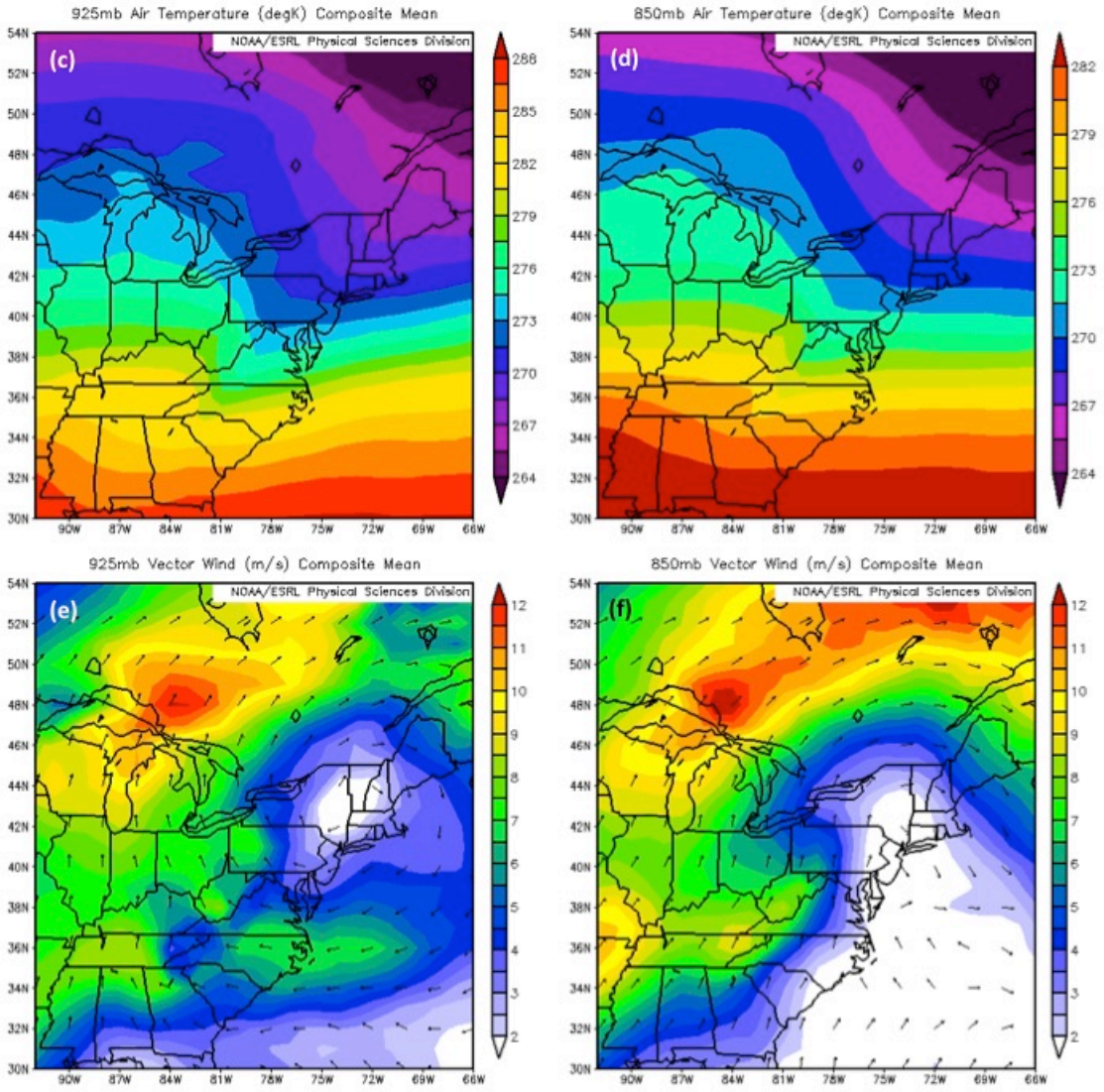


Figure 4.26a-f. Synoptic composites from the 14 days of cold season classic CAD that resulted in busted forecasts. Shown are (a) 925mb and (b) 850mb geopotential height (m), (c) 925mb and (d) 850mb air temperature (K), and (e) 925mb and (f) 850mb vector winds (m/s).

The wedge signature nearly disappears at the 850 millibar level (Figure 4.26b, Figure 4.26d) and southerly geostrophic winds dominate atmospheric flow (Figure 4.26f). Zonal flow at 500 millibars (not shown) further highlights the shallow nature of this phenomenon.

Composites for the 12 hybrid CAD days on which forecasts busted during the warm season reveal a similar pattern to that of problematic classic CAD, showing weaker pressure (Figure 4.27a) and temperature (Figure 4.27b) gradients but similar positioning.

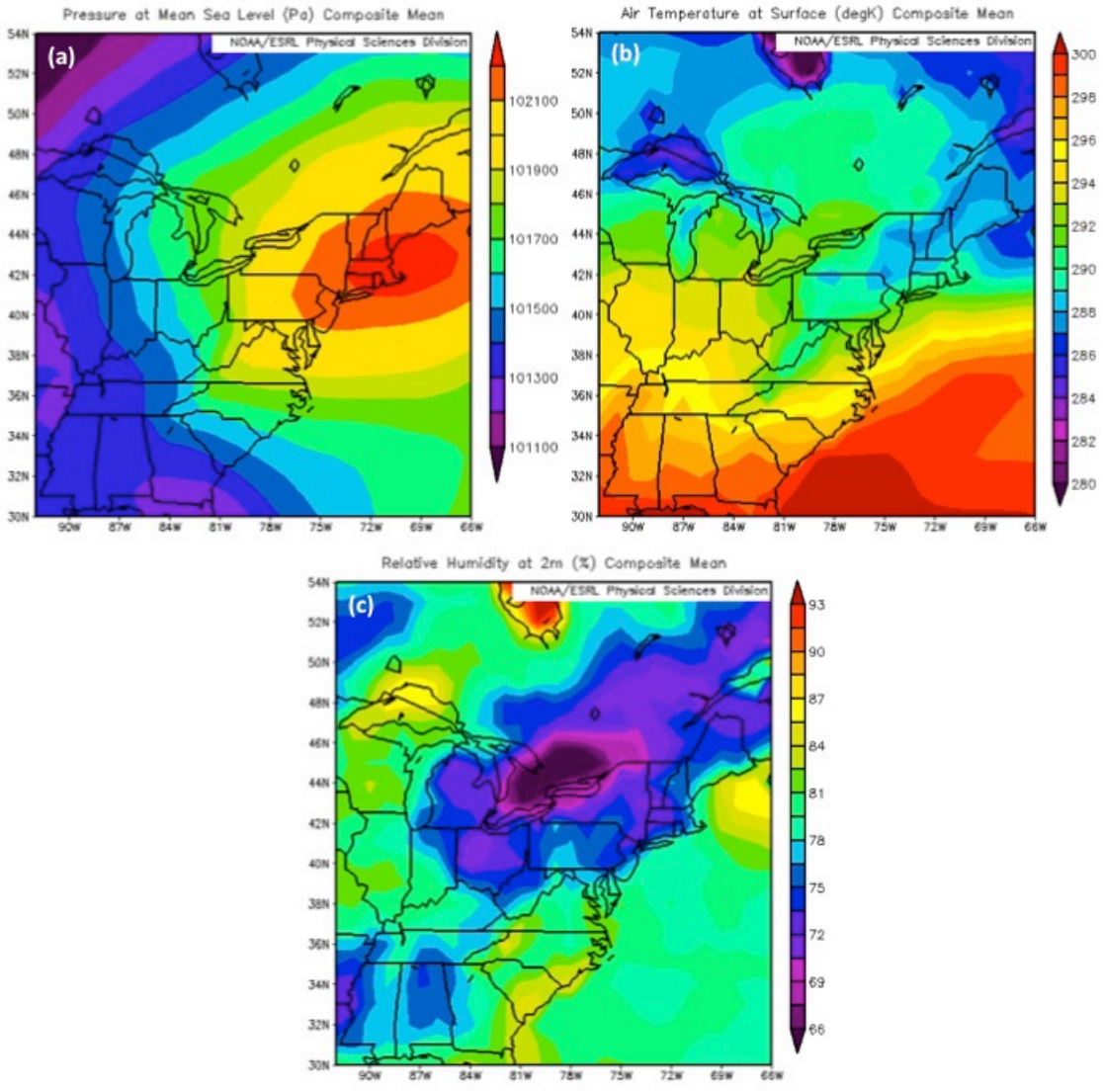
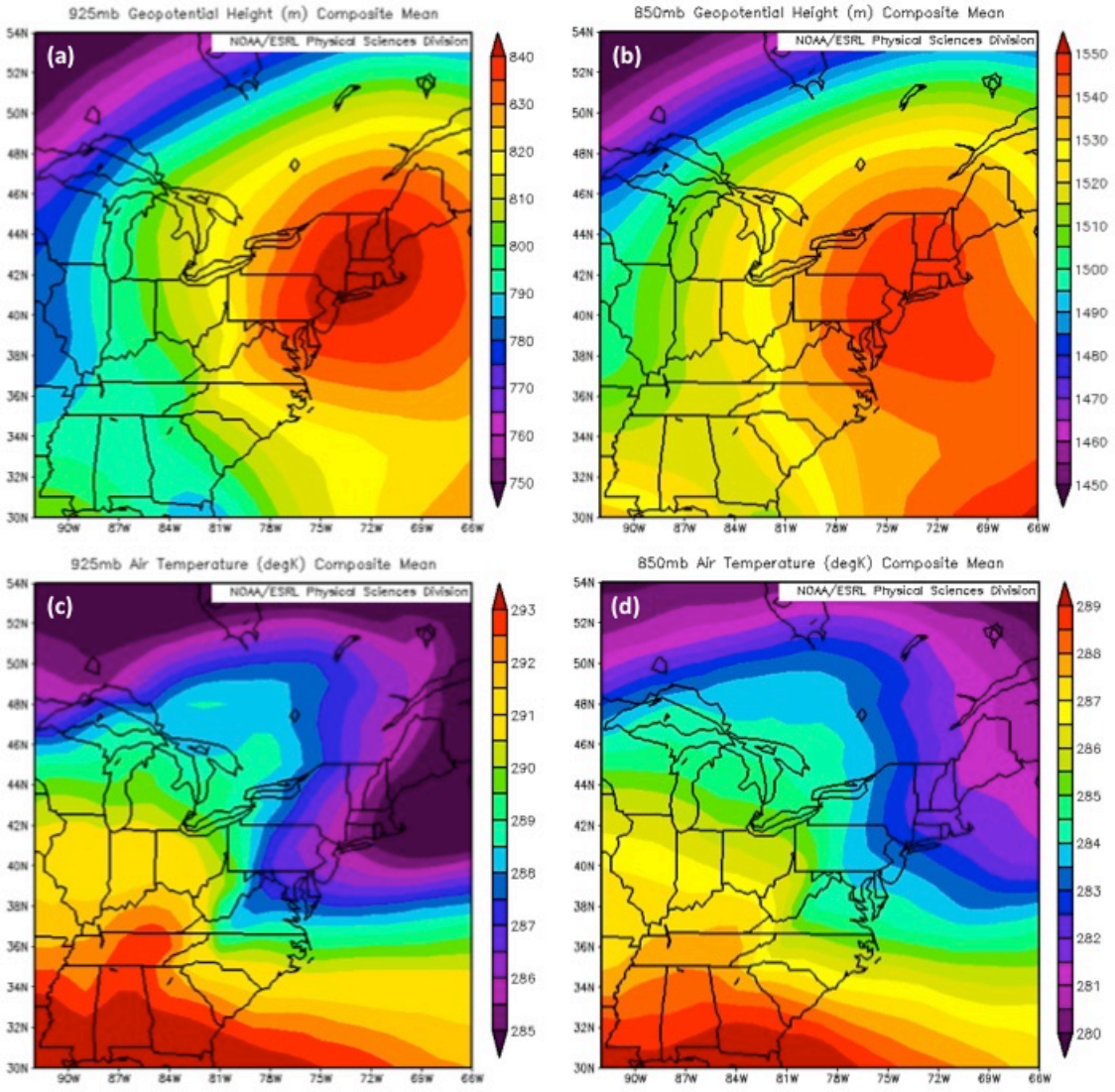


Figure 4.27a-c. As in 4.25a-f, except for the 12 days of problematic warm season hybrid CAD cases.

