



# Analysis of Model Thermal Profile Forecasts Associated with Winter Mixed Precipitation within the United States Mid-Atlantic Region

ANDREW W. ELLIS

*Department of Geography, Virginia Tech, Blacksburg, Virginia*

STEPHEN J. KEIGHTON

*NOAA/National Weather Service, Blacksburg, Virginia*

STEPHANIE E. ZICK

*Department of Geography, Virginia Tech, Blacksburg, Virginia*

ANDREW S. SHEARER

*School of Meteorology, The University of Oklahoma, Norman, Oklahoma*

CASEY E. HOCKENBURY

*Department of Geography, Virginia Tech, Blacksburg, Virginia*

ANITA SILVERMAN

*NOAA/National Weather Service, Blacksburg, Virginia*

(Manuscript received 20 June 2021; review completed 2 February 2022)

## ABSTRACT

Winter mixed-precipitation events across the mid-Atlantic region of the United States from 2013–2014 through 2018–2019 were used to analyze common short-term model forecasts of vertical atmospheric thermal structure. Using saturated forecast soundings of the North American Mesoscale (NAM), higher-resolution nested NAM (NAMnest), and the Rapid Refresh models—corresponding with observed warm-nose precipitation events (WNPEs)—several thermal metrics formed the basis of the analysis of observed and forecast soundings, including Bourgoignie positive and negative areas. While the three models accurately forecast the general thermal structure well during WNPEs, a warm bias is evident within each. Well forecast are maximum and minimum temperatures within the warm nose and surface-based cold layer, respectively, but the cold layer is commonly too thin for each of the models, and the warm nose is regularly too thick, particularly within NAM and NAMnest forecasts. Forecasts of a cold layer that is too shallow tend to coincide with observations of stronger synoptic-scale upward motion, a deeper cold surface-based layer, and a higher isentropic surface. Forecasts of a warm nose that is too thick tend to coincide with observations of weaker upward motion, a shallower cold surface-based layer, and a lower isentropic surface across the region. Two-thirds of precipitation-type estimates from model soundings agreed with those derived from observed soundings, with the remaining third predominantly representing a warm bias in precipitation type.

## 1. Introduction

Winter mixed-precipitation events, particularly those involving freezing rain, are among the most disruptive and costly weather hazards. Billion-dollar

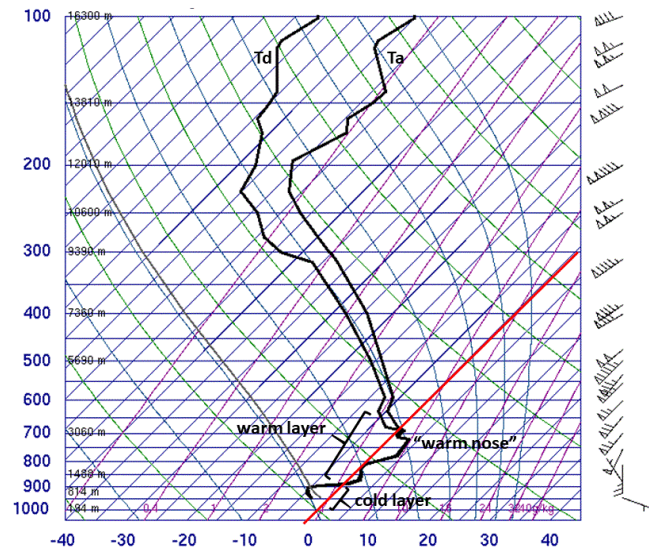
damage events in the United States and Canada have been well documented (McCray et al. 2019), while the study of insured property losses in the United States due to freezing rain revealed >\$16 billion in damages over a 52-yr period (Changnon 2003). The most

obvious impacts are interruption of travel and damage to power infrastructure, which can result in prolonged power loss. For example, an ice accretion of 0.25–1.27 cm (0.1–0.5 in) from an autumn 2018 event across southwestern Virginia within the eastern United States disrupted power for 77% of utility customers in one rural county while triggering 59 traffic accidents on state roads within a 4-h period across a four-county area (Gendreau 2021). Areas of the eastern United States east of the Appalachian Mountains are susceptible to mixed-precipitation events as low-level cold air can become entrenched against the mountains beneath warm flow aloft (Keeter et al. 1995; Rauber et al. 2001; Changnon and Karl 2003).

Freezing rain and ice pellets commonly occur when ice crystals aloft fall through a warm ( $>0^{\circ}\text{C}$  or  $32^{\circ}\text{F}$ ) layer above surface-based cold ( $<0^{\circ}\text{C}$  or  $32^{\circ}\text{F}$ ) air. The degree to which the ice crystal melts dictates the surface precipitation type, with partial melting leaving an ice nucleus that aids refreezing to an ice pellet, and complete melting requiring the aid of a cold surface for refreezing as accretion, or freezing rain (Forbes et al. 1987). Light freezing rain also can occur in a homogeneously cold atmosphere within which moisture/saturation is shallow, and thus too warm to generate the ice crystals necessary for atmospheric ice formation. Limited moisture and weaker vapor diffusion without ice crystal nuclei generally yields very light freezing rain or drizzle (McCray et al. 2019).

Warm southerly flow above a surface cold layer is relatively common within the eastern United States east of the Appalachian Mountains, where cold-air advection directed from the northeast by an anticyclone to the north “dams” cold air against the mountains. Cold-air damming yields ageostrophic flow of cold air southward along the eastern mountain slopes while inducing higher atmospheric pressure with an accumulation of mass and a depth of cold air sufficient for frozen or freezing precipitation. Warm-air advection above the cold layer can result from southerly geostrophic flow, or from southerly flow originating lower in the atmosphere in association with passage of a mid-latitude cyclone, or weaker low-pressure center.

Within the thermal profile common to ice-pellet and freezing rain events, the warm air aloft often is referenced as the “warm nose” because of the shape of the profile on a thermodynamic diagram. In the example from 1200 UTC 18 February 2021 at Blacksburg, Virginia (KRNK, Fig. 1), a deep warm layer, with a maximum temperature of  $4.4^{\circ}\text{C}$ , extends from around



**Figure 1.** Atmospheric sounding at 1200 UTC 18 February 2021 at Blacksburg, VA (KRNK), as portrayed on a skew $T$ –log $p$  thermodynamic diagram. Vertical profiles of air ( $T_a$ ) and dewpoint ( $T_d$ ) temperatures are represented by the solid, black lines. The  $0^{\circ}\text{C}$  isotherm is in red (Obtained from <http://weather.uwyo.edu/upperair/sounding.html>). Click image for an external version; this applies to all figures and hereafter.