A weaker (1021 millibar) anticyclone at the surface is associated with less temperature wedging along the mountains. Higher relative humidity (Figure 4.27c)

near the surface denotes more saturated conditions than clasic setups. Pressure and temperature wedging at the 925 millibar level lessen along the eastern slopes of the mountains (Figure 4.28a, Figure 4.28c) among easterly wind flow (Figure 4.28e). The wedge signature remains at 850 millibars and is evident in weak wedging (Figure 4.28b, Figure 4.28d) despite southerly winds over the region (Figure 4.28f).



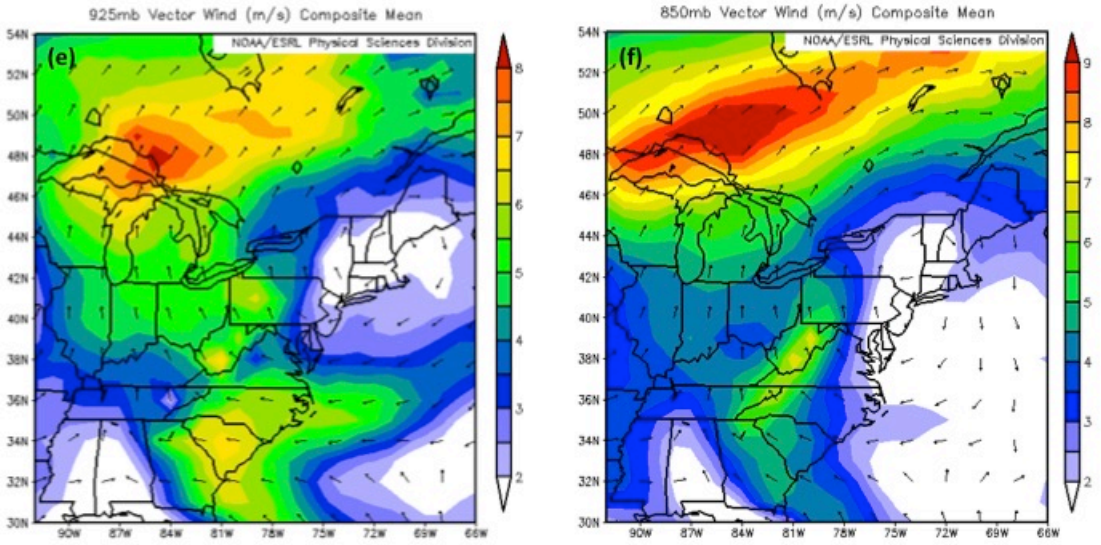
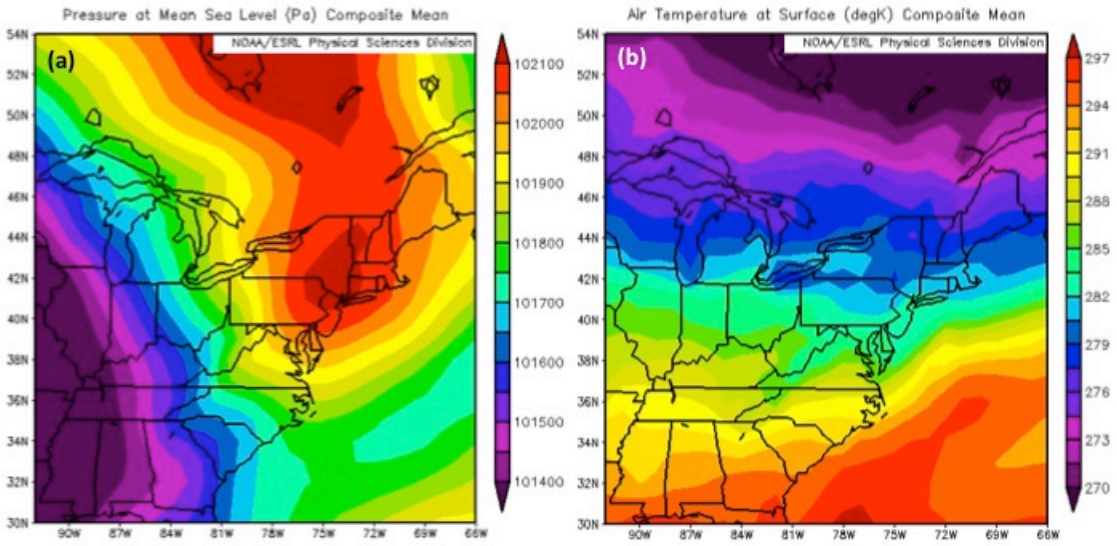


Figure 4.28a-f. As in Figure 4.26a-f, except for the 12 problematic warm season hybrid CAD scenarios.

Synoptic composites for the 6 days of problematic cold season hybrid CAD reveal a similar surface pressure pattern to that of classic busts, though the parent high pressure center over the northeastern United States branches out into eastern Canada (Figure 4.29a).



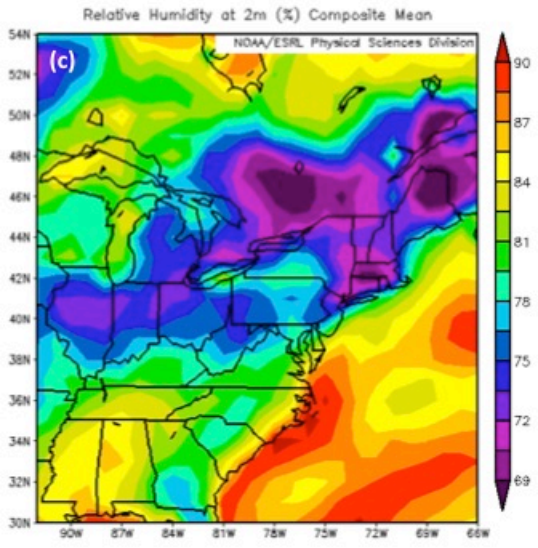
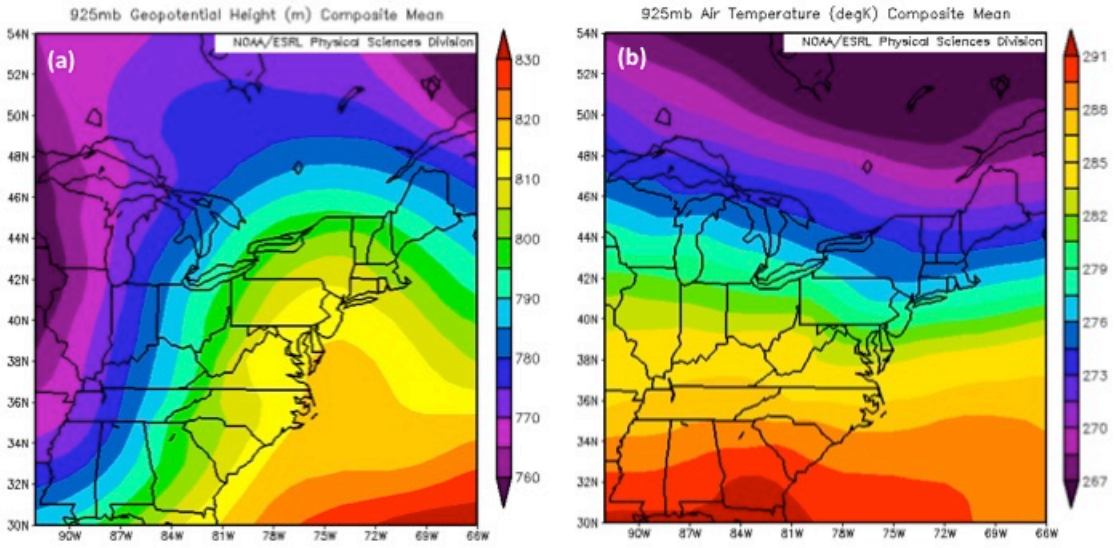


Figure 4.29a-c. As in 4.25a-f, except for the 7 days of problematic cold season hybrid CAD cases.

Apart from a broadened anticyclone, temperature wedging (Figure 4.29b) and relative humidity values (Figure 4.29c) at the surface of cold season hybrid busts remain consistent to that of warm season hybrid CAD bust composites and are similar to cold season classic problematic CAD composites.



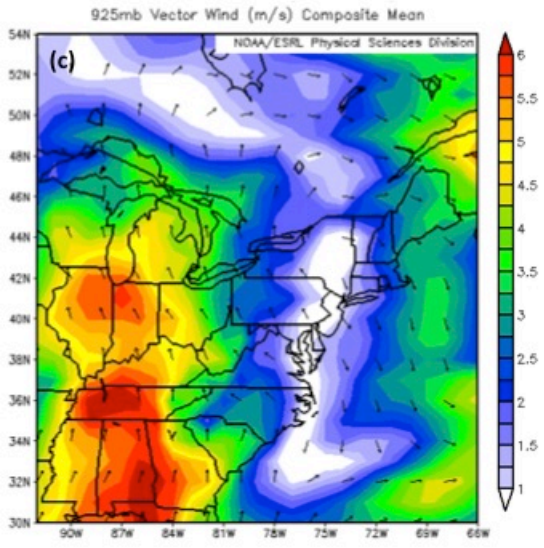


Figure 4.30a-c. Synoptic composites from the 14 days of cold season classic CAD that resulted in busted forecasts. Shown are (a) 925mb geopotential height (m), (b) 925mb air temperature (K), and (c) 925mb vector winds (m/s).

The wedge signature of cold season hybrid CAD busts dissolves by 925 millibars as a pressure ridge (Figure 4.30a) over the northeast directs southeasterly wind flow over the central Appalachians (Figure 4.30c). Temperature wedging appears to diminish at 925 millibars (Figure 4.30b), suggesting the hybrid CAD scenarios during the cold season in which forecasters struggled were marginal cases of the phenomenon.

Synoptic composites for the 7 days of problematic in-situ CAD during the warm season reveal a broader parent anticyclone (Figure 4.31a) positioned lower over the mid-Atlantic with a central strength averaging 1024 millibars.

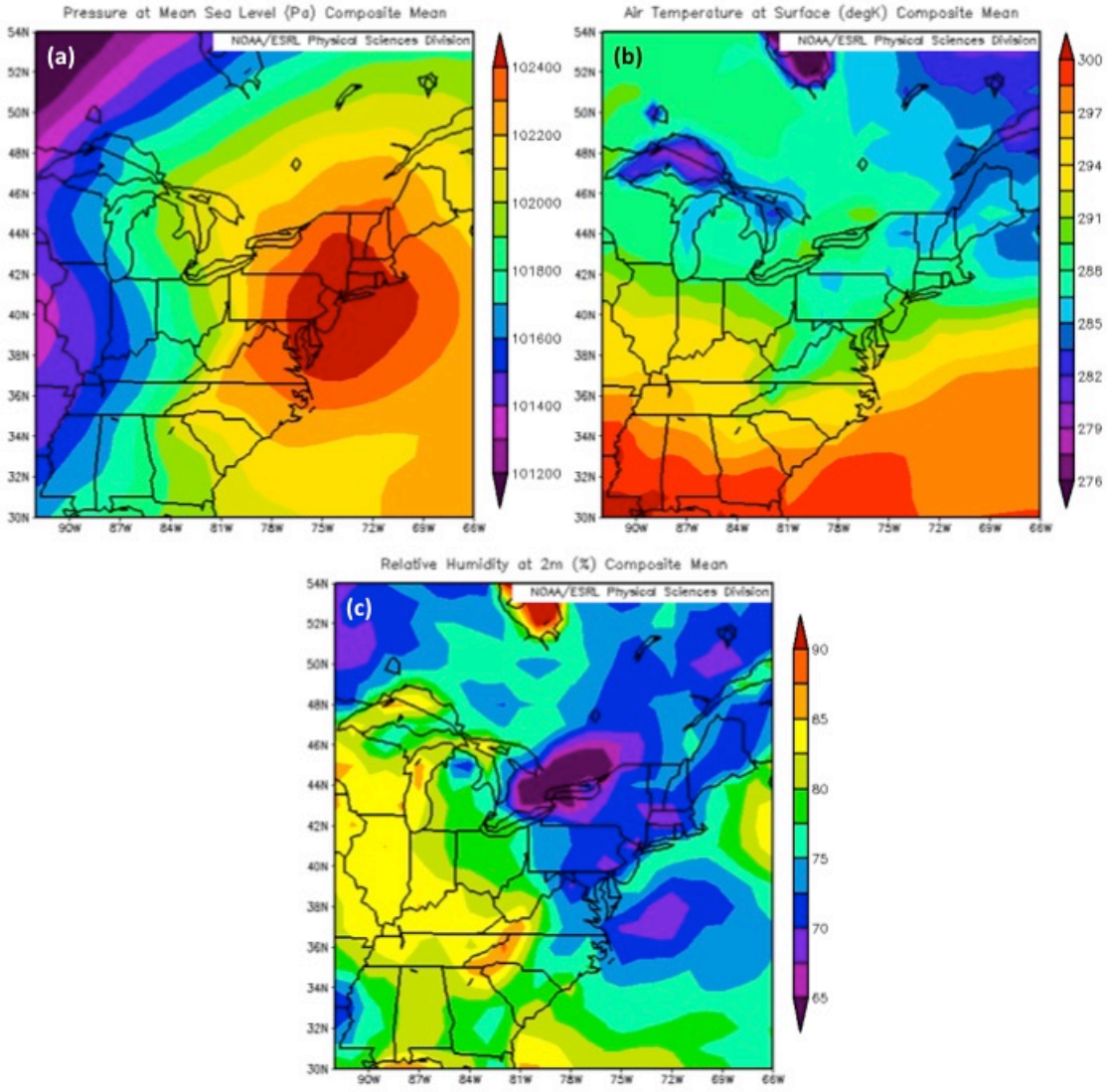


Figure 4.31a-c. As in 4.25a-c, except for the 7 days of problematic warm season in-situ CAD scenarios.