890 to 690 hPa above surface-based cold layer (946–890 hPa) with a minimum temperature of  $-5.3^{\circ}\text{C}$ . At or close to saturation through about 700 hPa, the thermal profile yielded a mix of freezing rain and ice pellets. For this form of thermal profile, precipitation type is predicated on the temperature and thickness of each of the layers—freezing rain (ice pellets) from a generally warmer (cooler) atmosphere consisting of a thicker (thinner) warm layer and a thinner (thicker) cold layer.

While the vertical thermal profile and synoptic atmospheric conditions for mixed-precipitation events in the eastern United States are rather well understood in theory, precipitation-type forecasting remains a practical challenge (Ralph et al. 2005; Ikeda et al. 2017; McCray et al. 2019). As Hux et al. (2001) noted, “the proper forecast of type and duration of winter precipitation is one of the most difficult challenges in operational meteorology.” Forecasts of thermodynamic structure are key, and it is important to understand interregional variability in model and algorithm performance (Robbins and Cortinas 2002). Guidance for forecasters as to how individual model biases might influence model blends of first-guess fields (e.g., warm-nose strength) that determine precipitation-type output is useful (Wandishin et al. 2005). Ikeda et al.

(2013) evaluated the ability of the National Oceanic and Atmospheric Administration's (NOAA) High-Resolution Rapid Refresh (HRRR) model to determine precipitation phase as rain, snow, or mixed during the 2010–2011 winter season. They found good qualitative model performance in terms of the spatial coverage in rain–snow transition areas and freezing–precipitation cases. However, HRRR model performance was less skillful quantitatively in regions with mixed precipitation (Ikeda et al. 2013). Their results signaled a need for a deeper diagnosis of model shortcomings in depicting mixed-precipitation events. Subsequently, Ikeda et al. (2017) re-engaged the HRRR model's ability to forecast mixed-precipitation using surface observations and soundings across two winter seasons. They found the model to forecast the observed temperature profile associated with mixed-precipitation events reasonably well, but they identified a warm bias—particularly at the surface ( $<2^{\circ}\text{C}$ ) and in cold-air damming situations in the eastern United States ( $\sim 4^{\circ}\text{C}$ )—yielding errors in precipitation-type forecasts.

In this study, we analyze model forecast soundings associated with warm-nose precipitation events (WNPEs) for three numerical weather prediction (NWP) models operated by the National Centers for Environmental Prediction (NCEP): North American Mesoscale (NAM), higher-resolution nested NAM (hereafter, NAMnest), and Rapid Refresh (RAP). These represent the three relatively high-resolution models for which forecast sounding archives are readily available, while they also represent the model solutions used in the short term by the National Weather Service (NWS) National Blend of Models (Craven et al. 2020; [https://www.meted.ucar.edu/education\\_training/course/52](https://www.meted.ucar.edu/education_training/course/52)). All three models are generated with Weather Research and Forecasting (WRF) model dynamical solvers. The two NAM models use the WRF Nonhydrostatic Mesoscale Model (WRF-NMM) solver and run four times per day (0000, 0600, 1200, and 1800 UTC), out to 84 h for the NAM and 60 h for the NAMnest (NCEP 2020). The NAM and NAMnest models are run with 12- and 3-km horizontal grid spacings, respectively, with 60 vertical layers. The RAP model uses the Advanced Research WRF (WRF-ARW) core and operates on a 3-km spatial grid with 50 vertical layers and executes four times per day with hourly output to 51 h; the WRF-ARW also runs on an hourly schedule out to 21 h (NCEP 2020).

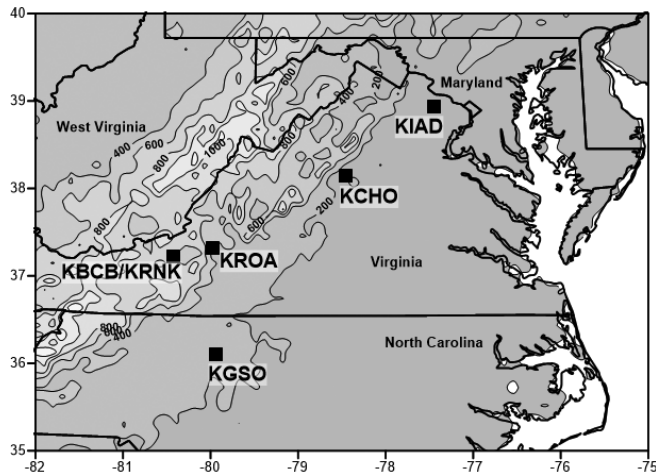
To focus on contemporary model operation, the period of study is limited to the six cool seasons

(November through April) from 2013–2014 through 2018–2019. Still, changes to the models during the study period could influence skill or bias. We reviewed the NCEP documentation of changes to the models and the implementation dates (<https://www.nco.ncep.noaa.gov/pmb/changes/>), and identified five modifications that could impact the accuracy of thermal profile forecasts. These modifications include: (1) February 2014 (RAP model)—upgrade of data assimilation, upgrade to WRF core, upgrade to microphysics, and new planetary boundary layer scheme; (2) August 2014 (NAM model)—upgrade of microphysics, radiation, and convective parameterization scheme, and changes to data assimilation; (3) August 2016 (RAP model)—changes to assimilation of hydrometeor and mesonet data, and updates to WRF and microphysics package; (4) March 2017 (NAM/NAMnest model)—NAMnest grid spacing change (4 km to 3 km), improved radiation scheme, quantitative precipitation forecast bias correction, changes to data assimilation, and reduction of terrain smoothing; (5) July 2018 (RAP model)—data assimilation upgrade, WRF update, new hybrid vertical coordinate system, improved terrain representation, improved simulation of air temperatures over terrain, and improved microphysics for upper clouds. It is beyond the scope of this study to examine whether a given change alters model performance related to the thermal profile, or to what degree performance is altered. Instead, we analyze the overall model performance with the caveats that continuous model upgrades to operational models can influence any biases that we uncover during the study period. Similar studies have been conducted to evaluate the forecast precipitation phase in multiple versions of the HRRR model (Ikeda et al. 2013, 2017) and for comparing multiple precipitation-type algorithms over multiple winter seasons (Reeves et al. 2014). Such an approach can be useful for understanding systemic biases for a given model (in this case with a focus on mixed-precipitation events in a particular geographic region).