Surface temperature wedging is evident from Pennsylvania southward through western North Carolina (Figure 4.31c). Relative humidity values (Figure 4.31c) are the highest among the three CAD classifications, suggesting the highest saturation levels among CAD classes are found in in-situ scenarios. Southeasterly winds dictate flow at 925 millibars (Figure 4.32c), as damming signatures at this level dissipate

(Figure 4.32a, Figure 4.32b); this reflects the extremely shallow nature of the problematic in-situ cases.

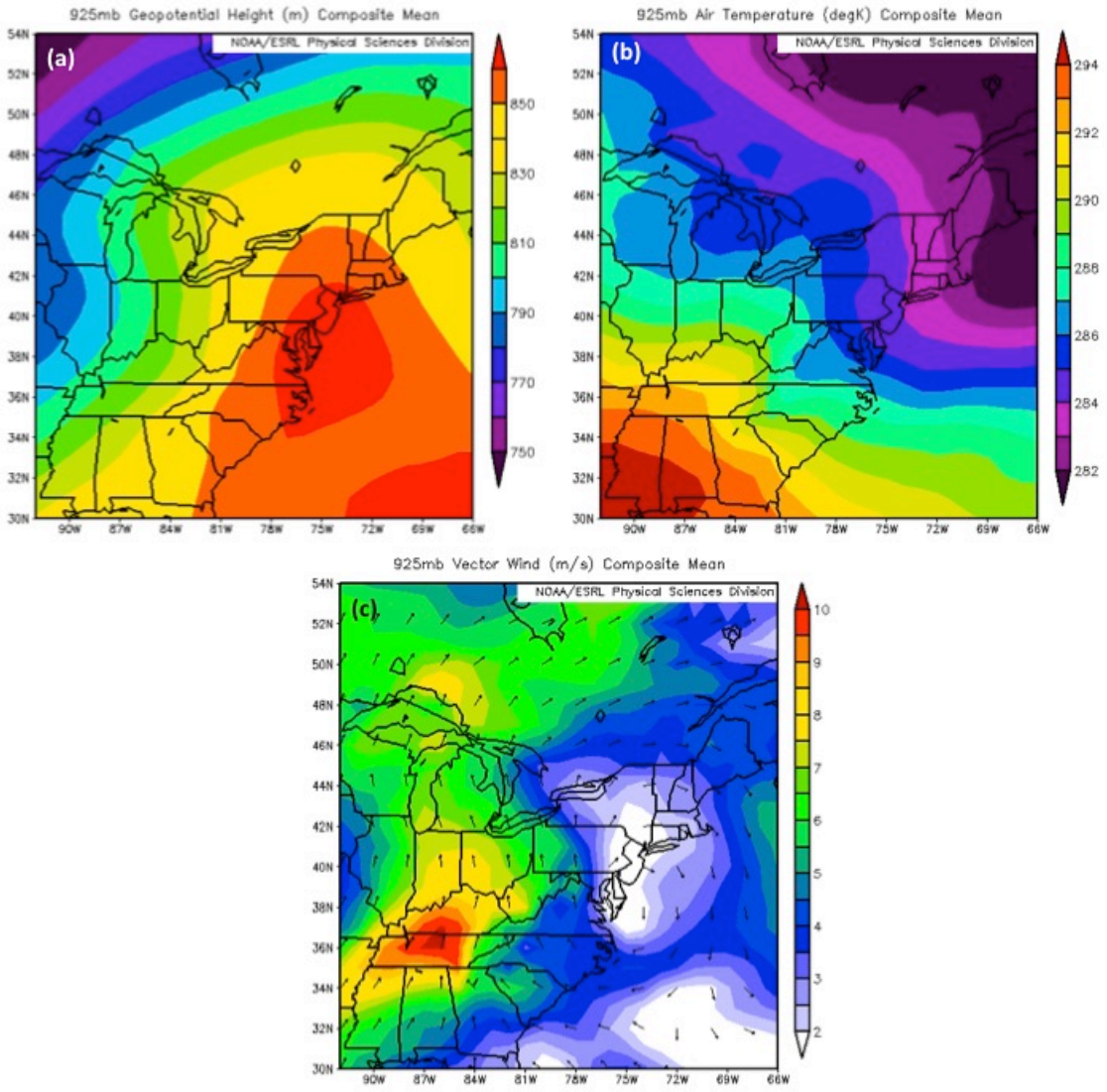


Figure 4.32a-c. As in Figure 4.30a-c, except for the 7 warm season in-situ problematic CAD scenarios.

4.7.2 Synoptic Composites of Cold-air Damming Onset

Synoptic composites generated from the 11 days on which busted cold season forecasts were associated with CAD onset lack strong wedging along the Appalachian Mountains (Figure 4.33a, Figure 4.33b), implying these composites

occur early-on in the onset process. Furthermore, high pressure at the surface is centered much farther to the west than the above classifications, confirming that the days classified as onset occur during the initialization of CAD. Weak evidence of shallow pressure wedging initializing at the surface suggest northerly surface winds have not yet begun to dam against the mountains.

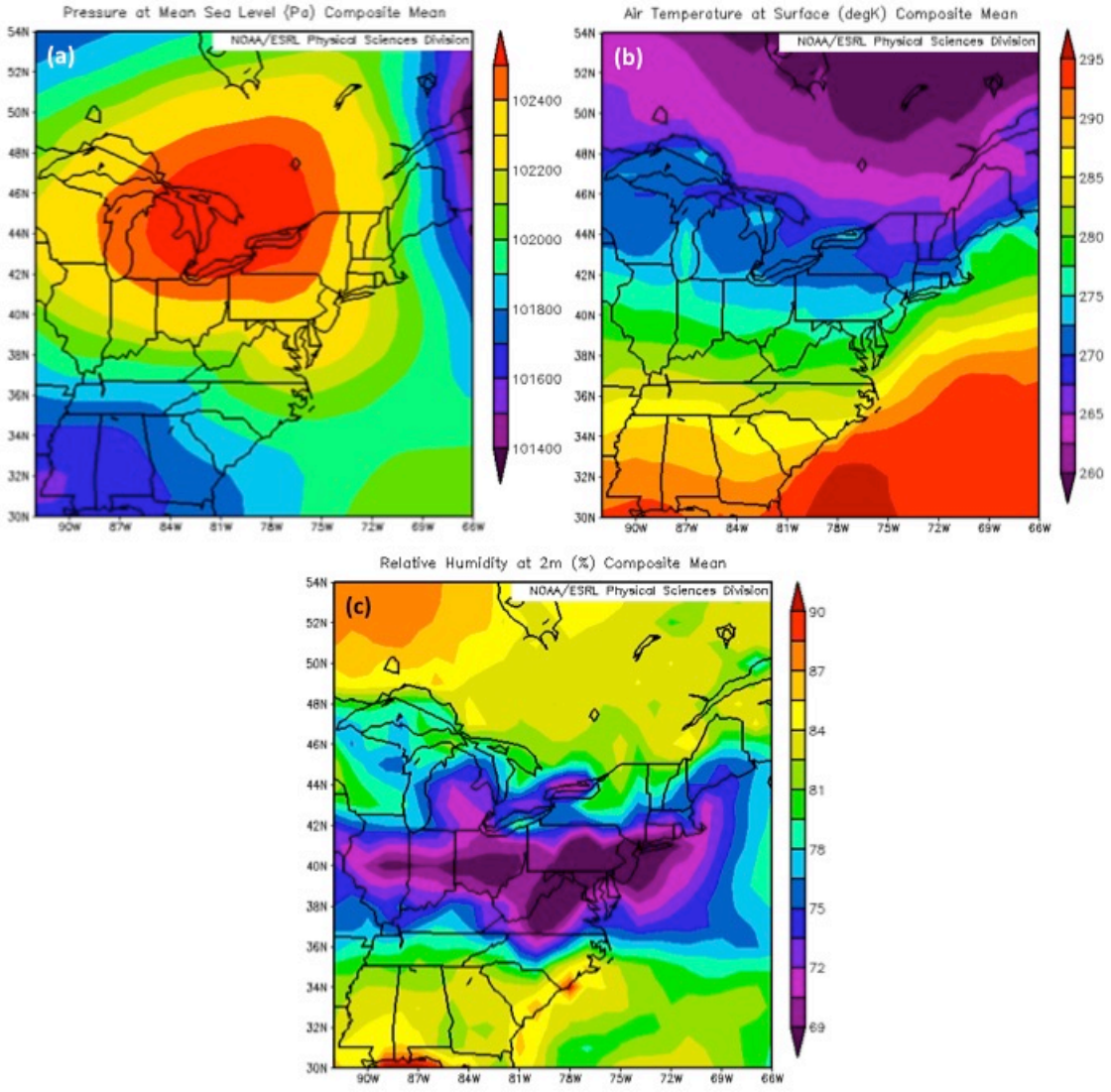


Figure 4.33a-c. Synoptic composites from the 11 days of cold season CAD onset that resulted in busted forecasts during the cold season. Shown are (a) sea level pressure (Pa), (b) surface air temperature (K), and (c) 2-meter relative humidity (%).

A pressure ridge forms over the region at 925 millibars (Figure 4.34a) while temperature wedging along the eastern Appalachian slopes is not yet evident (Figure 4.34b), and clockwise rotating winds at the 925 (Figure 4.34c) and 850 millibar levels (not shown) reinforce the presence of an anticyclone propogating towards the northeast.