The spatial domain for the study is an area of the mid-Atlantic region, east of the Appalachian Mountains, that is frequently susceptible to winter precipitation-type forecast challenges. Specifically, this study focuses on areas represented by three sounding locations [KRNK; Greensboro, North Carolina (KGSO); and Sterling, Virginia (KIAD); see Fig. 2] and five Automated Surface Observing System (ASOS)/Automated Weather Observing System (AWOS) locations (Table 1, Fig. 2). While surface observations of precipitation type are

**Table 1.** The four ASOS and one AWOS (KBCB) station locations (Fig. 1). The sounding station linked to each location is indicated.

| Location            | Station ID | Elevation (m) | Sounding Station |
|---------------------|------------|---------------|------------------|
| Chantilly, VA       | KIAD       | 93            | KIAD             |
| Charlottesville, VA | KCHO       | 195           | KRNK, KIAD       |
| Greensboro, NC      | KGSO       | 270           | KGSO             |
| Roanoke, VA         | KROA       | 358           | KRNK             |
| Blacksburg, VA      | KBCB       | 654           | KRNK             |



**Figure 2.** The locations of the five surface stations (KBCB, KCHO, KGSO, KIAD, KROA) and three upper-air stations (KGSO, KIAD, KRNK; Table 1) across surface elevation (200-m interval) within the mid-Atlantic region.

used to identify potentially impactful events, this study focuses on those with an identifiable warm nose in the sounding data, which represents the underlying cause of the precipitation type and is the key forecast element.

## 2. Data and methods

### a. Identification of warm-nose precipitation events

Raw Meteorological Aerodrome Reports (METAR) data for each of the five ASOS/AWOS study locations were downloaded from the Iowa Environmental Mesonet (IEM) data archive at Iowa State University (<https://mesonet.agron.iastate.edu/request/download.phtml>) for November through April for the 6-yr study period 2013–2014 through 2018–2019. Any of nine METAR weather report codes associated with precipitation (Table 2) in combination with a surface air temperature  $\leq 1.7^{\circ}\text{C}$  ( $35^{\circ}\text{F}$ ) were used to define the possibility of frozen or freezing precipitation.

The METAR code, unknown precipitation (UP), was included because of the known difficulty in mixed-precipitation identification by ASOS/AWOS sensors (Landolt et al. 2019). Potential WNPEs included all periods of  $\geq 6$  h [a threshold previously used to define a long-duration event (McCray et al. 2020)] with one or more precipitation types (Table 2) coincident with an air temperature  $\leq 1.7^{\circ}\text{C}$ . For inclusion as a WNPE,  $\geq 50\%$  of the observations within the period needed to meet the weather code and temperature criteria, with  $\geq 10\%$  of observations of a precipitation type other than snow or rain. The observation threshold for snow/rain eliminated events that were predominantly snow, cold rain, or mixed rain/snow.

The basis of a subjective classification of the synoptic pattern associated with each potential WNPE was the inspection of 3-h surface weather maps from the Weather Prediction Center’s Surface Analysis Archive ([https://www.wpc.ncep.noaa.gov/archives/web\\_pages/sfc/sfc\\_archive.php](https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive.php)), 12-h surface weather maps from the Storm Prediction Center’s archive (<https://www.spc.noaa.gov/obswx/maps/>), and archived radar data. Four synoptic patterns conducive to warm-air advection over a cold surface layer, each with some degree of cold-air damming, constituted the classification. These include: (1) passage of a low-pressure center to the west of the study area (southerly sourced warm advection), (2) Miller-A type low-pressure center evolution (Miller 1946), with cyclone development and emergence from the southeastern United States and passage along the Atlantic Ocean coastline (Atlantic-sourced warm advection), (3) Miller-B type low-pressure center evolution (Miller 1946), with cyclone passage from west-to-east over or to the south of the study area followed by intensification or secondary cyclogenesis near the coastline (Atlantic-sourced and/or southerly sourced warm advection), and (4) high-pressure center passage to the north in a classic cold-air damming scenario east of the Appalachian Mountains beneath isentropic lift of warmer southerly flow aloft (colloquially referred to as

**Table 2.** The nine METAR weather reports used to identify precipitation and their definitions. The codes may include a qualifying code of freezing (FZ) and modifiers to indicate intensity as heavy (+) or light (-).

| Code | Definition                      |
|------|---------------------------------|
| BR   | Mist                            |
| DZ   | Drizzle                         |
| GS   | Small Hail/Snow Pellets (<5 mm) |
| IC   | Ice Crystals                    |
| PL   | Ice Pellets                     |
| RA   | Rain                            |
| SG   | Snow Grains                     |
| SN   | Snow                            |
| UP   | Unknown Precipitation           |

“overrunning”). Stratification by synoptic pattern was intended to support a more granular analysis of forecast soundings associated with WNPEs.

Observational data from the array of soundings within the timeframe of each potential WNPE (and ultimately exhibiting a warm nose coincident with precipitation) were downloaded from the University of Wyoming upper-air data portal (<http://weather.uwyo.edu/upperair/sounding.html>). Soundings from KRNK were linked to potential WNPEs at three ASOS/AWOS locations (KBCB, KROA, KCHO), while soundings at KIAD were linked to potential WNPEs at two locations (KIAD, KCHO), and soundings at KGSO were only linked to potential WNPEs at the KGSO ASOS site (Fig. 2). At KCHO, a mixed-precipitation event led to review of coincident soundings at both KRNK and KIAD. As the interest is in WNPEs, only observed soundings exhibiting a warm nose above a freezing layer, and coincident with saturation of the lower atmosphere, were retained for comparison with model soundings. Saturation was defined as the occurrence of precipitation at the co-located ASOS/AWOS site at the time of the sounding or within the preceding hour (coinciding with sounding initiation).

Several metrics were calculated to characterize each sounding. In addition to surface air temperature, atmospheric thickness between key pressure levels was calculated as an indication of layer temperature. Simple linear interpolation was employed between observed data points to estimate the vertical position of key thresholds (e.g., 0°C air temperature) that fell between observations. Thicknesses were calculated for the 1000–500-, 1000–700-, 850–700-, and 1000–850-hPa layers. Multiple variables were recorded or calculated

for both the surface-based cold layer and elevated warm nose, including minimum and maximum cold- and warm-layer temperature, mean temperature, mean wind direction and speed, pressure and height bounds, and the negative and positive areas using the Bourgoiuin method (Bourgoiuin 2000). The Bourgoiuin method can be used to predict probable precipitation type based on thermal condition and depth of the warm and cold portions of the thermal profile, or the areas of the warm (positive area) and cold (negative area) layers on a thermodynamic diagram (e.g., tephigram or skewT–logp). We used the original Bourgoiuin method because the modified version from Birk et al. (2021) was not available at the time of data analysis. The Bourgoiuin method requires only mean layer temperature ( $T_l$ ) and potential temperature ( $\theta$ ) at the top and bottom of the layer. The Bourgoiuin layer area is calculated as

$$C_p |Area| = C_p \oint d \ln \theta = C_p \bar{T}_l \ln \left( \frac{\theta_{top}}{\theta_{bottom}} \right), \quad (1)$$

where  $C_p$  is the specific heat of air at constant pressure (1004 J kg<sup>-1</sup> K<sup>-1</sup>). Applied to the cold and warm layers separately, the computation yields a solution for area on a thermodynamic diagram in units of specific energy (J kg<sup>-1</sup>). Bourgoiuin (2000) used precipitation-type observations and a linear relationship between corresponding positive and negative areas to establish a predictor of precipitation type [ice pellets (IP) versus freezing rain (FZRA)] based on the relative sizes of the two areas. A small positive area and large negative area supports ice pellets, while a large positive area and small negative area supports freezing rain. The Bourgoiuin thresholds for determining probable precipitation type based on negative area (NA) and positive area (PA) values are:

$$IP: NA > (66 + 0.66PA) \quad (2)$$

$$FZRA: NA < (46 + 0.66PA) \quad (3)$$

$$FZRA/IP: (46 + 0.66PA) \leq NA \leq (66 + 0.66PA) \quad (4)$$

The Bourgoiuin method is an element of the BUFKIT forecast profile visualization and analysis toolkit (<https://training.weather.gov/wtd/tools/BUFKIT/index.php>, Mahoney and Niziol 1997) used in many NWS forecast offices.

### b. Observed versus forecast sounding analysis

For the three numerical weather prediction models that are the focus of this study (NAM, NAMnest, and RAP), forecast model sounding data associated with each observed warm-nose sounding were downloaded from the IEM archive site (<https://mtarchive.geol.iastate.edu/>). The data are Binary Universal Form for the Representation of meteorological data (BUFR) soundings converted to BUFKIT format and represent the native model vertical grid spacing. For each of the three models, data were downloaded for model runs at the observed sounding time (i.e., model initialization) and 12 h before the observed sounding. For the longer-horizon NAM and NAMnest models, data also were downloaded for model runs at 24 and 36 h before the observed sounding. For the shorter-horizon RAP model, data also were downloaded for model runs at 3 and 6 h before the observed sounding. Thus, four model runs were the focus for each of the NAM (0 h, +12 h, +24 h, and +36 h), NAMnest (0 h, +12 h, +24 h, and +36 h) and RAP (0 h, +3 h, +6 h, and +12 h) models. The higher-resolution models and shorter-term forecast horizons were chosen to focus on the winter-storm warning period and the near-term update periods for the higher-resolution models used by the NWS. While the RAP model generates forecasts out to 51 h four times per day, archived BUFR data are only available for the hourly model runs with forecasts out to 21 h, and it is unclear whether forecasters are aware when the blended cycles they are using include the extended RAP cycles.

Observed and forecast saturated soundings during WNPEs were compared for each of the three sounding locations. Comparisons were made for each of the four model runs of each of the three models. With the research focus on forecasts of the vertical thermal profile rather than error in forecast timing, flexibility in the timing of a saturated forecast sounding (model-generated precipitation) relative to the observed sounding time was instituted. Designated for comparison with an observed sounding was the saturated forecast model sounding nearest the observed sounding time, but with a limit of  $\leq 3$  h from the observed sounding. Beyond visual depictions of observed and forecast soundings, mean absolute error [MAE; Eq. (5)] and mean percent error [MPE; Eq. (6)] were calculated to reflect the general magnitude and bias of forecast error.

$$MAE = \frac{\sum_{i=1}^n |forecast - observed|}{n} \quad (5)$$

$$MPE = \frac{\sum_{i=1}^n \left( \frac{forecast}{observed} \right) \times 100\%}{n} \quad (6)$$

Error assessment focused on atmospheric thickness variables and the characteristics of the cold and warm layers (e.g., depths, positive/negative areas, maximum/minimum/mean temperatures for a layer).

## 3. Analysis and discussion

### a. Warm-nose precipitation events

Within the study region, 45 single-site or multi-site (coincidence between two or more ASOS/AWOS locations) WNPEs were identified for the 6-yr period, or an average of  $7.5 \text{ yr}^{-1}$  (Table 3). The locations of highest elevation (KBCB) and largest latitude (KIAD) have the greatest number of WNPEs during the 6-yr study period (Table 3). Nearly all of the WNPEs occurred in four of the six winter seasons studied, as few were identified for each of the 2015–2016 and 2016–2017 seasons (Table 3). For the WNPEs at each location and regionally, individual precipitation observations during the event were predominantly a combination of ice pellets and freezing rain, ranging from 44% (KIAD) to 57% (KBCB; Table 3). For all precipitation observations across the region associated with the 45 regional WNPEs, 51% were either ice pellets or freezing rain, 28% were snow, and 21% were rain (Table 3).

Of the 45 regional events, 20 were single site while only two involved all five ASOS/AWOS locations (Table 4). Aside from the lengthier five-site events, the mean per-site duration of a WNPE was between 12 and 15 h, with precipitation occurring during 55–70% of the event extent (Table 4). Mean per-site event-total precipitation generally ranged between 9 and 14 mm (0.35 and 0.55 in) with a mean intensity of around  $1 \text{ mm h}^{-1}$  ( $0.04 \text{ in h}^{-1}$ ; Table 4). While each WNPE inherently involves cold-air damming, approximately 71% (32) of the 45 WNPEs were associated with a low-pressure center/mid-latitude cyclone. These involved either passage to the west of the study area (10), or evolutions of type Miller-A (passage to the east from the southeastern United States; 11) or Miller-B (passage from west-to-east with redevelopment to the east; 11, see Table 4). Seven WNPEs, including the two that involved all five ASOS/AWOS locations (Table 4), were characterized primarily by an anticyclone to the north and classic synoptically driven cold-air damming—

**Table 3.** Number of warm-nose precipitation events, November through April, at each ASOS/AWOS site for the seasons 2013–2014 through 2018–2019. Region values represent the total of single- and multi-location (coincident at two or more locations) events across the study region as a whole. Also included for each location is the mean percentage of event precipitation observations as snow (SN), rain (RA), or either ice pellets (IP) or freezing rain (FZRA).

| Location      | Total     | 2013-14   | 2014-15   | 2015-16  | 2016-17  | 2017-18   | 2018-19  | SN        | RA        | IP/FZRA   |
|---------------|-----------|-----------|-----------|----------|----------|-----------|----------|-----------|-----------|-----------|
| BCB           | 27        | 6         | 6         | 1        | 2        | 4         | 8        | 34        | 9         | 57        |
| CHO           | 13        | 5         | 3         | 1        | 1        | 0         | 3        | 20        | 28        | 52        |
| GSO           | 19        | 6         | 3         | 2        | 0        | 4         | 4        | 21        | 27        | 52        |
| IAD           | 24        | 7         | 7         | 1        | 1        | 5         | 3        | 30        | 26        | 44        |
| ROA           | 18        | 5         | 4         | 1        | 1        | 3         | 4        | 22        | 29        | 49        |
| <b>Region</b> | <b>45</b> | <b>11</b> | <b>11</b> | <b>3</b> | <b>2</b> | <b>10</b> | <b>8</b> | <b>28</b> | <b>21</b> | <b>51</b> |