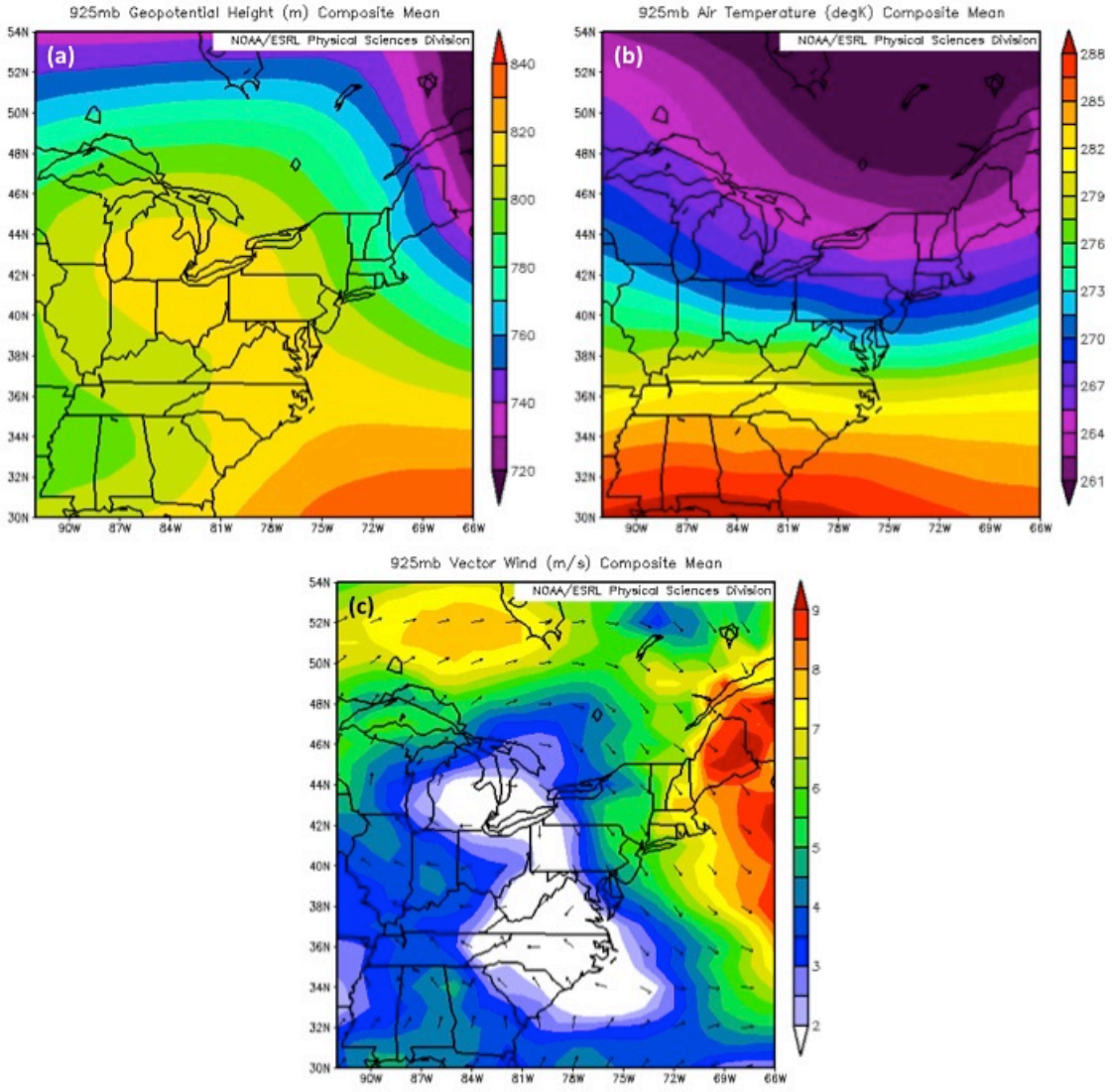


Figure 4.34a-c. Synoptic composites from the 11 days of CAD onset that resulted in busted forecasts during the cold season. Shown are (a) 925mb geopotential height (m), (b) 925mb air temperature (K), and (c) 925mb vector winds (m/s).

Synoptic composites from the 7 days of CAD onset that resulted in busted forecasts in the warm season, although not shown, exhibit more of the same characteristics typical of CAD onset.

4.7.3 Synoptic Composites of Cold-air Damming Erosion

Surface composites generated from the 40 days on which busted forecasts were associated with the erosion of CAD during the cold season show a strong pressure gradient ahead of an approaching low pressure center (Figure 4.35a). This advection pattern (Figure 4.35b) verifies CAD erosion during these days.

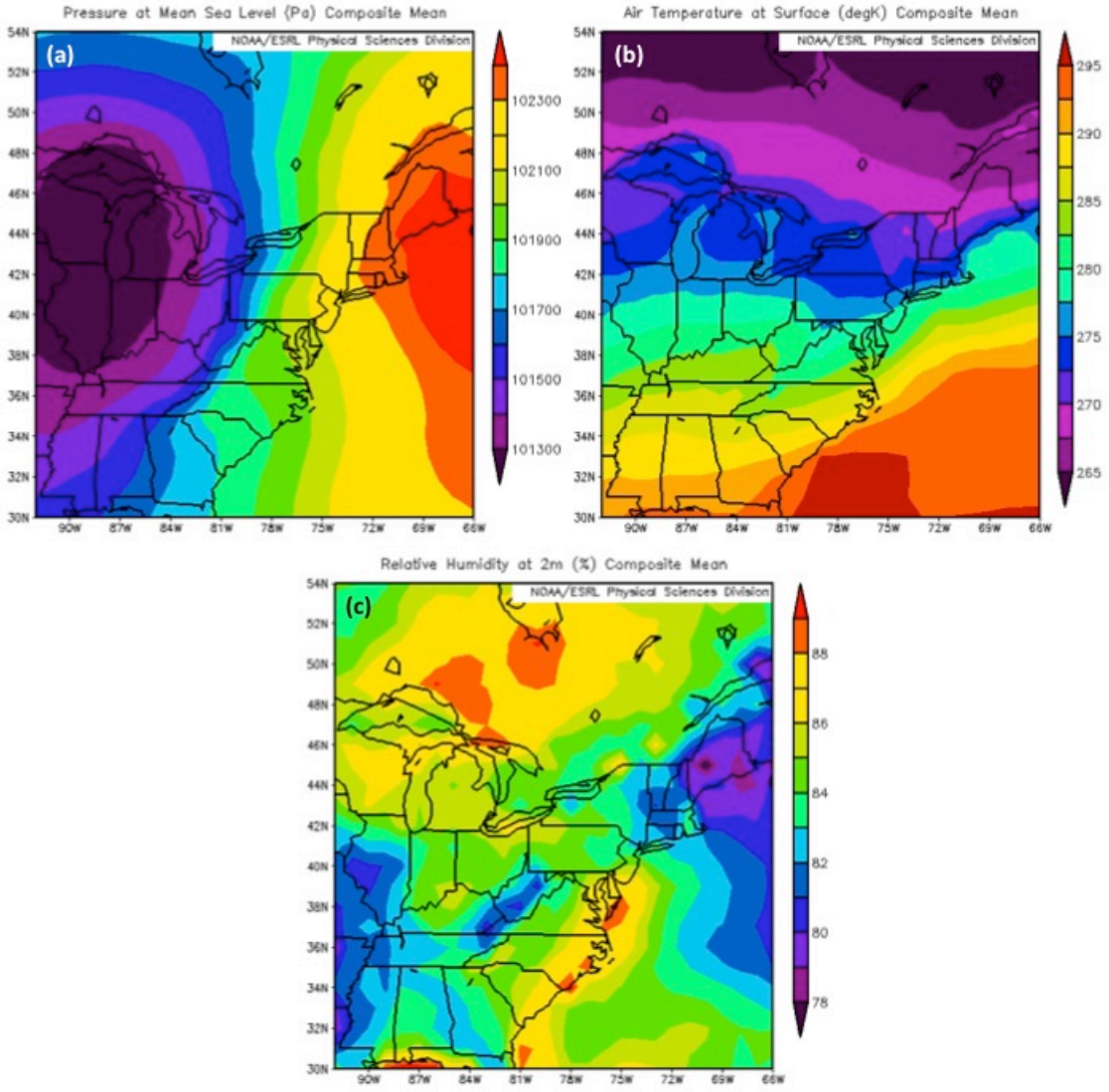


Figure 4.35a-c. Synoptic composites from the 40 days of cold season CAD erosion that resulted in busted forecasts. Shown are (a) sea level pressure (Pa), (b) surface air temperature (K), and (c) 2-meter relative humidity (%).

Relatively strong southerly flow dictated by a northeast-southwest pressure gradient (Figure 4.36a, Figure 4.36c) lacks evidence of CAD at 925 millibars; however, a weak residual cold pool persists along the mountains at this level (Figure 4.36c) yet dissolves by 850 millibars (not shown).

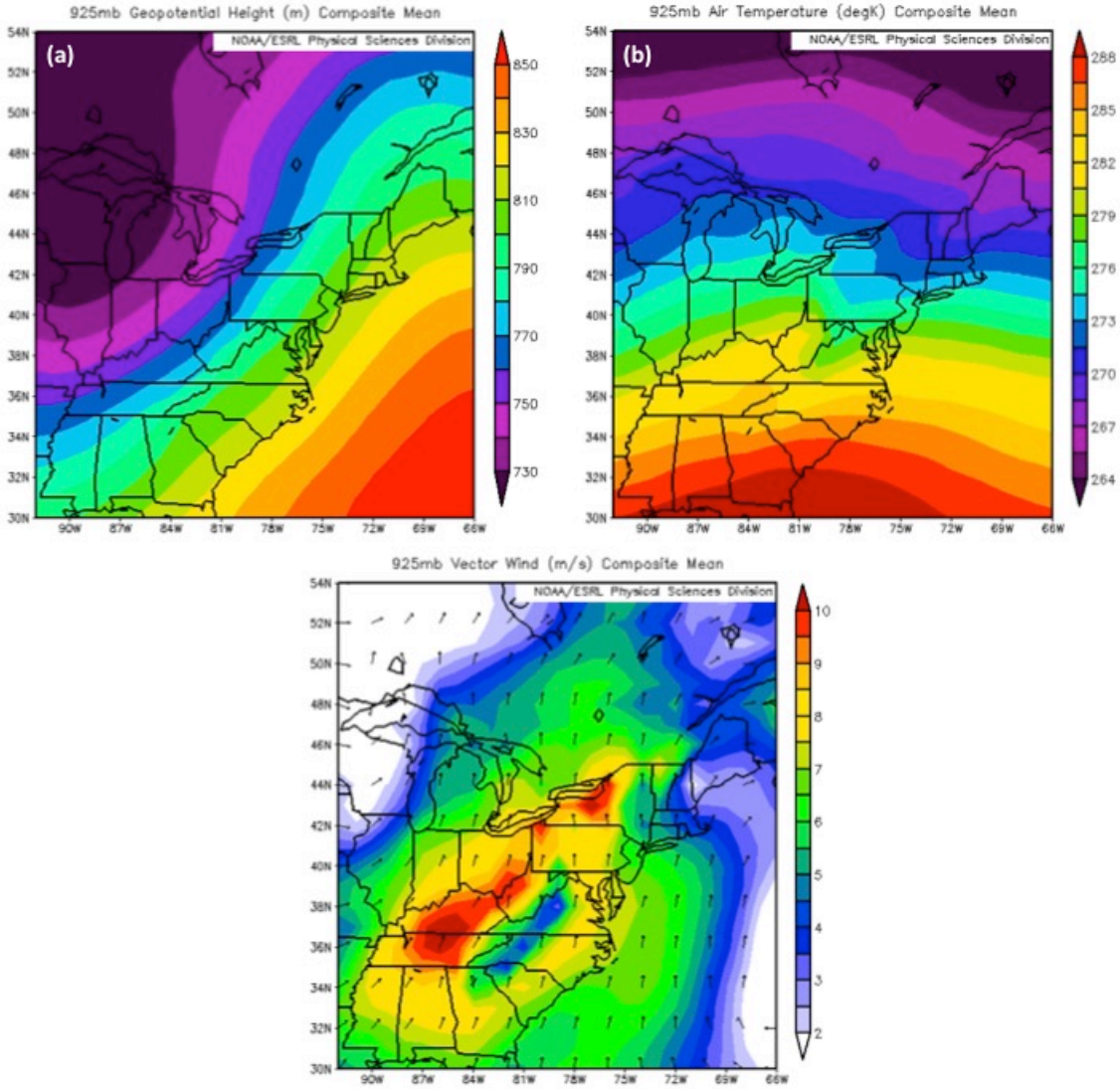


Figure 4.36a-c. Synoptic composites from the 40 days of cold season CAD erosion that resulted in busted forecasts. Shown are (a) 925mb geopotential height (m), (b) 925mb air temperature (K), and (c) 925mb vector winds (m/s).

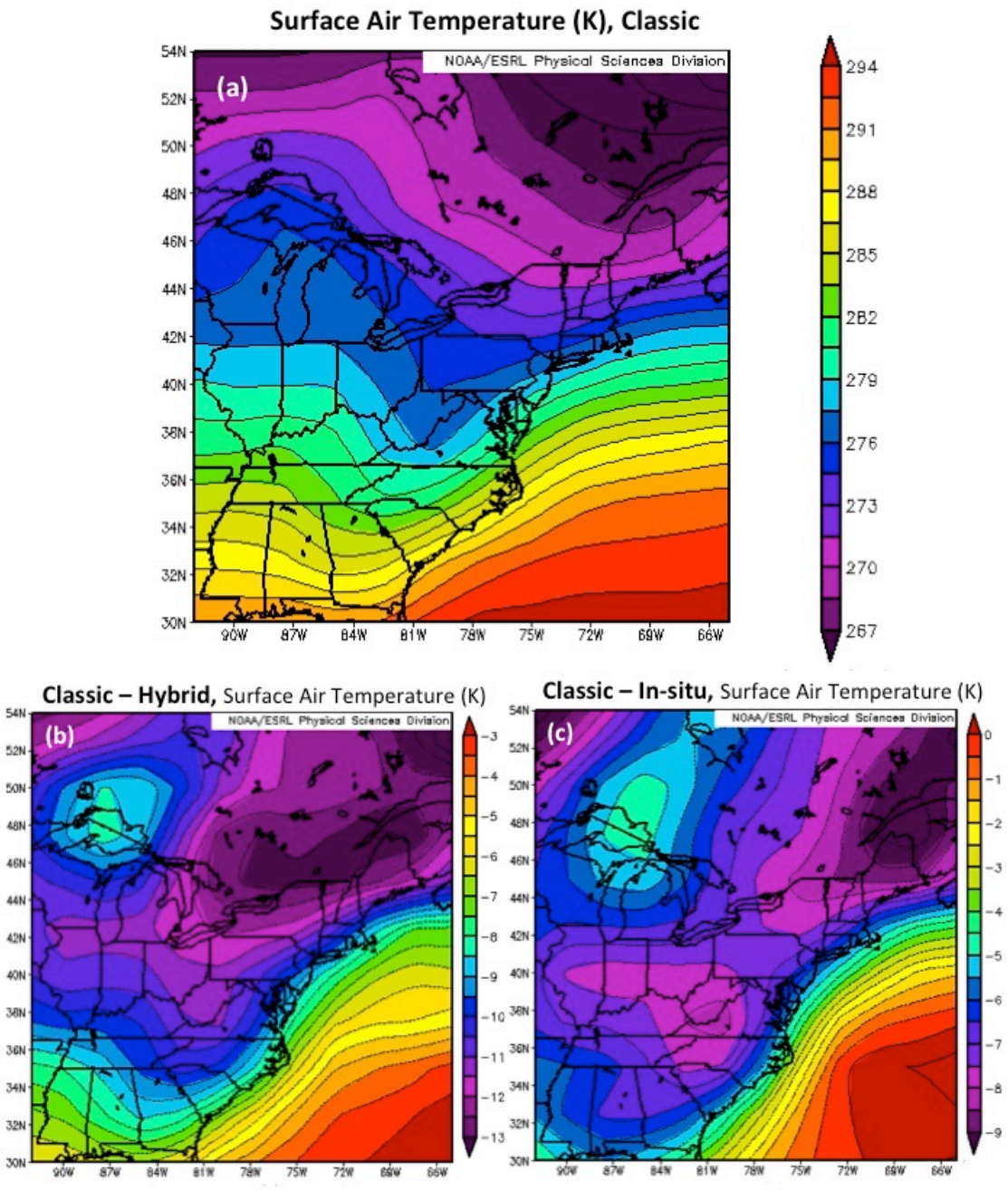
This suggests a common top-down erosion pattern during the busted CAD events. Synoptic composites from the 7 days of CAD erosion that resulted in busted forecasts in the warm season, although again not shown, exhibit more of the same characteristics typical of CAD erosion.

4.7.4 Summary

The composites within each classification are consistent with synoptic set-ups described in previous literature including Ellis et al. (2017), Bailey et al. (2003), and Bell and Bosart (1988), without any major differences. Meanwhile, synoptic composites of the onset of problematic CAD reflect those studied by Bailey et al. (2003) and Bell and Bosart (1988), and the erosion composites in this study are consistent with cold-frontal passage and northwestern low erosion mechanisms studied by Stanton (2003). Overall, while the composites seem to confirm the assigned CAD type (classic, hybrid, in-situ, onset, erosion), they lack obvious clues as to why these problematic CAD events would have been particularly difficult to forecast at a synoptic scale.

4.7.5 Synoptic Difference Composites Based on Classic Scenarios

Differences in surface air temperature for each CAD classification relative to classic CAD indicate warmer conditions across the CWA, suggesting classic CAD as the strongest wedge type among the population of problematic CAD forecast scenarios.


Surface Air Temperature (K), Classic
NOAA/ESRL Physical Sciences Division
(a)
54N
52N
50N
48N
46N
44N
42N
40N
38N
36N
34N
32N
30N
90W
87W
84W
81W
78W
75W
72W
69W
66W
294
291
288
285
282
279
276
273
270
267
Classic – Hybrid, Surface Air Temperature (K)
NOAA/ESRL Physical Sciences Division
(b)
−3
−4
−5
−6
−7
−8
−9
−10
−11
−12
−13
Classic – In-situ, Surface Air Temperature (K)
NOAA/ESRL Physical Sciences Division
(c)
0
−1
−2
−3
−4
−5
−6
−7
−8
−9

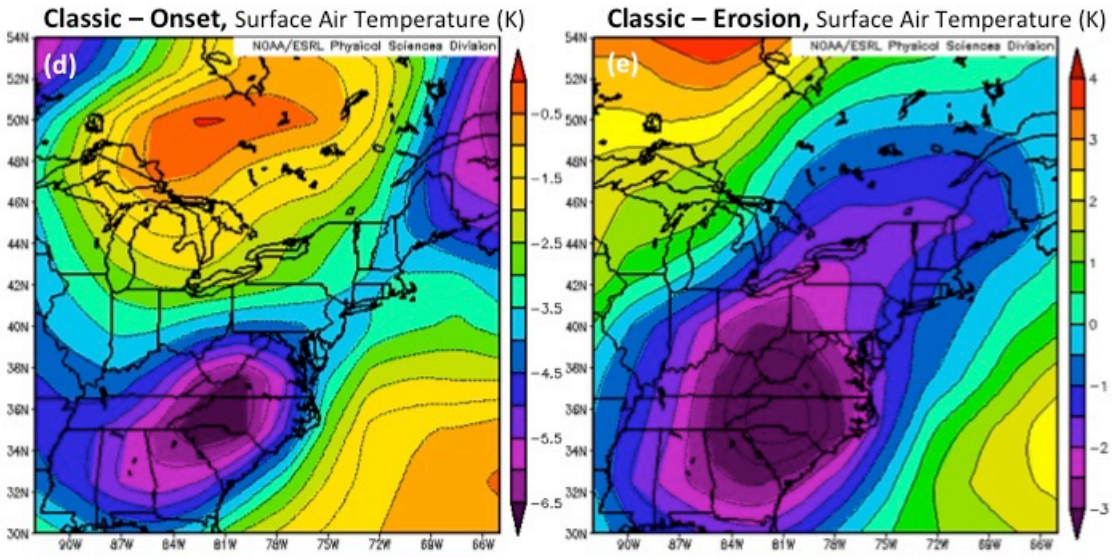
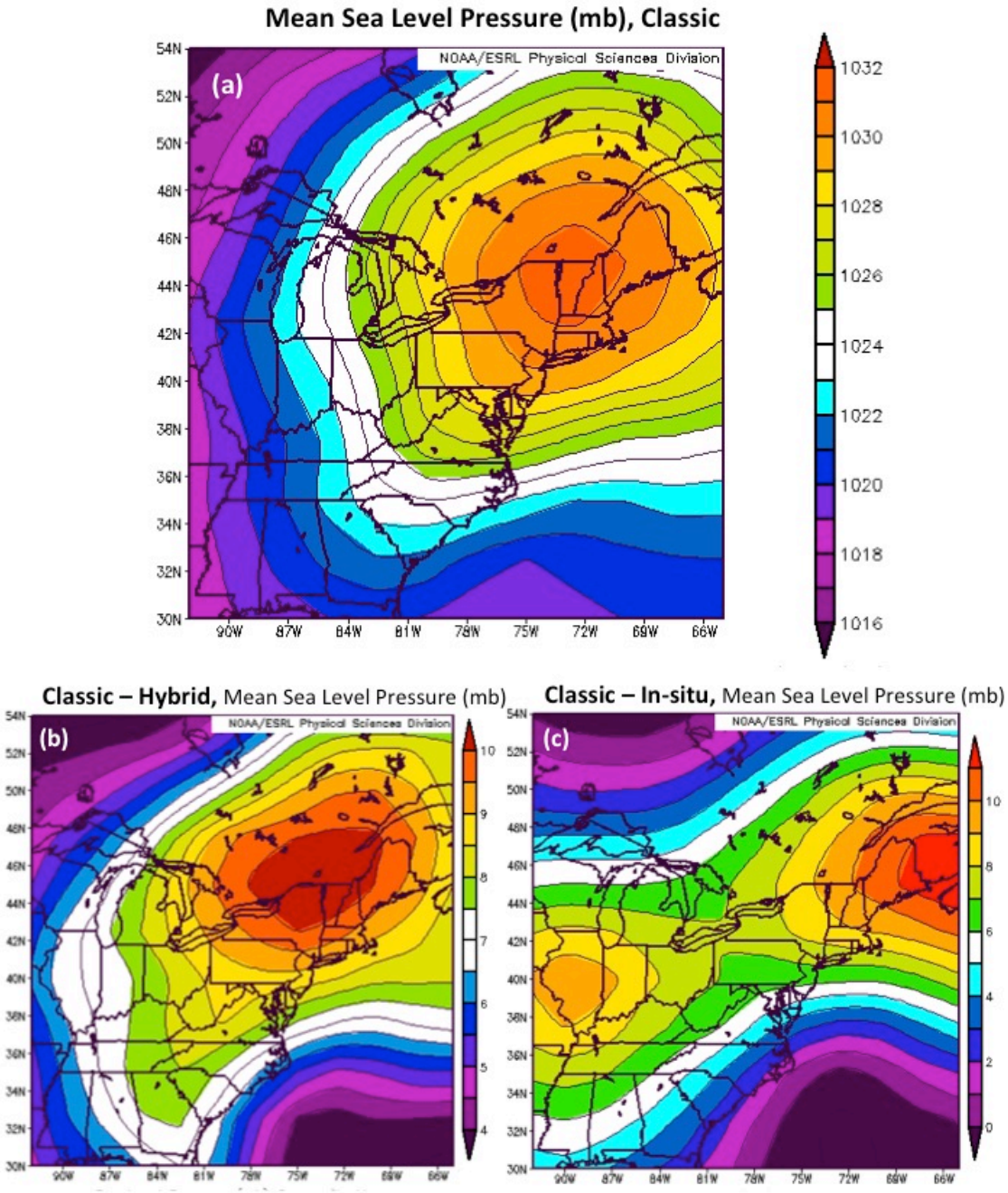


Figure 4.37a-e. Synoptic difference composites of surface air temperature (K) for (b) hybrid, (c) in-situ, (d) onset, and (e) erosion scenarios based on that of (a) problematic classic cold-air damming events.

On average, hybrid cases are considereably warmer (11K) across the CWA than classic CAD cases (Figure 4.37b), while in-situ cases are also much warmer (8K) (Figure 4.37c). This implies these cases were much weaker than their classic counterparts.

Differences in sea level pressure composites reveal that classic CAD cases have much stronger parent anticyclonic centers (by at least 10 millibars) over the northeast United States than do the other classifications (Figure 4.38b, Figure 4.38c).


Mean Sea Level Pressure (mb), Classic
(a)
NOAA/ESRL Physical Sciences Division
54N
52N
50N
48N
46N
44N
42N
40N
38N
36N
34N
32N
30N
90W
87W
84W
81W
78W
75W
72W
69W
66W
1032
1030
1028
1026
1024
1022
1020
1018
1016
Classic – Hybrid, Mean Sea Level Pressure (mb)
(b)
NOAA/ESRL Physical Sciences Division
10
9
8
7
6
5
4
Classic – In-situ, Mean Sea Level Pressure (mb)
(c)
NOAA/ESRL Physical Sciences Division
10
8
6
4
2
0

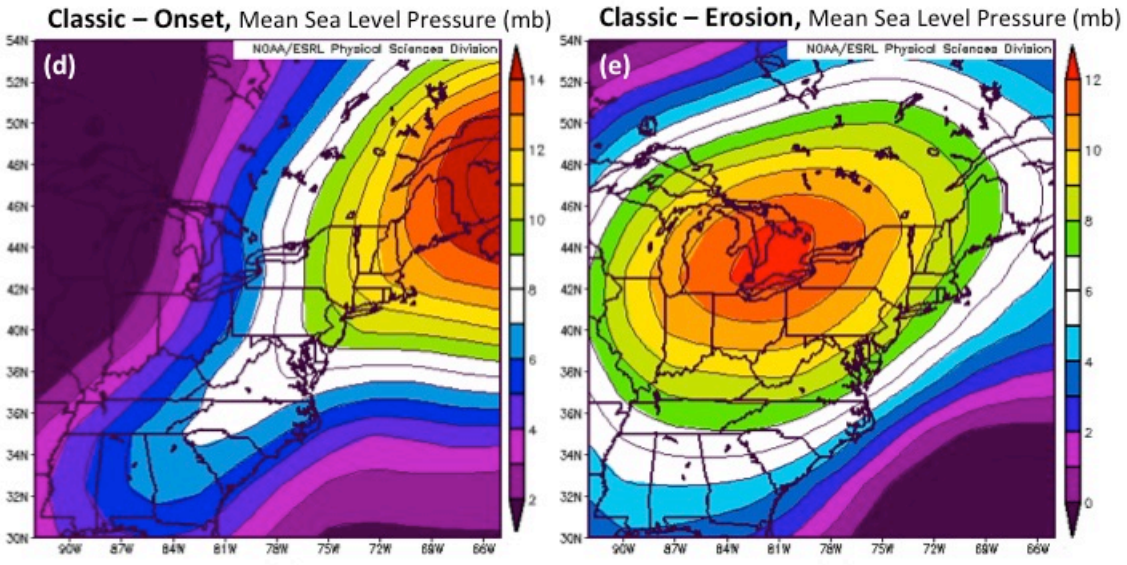


Figure 4.38a-e. As in Figure 4.37a-f, except for mean sea level pressure (mb).