**Table 4.** For the 45 regional WNPE events evident at one or more ASOS/AWOS sites (*n* Sites), the number of events (*n* Events), and from all events in each category of *n* Sites, the per-site mean values of the event length, percent of event hours with precipitation (Percent *P*), total precipitation, and precipitation intensity. Also listed is the number of events within each category of *n*Sites associated with overrunning (OR), with cyclone passage to the west of the study region (West), with Miller-A or -B cyclone evolution, and with an unclassifiable synoptic pattern (None).

|                |                 | Mean Event Characteristics |                  |                     |                            | Synoptic Pattern Frequency |           |           |           |          |
|----------------|-----------------|----------------------------|------------------|---------------------|----------------------------|----------------------------|-----------|-----------|-----------|----------|
| <i>n</i> Sites | <i>n</i> Events | Length (h)                 | Percent <i>P</i> | Total <i>P</i> (mm) | <i>P</i> Intensity (mm/hr) | OR                         | West      | Miller-A  | Miller-B  | None     |
| 1              | 20              | 13.4                       | 62.6             | 11.2                | 0.9                        | 2                          | 4         | 3         | 8         | 3        |
| 2              | 7               | 12.2                       | 69.5             | 14.6                | 1.5                        | 0                          | 2         | 4         | 0         | 1        |
| 3              | 10              | 12.3                       | 55.6             | 8.8                 | 1.0                        | 2                          | 1         | 3         | 2         | 2        |
| 4              | 6               | 15.2                       | 63.6             | 13.4                | 1.2                        | 1                          | 3         | 1         | 1         | 0        |
| 5              | 2               | 21.5                       | 52.9             | 12.8                | 1.1                        | 2                          | 0         | 0         | 0         | 0        |
| <b>Total</b>   | <b>45</b>       | <b>14.3</b>                | <b>60.6</b>      | <b>11.7</b>         | <b>1.1</b>                 | <b>7</b>                   | <b>10</b> | <b>11</b> | <b>11</b> | <b>6</b> |

the closest evidence of a surface cyclone being far to the south of the study area—and overlaid by broad southerly isentropic lift, or overrunning. Six WNPEs were associated with a nondescript synoptic pattern, possibly characterized by a weak wave of low pressure or weak frontal passage. These were left unclassified (Table 4).

For the 101 potential WNPEs across the five ASOS/AWOS locations (Table 3), 88 coincided with  $\geq 1$  observed sounding with a distinct warm nose (nine without a warm nose, three missing soundings, and one without a cold surface layer). Some potential WNPEs had multiple soundings within their timeframes, yielding a total of 132 potential soundings for study. However, as soundings at KRNK and KIAD were linked with multiple ASOS/AWOS sites, multi-site events eliminated 37 duplicate soundings to produce a population of 95 soundings for potential analysis. Finally, the requirement of observed precipitation at

or within 1 h preceding the sounding time reduced the number of soundings for further study to 55 soundings: 19 at KRNK (representing events at KBCB, KCHO, and KROA), 17 at KGSO, and 19 at KIAD (representing events at KCHO and KIAD).

Composite soundings for each location depict the warm nose and the veering of low-level winds with height that is symbolic of warm-air advection (Figs. 3a–c). Mean values of sounding metrics at the three locations (Table 5) illustrate the relative strength of the warm nose compared to the surface-based cold layer. At both KGSO and KIAD, the depth of the warm nose is considerably greater than the depth of the cold air beneath (Table 5). The greater depth and/or greater difference from freezing for the warm nose relative to the cold layer at the three locations equates to positive area values (warm layer) that are 1.7 (KRNK) to 2.5 (KGSO, KIAD) times larger than negative-area values (cold layer; Table 5). Mean surface winds are from the

**Table 5.** Mean values from the observed soundings identified for forecast analysis at KRNK ( $n = 19$ ), KGSO ( $n = 17$ ), and KIAD ( $n = 19$ ). In addition to surface air temperature ( $T_a$ ), metrics for the surface-based cold layer and overlying warm layer are shown, including the negative and positive areas calculated using the Bourguoin method, along with thickness values for key pressure ranges. 1 kt = 0.5144 m s<sup>-1</sup>.

| Layer      | Metric                              | KRNK   | KGSO   | KIAD   |
|------------|-------------------------------------|--------|--------|--------|
| Surface    | Surface $T_a$ (°C)                  | -0.9   | -0.8   | -0.3   |
| Cool Layer | Minimum $T_a$ (°C)                  | -3.2   | -3.7   | -2.9   |
|            | Mean Layer $T_a$ (°C)               | -2.0   | -2.3   | -1.8   |
|            | Negative Area (J kg <sup>-1</sup> ) | 76.2   | 81.8   | 60.5   |
|            | Mean Layer Wind Direction (degrees) | 166.1  | 82.1   | 126.1  |
|            | Mean Layer Wind Speed (kt)          | 13.2   | 20.3   | 17.3   |
|            | Thickness (gpm)                     | 628.5  | 683.7  | 701.7  |
|            | Top Layer Pressure (hPa)            | 865.2  | 904.8  | 915.4  |
| Warm Layer | Maximum $T_a$ (°C)                  | 3.7    | 4.9    | 4.1    |
|            | Mean Layer $T_a$ (°C)               | 2.3    | 2.9    | 2.4    |
|            | Positive Area (J kg <sup>-1</sup> ) | 132.9  | 202.5  | 149.6  |
|            | Mean Layer Wind Direction (degrees) | 211.2  | 175.2  | 164.3  |
|            | Mean Layer Wind Speed (kt)          | 35.6   | 36.1   | 31.0   |
|            | Thickness (gpm)                     | 582.5  | 1395.0 | 1248.0 |
|            | Top Layer Pressure (hPa)            | 746.6  | 765.1  | 787.2  |
| Thickness  | 1000–500 hPa (gpm)                  | 5472.7 | 5474.8 | 5457.1 |
|            | 1000–700 hPa (gpm)                  | 2862.8 | 2867.4 | 2864.2 |
|            | 850–700 hPa (gpm)                   | 1561.8 | 1563.5 | 1560.6 |
|            | 1000–850 hPa (gpm)                  | 1301.1 | 1303.8 | 1303.6 |

northeast at each location, but the mean wind from the cold-to-warm layers veers from east to south (Figs. 3a–c; Table 5). Atmospheric thickness, reflecting mean temperature, is rather consistent across the three locations for the 1000–500-hPa portion of the atmosphere, and especially so for the partial layers of 1000–700, 850–700, and 1000–850 hPa (Table 5).