Large sea level pressure differences associated with onset and erosion composites reveal a wedge either has not fully initialized in the Blue Ridge or has eroded from the region at the time of the composites (Figure 4.38d, Figure 4.38e). Much weaker hybrid and in-situ patterns suggest more localized impacts that affected these mis-forecasted cases as opposed to broad-scale atmospheric flow. Additionally, in-situ cases appear to be reinforced by a low pressure system to the west that is not present in the other classifications. While models may have anticipated stronger pressure patterns in hybrid and in-situ cases, a sizeable difference in surface pressure suggests models may have also entirely undetected weak events and forecast temperatures accordingly.

4.7.6 Synoptic Composites of Days Before Cold-air Damming Events

Surface pressure composites of days 1 through 3 before each of the CAD classifications with erred forecasts (not shown) reveal the wedge begins to form in the Blue Ridge on average 2 to 3 days before each forecast period, providing models

with sufficient time to detect a wedge scenario. A lack of sudden onset paired with no significant differentiating features from CAD scenarios throughout literature suggests these busted cases were not particularly problematic at a synoptic scale, and localized factors may have played a larger role in these erred forecasts.

4.8 Sounding Composites

Composite vertical atmospheric profiles of observed air and dew point temperatures and wind profiles during warm-season (May to October) and cold-season (November to April) busted CAD events at the Blacksburg NWSFO upper air site (KRNK) are plotted to help identify any atmospheric characteristics atypical of CAD. Profiles not shown, including cold season in-situ and warm season classic CAD busts, did not have sufficient sample sizes to produce meaningful results in the form of synoptic composites. Due to the shallow nature of the wedge, a 25-meter resolution was used in these composites to observe the lowest levels of the atmosphere at a high resolution that still yields a reasonable n-size on observations used to create each composite value. Since maximum daily temperatures generally fall between 12Z and 0Z, upper air soundings launched at both of these time periods are used for erred maximum temperature forecast composites; the same goes for minimum temperature forecasts commonly falling between 0Z and 12Z. These upper air balloon launches are performed at 654 meters above sea level, exhibiting a surface atmospheric pressure generally less than 950 millibars where the composites begin. This analysis observes conditions up to 500 millibars, as CAD is confined to the lower atmosphere.

The composite air and dew point temperatures of the 14 problematic classic CAD scenarios during the cold season are then plotted against the 71 cold season classic damming cases between 2007 and 2016 identified by Ellis et al. (2017) through synoptic weather typing (SSC) analysis. Since these particular CAD days identified by Ellis et al. are not in the NWSFO Blacksburg bust database, the argument can be made that these events were well forecast. Any duplicate dates between the busted CAD events and dates identified by Ellis et al. (2017) in which a wedge was present at the Blacksburg observation site are withdrawn from calculations for the SSC-identified, or well forecasted, composites. Though only 14 busted classic scenarios are used, this analysis aims to diagnose possible reasons as to why MAV and MET exhibit certain biases by comparing the differences between busted scenarios and an available database of seemingly well-forecast central Appalachian CAD cases.

Vertical atmospheric composites of both warm and cold season CAD forecast busts associated with onset highlight dry air at the surface being advected into the Bue Ridge by northeasterly surface winds (4.39a). Southwesterly flow caps a nearly isobaric temperature profile between 875 hPa and 850 hPa in warm season onset, and this warm capping flow becomes westerly by 725 hPa. Cold season composites of onset busts reveal a slight subsidence inversion between 900 hPa and 875 hPa, indicative of cold air beginning to pool alongside the eastern slopes of the Appalachians.

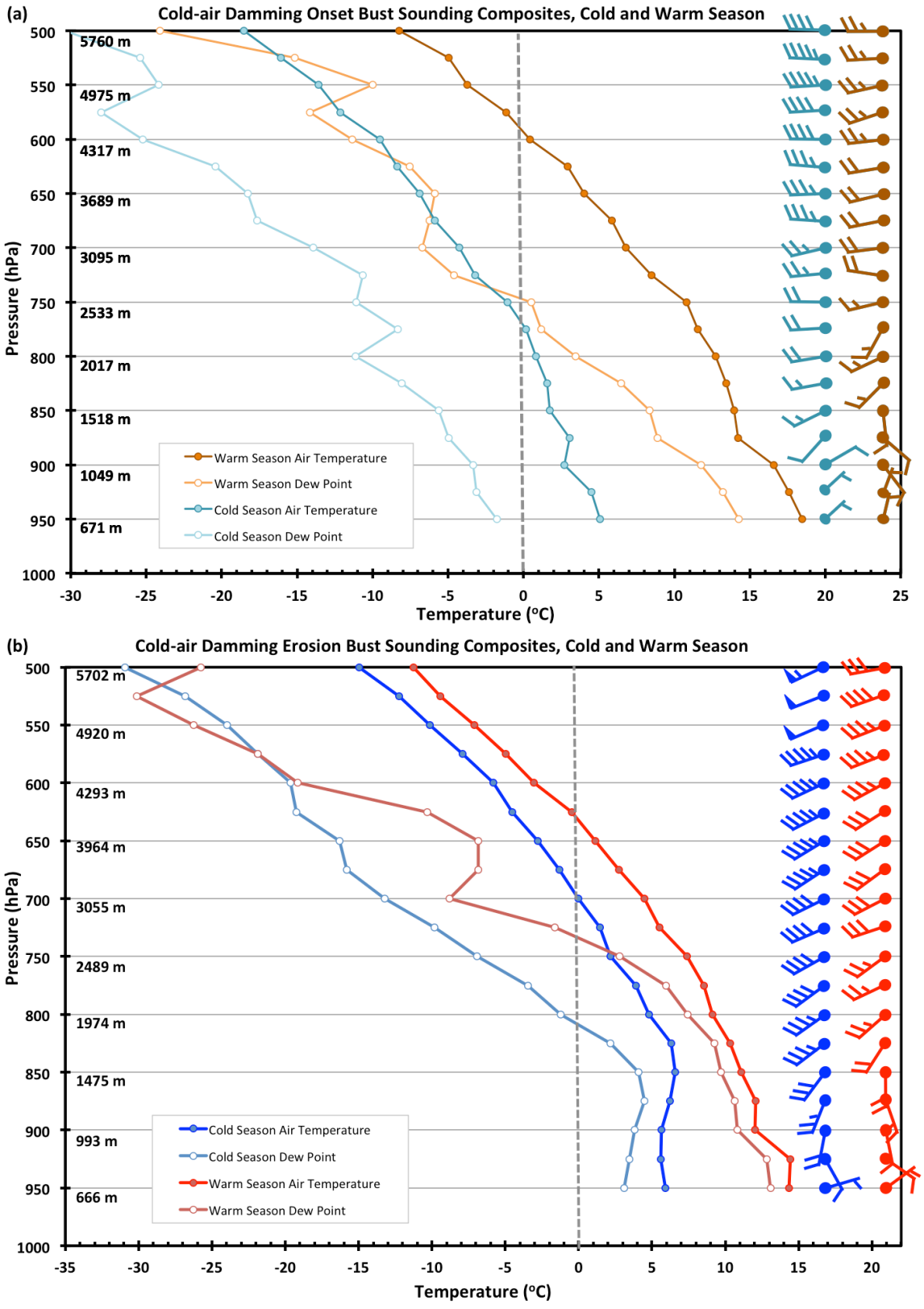

(a)
Cold-air Damming Onset Bust Sounding Composites, Cold and Warm Season
Pressure (hPa)
Temperature (°C)
5760 m
4975 m
4317 m
3689 m
3095 m
2533 m
2017 m
1518 m
1049 m
671 m
Warm Season Air Temperature
Warm Season Dew Point
Cold Season Air Temperature
Cold Season Dew Point
(b)
Cold-air Damming Erosion Bust Sounding Composites, Cold and Warm Season
Pressure (hPa)
Temperature (°C)
5702 m
4920 m
4293 m
3964 m
3055 m
2489 m
1974 m
1475 m
993 m
666 m
Cold Season Air Temperature
Cold Season Dew Point
Warm Season Air Temperature
Warm Season Dew Point

Figure 4.39a-b. Sounding composites of actual atmospheric conditions during busted cold-air damming forecasts associated with (a) onset and (b) erosion, 2007 to 2016.

As for the CAD erosion bust composites, a near-saturated lower atmosphere in both seasons, but moreso during summer months, is capped by westerly warm air advection over the wedge (Figure 4.39b). This signature also indicates mixing aloft, typical of an encroaching cold front. These scenarios are characteristic of both the onset and demise of central Appalachian CAD.

Vertical temperature profiles of problematic warm season hybrid and in-situ CAD events show more of what is typical for this phenomenon (Figure 4.40).

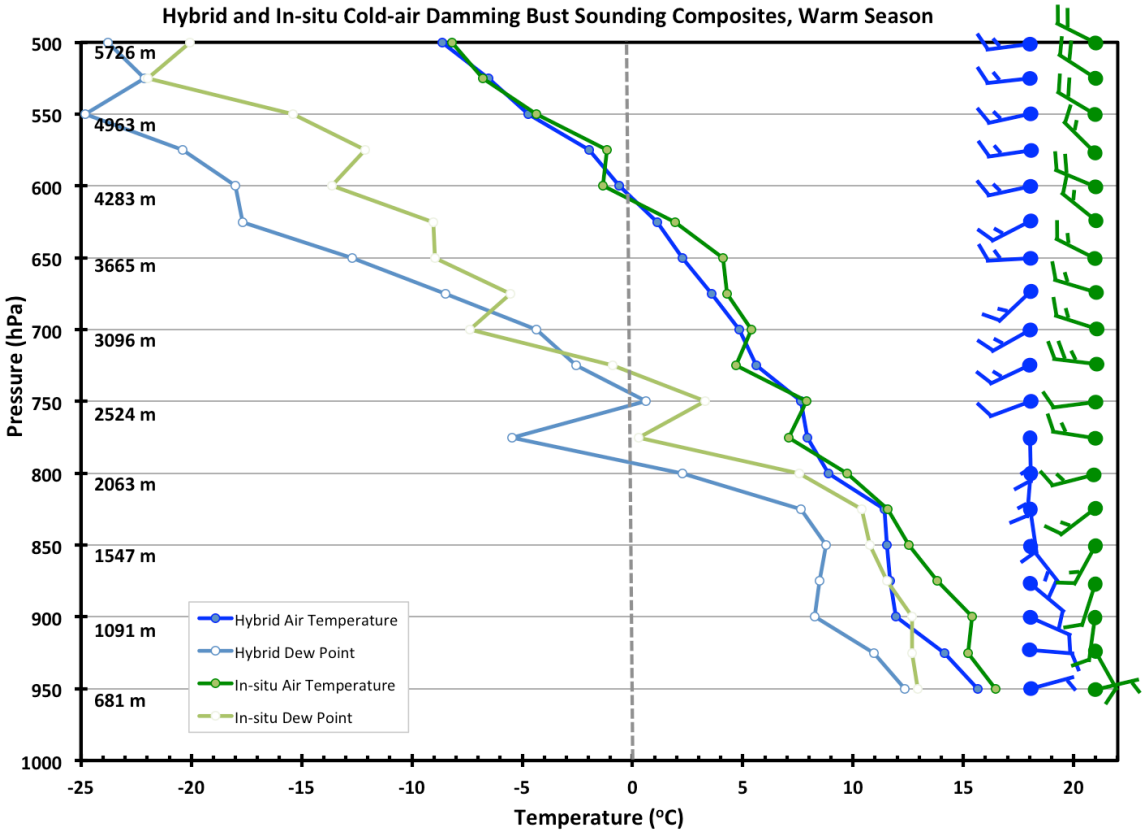


Figure 4.40. Sounding composites of actual atmospheric conditions during warm season hybrid and in-situ busted cold-air damming forecasts, 2007 to 2016.

Hybrid and in-situ composites reveal similar moist layers at the surface, and in-situ busts experience a weak subsidence inversion at 925 hPa. Hybrid temperatures remain nearly isothermal between 900 hPa and 825 hPa, suggesting these CAD cases were rather marginal. Alternately, as the sample size for warm season hybrid composites only includes 12 cases, any surface inversions may have been masked within the composites. In-situ profiles remain slightly warmer than that of hybrid CAD, indicating these cases may be weaker than hybrid bust scenarios. Westerly capping flow is apparent across the board, though in-situ wind direction displays stronger southerly components at the lowest levels of the atmosphere.

Cold season classic events are colder and drier than that of cold season hybrid cases and exhibit a stronger and more consistent subsidence inversion (Figure 4.41).

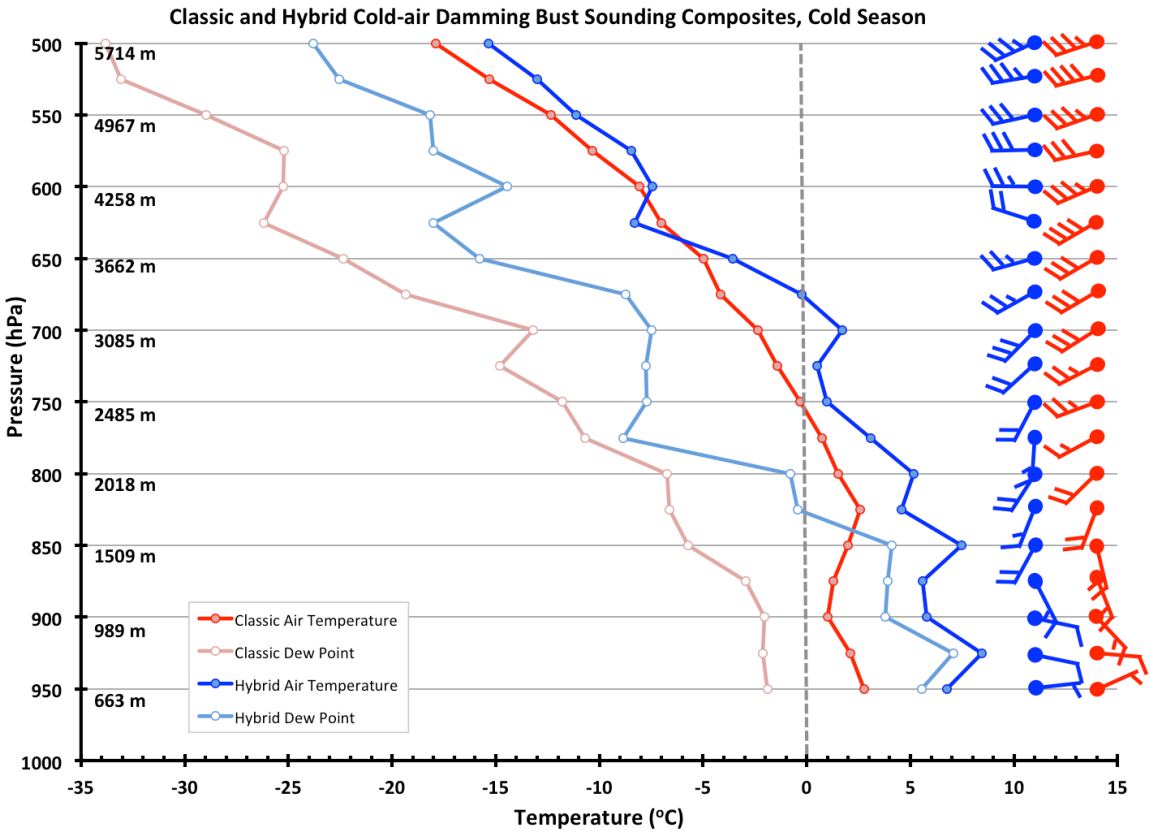


Figure 4.41. Sounding composites of actual atmospheric conditions during cold season classic and hybrid busted cold-air damming forecasts, 2007 to 2016.

Composites reveal a very moist, near-saturated surface layer in cold season hybrid events and a strong subsidence inversion in classic CAD cases. Hybrid cases exhibit an erratic temperature profile with several inversions that may result from a limited sample size of 6 cases. Westerly capping flow above the CAD wedge is evident in both cold season classes, though classic CAD busts experience slightly stronger wind shear in the atmosphere's lowest levels. Both warm and cold season composites reflect conditions standard of Blue Ridge CAD episodes, lacking any strong evidence as to what set these problematic cases apart from well-forecast CAD events.

4.8.1 Sounding Comparison of Busted Classic & SSC-Identified Well-Forecast Classic Cold-air Damming Events

Vertical atmospheric profiles of busted classic CAD cases reveal conditions to be slightly colder and moderately drier than that of well-forecasted wedge scenarios, while composite wind speed and direction of the two are nearly identical (Figure 4.42), indicating a similar degree of cold air advection for both scenarios.

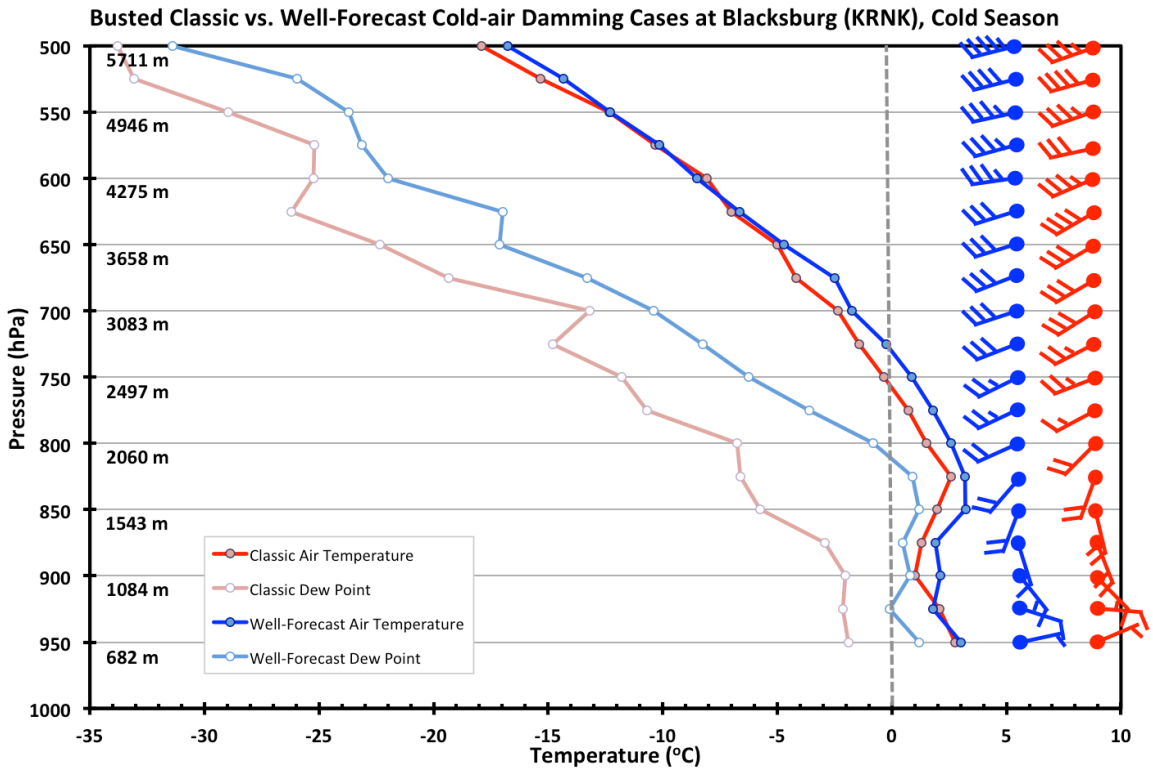


Figure 4.42. Atmospheric sounding composites of actual conditions during cold season (November to April) busted classic cold-air damming cases versus well-forecasted classic cases at Blacksburg (KRNK) to 500mb, 2007 to 2016.

Paired with depressed relative humidity values in comparison to well-forecast cases, classic bust composites also consistently exhibit lower mixing ratio values (not shown) than opposing SSC-identified cases, indicative of a drier lower atmosphere during these forecast busts. A drier atmosphere during cold-season problematic classic cases suggests models struggle with drier wedge scenarios, potentially originating from model difficulties in parameterizing atmospheric

moisture during these setups and predicting atmospheric conditions to be more moist than in actuality. Insufficient modeling of Atlantic moisture fetch when cold air is advected from the Northeast may perhaps cause this model error, while an overestimation of diabatic processes in wedge initiation and maintenance may also provide some source of inaccuracy. However, without a comprehensive comparison of CAD scenarios that did not bust in the CWA between 2007 and 2016, no conclusions can be drawn. Since the cases identified by Ellis et al. (2017) detect the purest, most obvious instances of CAD (i.e., classic CAD), the well-forecast wedges in this comparison may simply be stronger and more moist by nature. The drier nature of classic CAD busts alligns with MAV and MET temperature errors, condusive of a warm bias during the day in maximum temperature forecasts and a cold bias at night in minimum temperature forecasts.

Chapter 5. Conclusions

5.1 Summary

Cold-air damming is a prevalent Mid-Atlantic weather phenomenon that occurs when cold, dense air is orographically blocked, or dammed, alongside the eastern slopes of the central Appalachian Mountains. Lower-than-normal maximum and higher-than-normal minimum temperatures, increased and prolonged cloud cover, and precipitation that produces hazardous impacts are common features of this weather event, which are well known for presenting difficulties to both human forecasters and weather prediction models. The Blue Ridge Mountains region has not been the focus of previously published CAD research, and no study has specifically examined the forecasting accuracy of MOS guidance during different CAD scenarios or wedge onset and erosion.

Meteorologists in the central Appalachian Mountains often face challenges when formulating CAD forecasts. Limited surface observations and poor model guidance combined with difficulties in predicting the timing of onset and erosion give this shallow synoptic-to-mesoscale phenomenon a problematic reputation among forecasters. Numerical weather prediction models do not always parameterize CAD accurately, and MOS guidance is notorious for inconsistent forecasts when a wedge enters the Blue Ridge. The NWSFO in Blacksburg, Virginia archives especially difficult-to-forecast CAD cases using an 8°F threshold for erred temperature forecasts. This CAD bust database reflects the forecast challenge this weather situation presents, and using this database to assess how MOS MAV and MET guidance have handled these forecasts in which operational meteorologists

struggled sheds light on the biases that occur in their parent models. This study focused on biases in problematic MAV and MET minimum and maximum air temperature forecasts, how they changed between 2007 and 2016, and why they might have occurred. This study also stipulates how forecasters might use this bias information operationally during central Appalachian CAD.

The CAD cases in the Blacksburg NWSFO bust database were fairly evenly distributed annually, suggesting no systematic changes in forecast difficulty during the study period. Monthly frequency of busted CAD forecasts is characterized by a peak in the spring and decline during summer months, consistent with previous research conducted by Bell and Bosart (1988) that indicated a greater frequency of springtime CAD cases. Broken into classifications, erosion accounted for a high proportion of total cases (47 of 110), reinforcing the difficult nature of forecasting the demise of CAD. Frequencies of busted forecasts associated with classic, hybrid, in-situ, and onset types of CAD across the ten-year period remained relatively similar through time with a majority of these cases falling within cold-season months. Three out of four busted forecasts occurred due to erroneous high temperature forecasts during the day between 12Z and 0Z, presenting high temperatures as a greater forecast challenge than nighttime lows, especially during CAD erosion. The number of forecast busts remained spatially consistent across all six TAF sites, lacking a correlation between wedge type and TAF site distribution. Most busted forecasts within a CAD event involved only one of the six TAF sites in the CWA, suggesting that the complex terrain of the region produces localized inconsistencies that offer challenges for both forecasters and models.