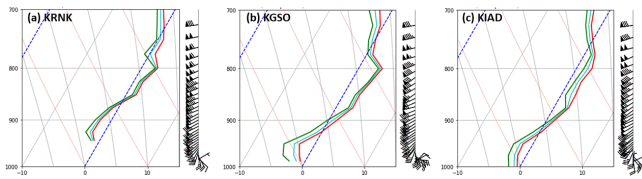
#### b. Observed versus forecast sounding analysis

For the 55 forecast soundings during WNPEs, each of the models generally simulates the thermal structure of the lower atmosphere well, but with thickness biases for the surface-based cold layer and the overlying warm nose. Because mean-layer temperature is skewed by error in depth, minimum cold-layer temperature and maximum warm-nose temperature are used to evaluate explicit temperature error. Commonly across the three models and three locations, surface air temperature and minimum/maximum air temperatures within the cold/warm layers are well forecast, with MPEs between 99 and 101% and standard deviations <1% (not shown). A slight warm bias is common, except for a cool bias

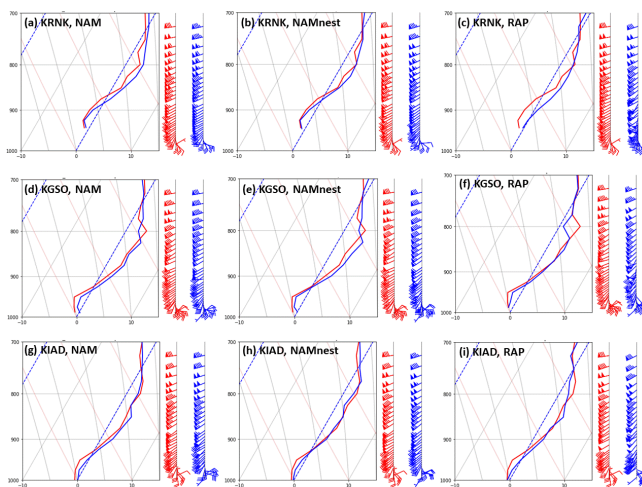
in maximum warm-nose temperature within the RAP model. Composite soundings (Fig. 4) for the +12-h runs of the NAM and NAMnest and +3-h cycle of the RAP (model runs prior to, but closest to, the observed sounding time) illustrate the mean accuracy of forecast thermal profiles. Note that differences in composite soundings are somewhat muted when computing mean values on a 25-hPa vertical grid from all soundings; but the relatively coarse vertical grid spacing is necessary to yield a robust computation of mean values at each pressure level.

Repeating the process for generating composite soundings, but focusing on the difference in forecast and observed air temperature (forecast minus observed), provides a clearer illustration of the vertical pattern of model error (Fig. 5). In terms of mean error, there is very good agreement above 800 hPa for each model at each location, while the variability in RAP forecast error (standard deviation of the temperature difference) is larger than that for the NAM or NAMnest (Fig. 5). At KRNK, the warm bias in NAM forecasts is most evident between 900 and 800 hPa (Fig. 5a), and while the same is true within NAMnest forecasts, the better



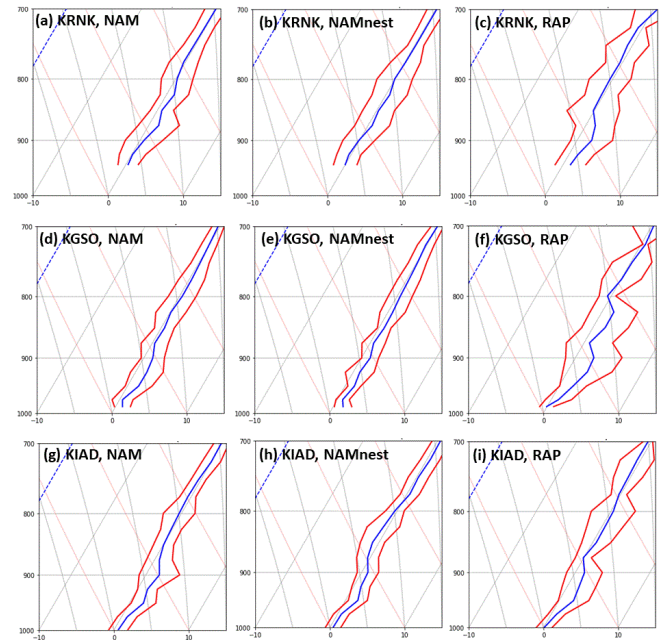


**Figure 3.** SkewT–log $p$  diagrams that illustrate composite soundings through the lower atmosphere (25-mb increment) for WNPEs at (a) KRNK ( $n = 19$ ), (b) KGSO ( $n = 17$ ), and (c) KIAD ( $n = 19$ ). Red lines represent air temperature ( $^{\circ}\text{C}$ ), green lines represent dewpoint temperature, and blue lines represent wet-bulb temperature. Wind direction and speed (kt;  $1 \text{ kt} = 0.5144 \text{ m s}^{-1}$ ) are indicated to the right of each diagram.



**Figure 4.** Composite observed (red) and forecast (blue) air temperature and wind speed/direction through 700 hPa for WNPEs at each location and each model. Forecasts for the NAM and NAMnest are for the model run 12 h before the observed sounding time, while composites for the RAP are for the model run 3 h before the observed sounding time.

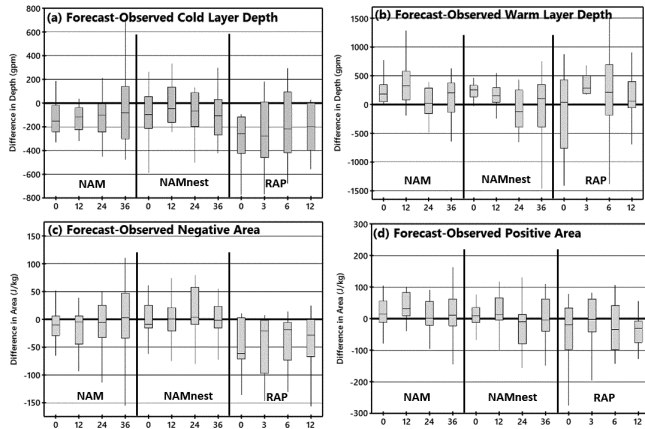
mean accuracy of NAMnest forecasts is evident (Fig. 5b). A mean warm bias in RAP forecasts at KRNK is evident between the surface and 850 hPa (Fig. 5c). At KGSO, a warm bias is evident from the surface up to about 875 hPa for NAM forecasts (Fig. 5d) and from the surface up to about 800 hPa for NAMnest forecasts (Fig. 5e), although the mean and variability in error is generally smaller for the NAMnest. A much larger variability in error and a warm bias extending from the surface up to 800 hPa is evident for RAP forecasts at KGSO (Fig. 5f). The patterns of larger variability in error for RAP forecasts and better mean accuracy with NAMnest forecasts also are evident at KIAD (Figs. 5g–i). As at KGSO, the warm bias in NAM forecasts



**Figure 5.** Mean of the difference in forecast and observed air temperature (blue line; forecast minus observed) through 700 hPa for WNPEs at each location and each model. Forecasts for the NAM and NAMnest are for the model run 12 h before the observed sounding time, while composites for the RAP are for the model run 3 h before the observed sounding time. Also shown are the temperature differences plus/minus one standard deviation (red lines).

at KIAD is evident from the surface to about 850 hPa (Fig. 5g), while that for RAP forecasts extends from the surface to about 800 hPa (Fig. 5i). Unlike for any other model–location combination, NAMnest forecasts at KIAD have a mean cool bias between about 875 to 800 hPa, but with the more typical warm bias beneath this layer (Fig. 5h).