Across the board, the most accurate forecasts tended to be produced 12 hours ahead of the forecast period. This may be a product of more recent surface observations enhancing forecast accuracy, or a possible result of MOS tendencies to weigh climatologically normal temperatures more heavily during earlier forecast cycles with greater lead time. Warm biases were evident in classic, hybrid, in-situ, and onset maximum temperature forecasts, as well as for minimum temperature forecasts associated with CAD erosion. Alternately, cold biases were exhibited in classic, in-situ, and onset minimum temperature forecasts, but maximum temperature forecasts associated with CAD erosion. MOS tendencies to underestimate or even fail to detect CAD scenarios in the central Appalachians are reflected in prevalent warm high-temperature and cold low-temperature forecast biases. Meanwhile, MOS may have tended to overpredict the lifetime of the cold dome during erosion scenarios reflected in cool-biased maximum and warm-biased minimum temperature forecasts. MET solutions were slightly more accurate for all classifications except erosion, in which MAV predictions outperformed MET forecasts. This could possibly stem from MOS-derived GFS and NAM tendencies to underestimate wedge strength, in which GFS-MOS guidance seems to particularly struggle with. Though neither model guidance performed statistically better than the other, forecasters may want to consider these nuances when formulating forecasts using MOS input.

Biases across the six TAF sites were highly variable, notably due to the region's diverse topographical features. Warm minimum and maximum temperature biases at Blacksburg may be an outcome of limited vertical resolution

in MOS parameterization, underestimating wedge strength at higher sites. This is also reflected in the difference in low temperature biases between Blacksburg and Roanoke, stressing the possibility of lacking vertical proficiency in model parameterization. Warm maximum and cold minimum temperature biases at Danville, Lewisburg, Lynchburg, and Roanoke reflect a possible over-exaggeration of diurnal temperature changes in MOS predictions, implying that MAV and MET solutions forecast climatologically normal conditions despite a present CAD set-up. Affected by frequent cloud cover and ample moisture from the Greenbrier River that sustains overnight warmth, the cold minimum temperature bias at Lewisburg may be enhanced by MOS parent model terrain smoothing during onset and incorrect assumptions that cold air encompasses the region uniformly across the rough terrain. During CAD erosion, higher elevations like Bluefield and Lewisburg exhibited a consistent cold bias while lower sites trended warm, highlighting potential errors in handling mountainous terrain. MOS solutions were rather inconsistent during CAD erosion, indicating that the GFS and NAM tend to prematurely erode the wedge and predict it to persist nearly equally. Again, despite small sample sizes, these TAF site biases propose location-based nuances to consider when producing forecasts during central Appalachian CAD scenarios.

Across the 10-year study period, both MAV and MET guidance follow the same bias trends when forecasting these problematic CAD cases. Neither model performed significantly better than the other, suggesting neither model should be favored over the other and both MET and MAV solutions should be incorporated into future CAD forecasts.

In attempts to diagnose why these biases may be present in MAV and MET busted forecasts, atmospheric composites during these CAD cases were generated at the surface using hourly data, significant synoptic levels, and vertically through archived upper air soundings. While synoptic soundings showed little evidence of characteristics dissimilar to that of the typical nature of Appalachian CAD, hourly surface composites indicated that Bluefield and Danville were commonly outside of the cold dome boundary. Synoptic difference composites based on classic types also denote that many of these cases were very weak in terms of surface pressure and temperature. The commonality of weak wedge scenarios among the problematic CAD cases may have led these episodes to progress poorly detected by the NAM and GFS, resulting in large model and human forecast error.

Though seasonal vertical atmospheric composites of CAD classifications failed to reveal differences from what might be expected from a cold-air wedge, comparisons of cold-season classic busts to well-forecasted cases identified by Ellis et al. (2017) imply that problematic cases are drier wedge scenarios. This points to a potential moisture parameterization issue in MOS equations that fail to properly predict atmospheric moisture content during more dry CAD setups, possibly stemming from poorly modeled Atlantic moisture fetch. However, this may also result from the nature of SSC-identified CAD events as being particularly moist, and these results are inconclusive as to explicitly identifying a problem within MOS forecast equations. However, since these CAD forecast busts look like typical central Appalachian wedge scenarios, these forecast problems are likely not made from

human error and perhaps stem from within MAV and MET guidance forecast equations.

5.2 Limitations and Future Work

Without a comprehensive database of CAD events that occurred in the Blacksburg CWA between 2007 and 2016, including well-forecast CAD, definitive comparisons cannot be made to determine what specifically caused these erred forecasts and temperature biases. Furthermore, a small n-size of cases in general is further narrowed when broken into classifications, causing a lack of meaningful results in certain calculations. CAD is a highly subjective event, and its definition may differ between forecasters who archive these events or even between different sources of literature. Lastly, the vertical sounding analysis was limited by a lack of model forecast sounding data in a user-friendly format, which would have elevated this study's findings as there were few 'well-forecasted' CAD cases to use for comparison.

Though the CAD 'bust' database archived by the NWSFO Blacksburg is a great starting point for researching the biases in MAV and MET temperature forecasts during central Appalachian CAD episodes, the compilation of a database of all CAD cases that occurred between 2007 and 2016 would provide more meaningful results. A further investigation of model success in forecasting precipitation and sky cover may enhance future results as these atmospheric variables heavily influence temperature, and the use of model soundings would provide more insight on characteristics of upper air sounding composites during these busted cases. Investigating the biases of more models would help forecasters understand a variety

of numerical weather prediction tools to use when incorporating them into wedge forecasts, and a gridded resolution of temperature versus site-specific data may provide more detail to depict localized idiosyncrasies within the CWA. This research presents a solid database of poorly forecasted CAD scenarios, and there are many different possibilities as to where to take further research to improve forecasts during this phenomenon. Most notably, the underlying equations of MAV and MET maximum and minimum forecasts necessitate further research to assess why MOS exhibits certain biases during CAD episodes in the central Appalachian Mountains.

Despite a lack of conclusive sources of forecast error during the problematic CAD cases in this study, numerous changes may have proved useful to better understand the conundrum that is Appalachian CAD forecasting. The compilation of a database of well-forecast CAD cases across the study area between 2007 and 2016 alongside the verification of busted forecasts would provide a solid foundation in which to make meaningful comparisons between datasets that could potentially identify more localized sources of forecast error. Furthermore, the available dataset of CAD busts uses human forecasts to determine erred CAD forecasts, whereas this study focuses on model-driven forecast error. Though in many cases the MAV and MET forecasts were outperformed by human predictions, this was not always the case. The best approach to expose model forecast error during Appalachian CAD would focus purely on MAV and MET errors of 8°F or more rather than relying on official NWS forecast error to flag a forecast bust. Additionally, focusing purely on the forecast accuracy during the erosion of problematic CAD events would prove highly important, as reflected in the high proportion of erosion events in this study;

this would also reduce the number of datasets with limited sample sizes. In addition, if given more time to manipulate archived forecast model soundings into a user-friendly format, the analysis of GFS and NAM forecast soundings at each of the six TAF sites could provide useful information into the subtleties of CAD forecasting across the CWA's complex terrain.

5.3 Final Thoughts

Forecasting CAD in the central Appalachian Mountains is no simple task, as demonstrated by the creation and upkeep of a wedge forecast 'bust' database within the NWSFO Blacksburg. MOS guidance generally handles these situations poorly, but an improved understanding of MAV and MET temperature forecast biases may aid in producing more accurate forecasts during this phenomenon. As a general rule based on this study, it's safe to undercut MOS-derived maximum temperatures by several degrees and raise guidance-suggested minimum temperatures across nearly all CAD types. Erosion, on the other hand, necessitates the opposite. Though MET solutions are slightly more accurate than MAV, it's best to consider both models when formulating temperature forecasts. Higher resolution models may provide more accurate interpretations of CAD, and considering elevation and topography is essential when formulating forecasts, as higher elevations may skew towards colder MOS biases depending on local geography. Though the nuanced results in this study do not identify a definitive solution to the conundrum of CAD forecast errors, they should still prove useful when forecasting CAD events in the Blacksburg CWA as they may help to improve wedge forecasts based on MOS guidance.

Ultimately, the CAD events archived into the bust database by operational meteorologists at the NWSFO Blacksburg appeared to be normal cases of the phenomenon lacking clear qualities distinguishing them from typical cases. Errors within these CAD forecasts are not necessarily formed by unusual atmospheric circumstances, but seem to result from problems within the MOS forecast equations themselves that guide forecasters to produce erroneous forecasts. MAV and MET solutions may have difficulty predicting levels of atmospheric moisture during dry CAD scenarios, though this is unsupported by a comparable database of well-forecasted wedge patterns. Furthermore, the cases in the bust database are further complicated by factors such as sky cover, precipitation, and evaporative cooling that add elements of difficulty to these temperature forecasts. The complexity of these cases extends beyond simple temperature forecasts that may be severely limited by NAM and GFS vertical resolutions. Forecaster error is likely not the source of the problem, and more research on how model equations parameterize Appalachian CAD is necessary to diagnose why these air temperature biases occur.

References

Bailey, C. M., Hartfield, G., Lackmann, G. M., Keeter, K., & Sharp, S. (2003). An objective climatology, classification scheme, and assessment of sensible weather impacts for appalachian cold-air damming. *Weather and Forecasting*, **18(4)**, 641-659.

Bell, G. D., & Bosart, L. F. (1988). Appalachian cold-air damming. *Monthly Weather Review*, **116(1)**, 137-161.

Bosart, L. F. (2003). Whither the weather analysis and forecasting process? *Weather and Forecasting*, **18**, 520-529.

Ellis, A. W., Marston, M. L. and Nelson, D. A. (2017). An air mass-derived cool season climatology of synoptically forced Appalachian cold-air damming. *International Journal of Climatology,* **38(2)**, 530-542.

Forbes, G. S., R. A. Anthes, and D. W. Thomson. (1987). Synoptic and mesoscale aspects of an Appalachian ice storm associated with cold-air damming. *Monthly Weather Review*, **115**, 564-591.

Glahn, H. R., & Lowry, D. A. (1972). The use of Model Output Statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology*, **11**, 1203-1211.

Kalnay, E. and Coauthors (1996). The NCEP/NCAR Reanalysis 40-year Project. *Bulletin of the American Meteorological Society*, **77**, 437-471.

Keeter, K., S. Businger, L. G. Lee, and J. S. Waldstreicher. (1995). Winter weather forecasting throughout the eastern United States. Part III: The effects of topography and the variability of winter weather in the Carolinas and Virginia. *Weather and Forecasting*, **10**, 42-60.

Koch, S. E. (2001). Real-time detection of split fronts using mesoscale models and WSR-88D radar products. *Weather and Forecasting*, **16(1)**, 35-55.

Mahoney, K. M., and G. M. Lackmann. (2006). The sensitivity of numerical forecasts to convective parameterization: A case study of the 17 February 2004 East Coast cyclone. *Weather and Forecasting*, **21**, 465-488.

NOAA - National Weather Service (2018). Model Output Statistics (MOS): What is MOS? Retrieved from https://www.weather.gov/mdl/mos_home

NOAA – National Weather Service (2017). [Map of Blacksburg CWA]. Retrieved from http://www.wx4rnk.org/images/rnk_cwa.jpg

NOAA – National Weather Service Forecast Office Blacksburg (2017). *Cold-air Damming Forecasting Practicum Lecture* [PowerPoint presentation]. Retrieved from NWSFO Blacksburg.

NOAA – NCEP Central Operations (2018). Changes to NCEP Models/Implementation Dates to NOAAPORT. Retrieved from http://www.nco.ncep.noaa.gov/pmb/changes/

NOAA – WPC Verification Threat Score and Bias Computation (2013). Retrieved from http://www.wpc.ncep.noaa.gov/html/scorcomp.shtml

Rackley, J. A., & Knox, J. A. (2016). A climatology of southern Appalachian cold-air damming. *Weather and Forecasting,* **31(2)**, 419-432.

Richwien, B. A. (1980). The damming effect of the southern Appalachians. *National Weather Digest*, **5(1)**, 2-12.

Stanton, W. M. (2003). An Analysis of the Physical Processes and Model Representation of Cold Air Damming Erosion. M.S. Thesis, North Carolina State University, 224 pp.

Stauffer, D. R., and T. T. Warner. (1987). A numerical study of Appalachian cold-air damming and coastal frontogenesis. *Monthly Weather Review*, **115**, 799-821.

Xu, Q. (1990). A Theoretical Study of Cold Air Damming. *Journal of the Atmospheric Sciences*, **47(24)**, 2969-2985.