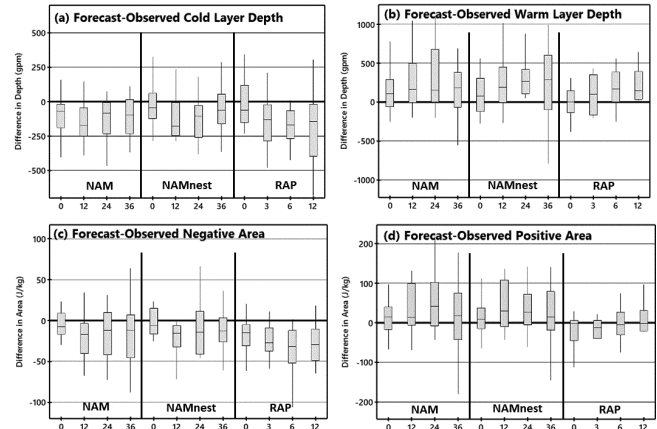
It is common for the surface-based cold layer to be too thin, and the warm nose to be too thick, across the three models and three locations. In some cases, errors decrease with shorter lead times (e.g., RAP warm-nose depth at KIAD), but in many cases, the biases remain large or even increase at the analysis time (e.g., NAM warm-nose depth at KIAD; Figs. 6–8). For WNPEs at KRNK, the surface-based cold layer is consistently too thin for all the models at all model forecast times (Fig. 6a). The error in RAP model forecasts is generally the largest of the three models, and arguably largest at the time of the observed sounding (0 h; Fig. 6a). For the NAM, the cold layer tends to be too thin at all forecast lead times (Fig. 6a). The forecast warm nose at KRNK



**Figure 6.** Box plots characterizing the population of forecast minus observed values for (a) cold and (b) warm layer depth (gpm), and (c) negative and (d) positive area ( $J kg^{-1}$ ) at KRNK for the NAM, NAMnest, and RAP models. The four model runs for the NAM and NAMnest (0 h, +12 h, +24 h, +36 h) and RAP (0 h, +3 h, +6 h, +12 h) models are ordered left-to-right from latest (0 h) to earliest (+36 h or +12 h) on the horizontal axis. Box plot whiskers represent the 10th and 90th percentiles.

is generally too thick, particularly at +12 h and nearer the observed sounding for the NAM and NAMnest, and for +3- and +6-h forecast soundings for the RAP, which also exhibits high variability in the error of forecast depth (Fig. 6b). Forecast errors for layer depth are realized in forecasts of negative and positive areas. For the negative area, variability in error associated with the NAM improves with model runs to an area that is slightly too small at 0 and 12 h before the observed sounding (Fig. 6c). The NAMnest generally forecasts the negative area better than the NAM, particularly in terms of reduced variability in error, while RAP forecasts of negative area are consistently too small and with a high variability in error (Fig. 6c). For the positive area at KRNK, NAM forecasts are only slightly too large, particularly at 12 h before the observed sounding, while NAMnest forecasts are too small initially and too large closer to the observed sounding time, but with a distinct reduction in error variability (Fig. 6d). In contrast, RAP forecasts of positive area are generally too small (Fig. 6d), owing to a cool bias in maximum warm-layer temperature that supersedes a thick bias in warm-layer depth. In addition to the small bias in the positive area, the error in RAP forecasts of the positive area is highly variable (Fig. 6d).

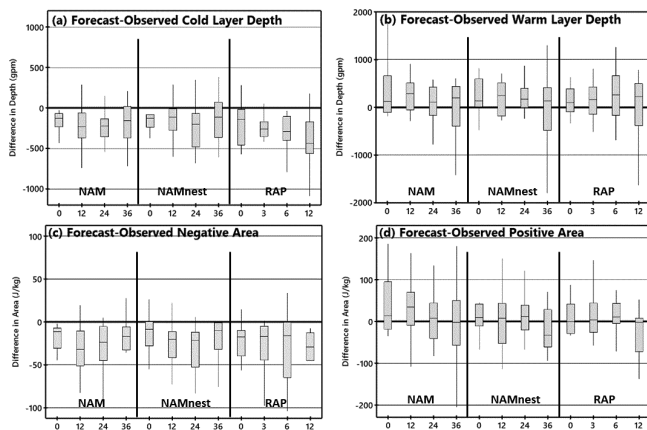
The general characteristics of model bias at KRNK—a cold layer/negative area that is too thin/



**Figure 7.** Same as Fig. 6, but for KGSO.

small and a warm nose/positive area that is too thick/large—are also evident at KGSO. Forecasts of the surface-based cold layer are consistently too thin for all three models, although RAP forecasts improve at the time of the observed sounding (0 h; Fig. 7a). Each model tends to forecast a warm nose that is too thick, but with improvement with model runs for the NAMnest, and improvement at the observed sounding time (0 h) for the NAM and especially for the RAP (Fig. 7b). Negative area forecasts for all three models are too small, but there is some improvement with shorter lead times for the NAM and NAMnest, and modest improvement through model runs for the RAP (Fig. 7c). Forecasts of the positive area are too large for the NAM and NAMnest, and each improves with shorter lead times (Fig. 7d). Similar to KRNK, RAP model forecasts of positive area at KGSO are slightly too small, owing to a cool bias in maximum warm-nose temperature; but unlike at KRNK, variability in model error is generally smaller than for the other models (Fig. 7d).

Forecast errors at KIAD generally conform to those at KRNK and KGSO, albeit with a more pronounced thin bias within the surface-based cold layer than at the other locations. Forecasts of cold-layer depth are consistently too thin, but variability in forecast error improves approaching the observed sounding time (Fig. 8a). Forecasts of the warm nose are too thick, with some improvement closer to the observed sounding time for the RAP model (Fig. 8b), which more accurately portrays warm-layer depth than cold-layer depth. Forecasts of the negative area at KIAD are consistently too small across all three models (Fig. 8c), while forecast biases in the positive area at KIAD are a bit more variable. NAM forecasts of the positive area are too large at +12 h and closer to the observed sounding time (Fig.



**Figure 8.** Same as Fig. 6, but for KIAD.

8d). NAMnest forecasts of positive area are too small early within the model cycle (+36 h), then slightly too large but with a lessening degree of variability in model error (Fig. 8d). RAP model forecasts of positive area at KIAD are generally more accurate than for the NAM and NAMnest, particularly at +6 h and closer to the observed sounding time (Fig. 8d). Unlike at KRNK and KGSO, RAP forecasts of positive area are too large nearer the sounding time, meaning that the warm-layer thick bias (Fig. 8b) outweighs the very slight cool bias in maximum warm-nose temperature.

Counts of large  $\pm$  biases in the thicknesses of the cold and warm layers illustrate skewness in the forecasts. We define a large model bias as an absolute error magnitude greater than the MAE. Focusing on the forecast model run nearest the observed sounding time for each model (+12 h for NAM and NAMnest, and +3 h for RAP), the thin bias in the surface-based cold layer in each model and the thick bias in the warm nose within the NAM and NAMnest are obvious (Table 6). For the three sounding locations taken together, 96% of the large NAM biases in the depth of surface-based cold layer are too thin, while 95% of large biases in warm-nose depth are too thick (Table 6). While still evident, the contrast is not quite as stark for the NAMnest, with 70% of large biases in the depth of cold air being too thin, and 75% of large biases in the warm nose as too thick (Table 6). For the RAP model, 94% of large biases in surface-based cold-layer depth are too thin, but large biases in warm-nose depth are near evenly split between too thick (58%) and too thin (42%; Table 6).

Counts of large biases in negative and positive areas reinforce those for layer depths. Collectively, for the three sounding locations, 86% of large NAM errors in negative area are for an area that is too small, and 86% of large errors in positive area are for an area that is too

large (Table 7). As with layer-depth biases, large errors in negative and positive areas within the NAMnest are slightly less skewed than for the NAM, but the contrasts are still very obvious. Of the large errors in negative area forecasts by the NAMnest, 76% are for an area that is too small, and of the large errors in positive area, 79% are for an area that is too large (Table 7). Large RAP model errors in the negative area are highly skewed toward an area that is too small (95% of the cases; Table 7); but large errors within RAP forecasts of positive area are rather evenly split between too large (54%) and too small (46%; Table 7).

Comparison of forecast and observed partial atmospheric thicknesses indicate larger errors lower within the cold-layer/warm-nose thermal structure. A larger warm bias within the RAP model at KRNK is evident in atmospheric-thickness biases for the deeper 1000–500-hPa layer and for the lower layers of 1000–700 and 1000–800 hPa (Fig. 9a). However, the better performance of the RAP model within the warm nose is apparent in the accuracy of 850–700-hPa thickness forecasts at KRNK (Fig. 9a). This also is the case for the RAP model at KGSO (Fig. 9b) and for all three models at KIAD (Fig. 9c), indicating that temperature across essentially the top half of the warm nose is rather well forecast. More problematic for all models at all locations is the warm bias illustrated by thickness forecasts for the lower layers of 1000–700 and 1000–800 hPa, which translates to pronounced error in the deeper 1000–500-hPa layer for all models at each of the locations (Fig. 9).

Composite data for observed and forecast wind speed/direction (not explicitly shown, but illustrated in Fig. 4) do not depict clear errors in wind to accompany the temperature error within the models. However, upon creating synoptic atmospheric composites associated with large warm model bias, it appears that error in the thickness of the surface-based cold layer and the overlying warm nose correspond differently to the strength of observed vertical motion. Using gridded NCEP/NCAR daily reanalysis data (Kalnay et al. 1996), mapped composites of mean observed daily vertical velocity in isobaric coordinates ( $\omega$ , Pa  $s^{-1}$ ) at the 850-hPa level were created for instances that coincided with large model errors in cold- and warm-layer thermal characteristics. These were contrasted with composites of vertical velocity coinciding with instances of relatively accurate forecasts of the cold and warm layers. Instances of large model error are defined as before—absolute error in the negative

**Table 6.** Frequency of large model error in the depth of the cold and warm layers within the vertical atmospheric sounding at KRNK, KGSO, and KIAD for the 12-h lead-time runs of the NAM and NAMnest models and for the 3-h lead-time run of the RAP model. Values represent frequency with which the depth exceeds the MAE magnitude in the positive (too thick) and negative (too thin) directions.

|       |       | NAM  |      | NAMnest |      | RAP  |      | Total |      |
|-------|-------|------|------|---------|------|------|------|-------|------|
|       |       | Cold | Warm | Cold    | Warm | Cold | Warm | Cold  | Warm |
| KRNK  | Thick | 0    | 8    | 4       | 4    | 0    | 4    | 4     | 16   |
|       | Thin  | 8    | 0    | 6       | 1    | 4    | 1    | 18    | 2    |
| KGSO  | Thick | 0    | 6    | 2       | 6    | 1    | 0    | 3     | 12   |
|       | Thin  | 9    | 1    | 8       | 0    | 5    | 2    | 22    | 3    |
| KIAD  | Thick | 1    | 6    | 2       | 2    | 1    | 3    | 4     | 11   |
|       | Thin  | 6    | 0    | 5       | 3    | 7    | 2    | 18    | 5    |
| Total | Thick | 1    | 20   | 8       | 12   | 1    | 7    | 11    | 39   |
|       | Thin  | 23   | 1    | 19      | 4    | 16   | 5    | 58    | 10   |

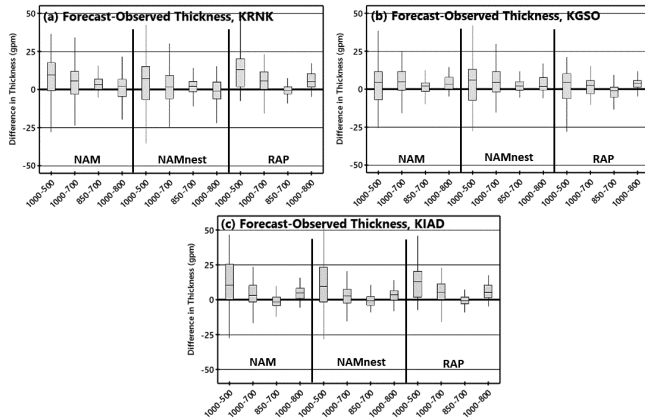
**Table 7.** Frequency of large model error in negative and positive areas of the vertical atmospheric sounding at KRNK, KGSO, and KIAD for the 12-h lead-time runs of the NAM and NAMnest models and for the 3-h lead-time run of the RAP model. Values represent frequency with which the area exceeds the MAE magnitude in the positive (too large) and negative (too small) directions.

|       |       | NAM  |      | NAMnest |      | RAP  |      | Total |      |
|-------|-------|------|------|---------|------|------|------|-------|------|
|       |       | Cold | Warm | Cold    | Warm | Cold | Warm | Cold  | Warm |
| KRNK  | Large | 2    | 7    | 3       | 8    | 0    | 3    | 5     | 18   |
|       | Small | 4    | 0    | 4       | 1    | 5    | 2    | 13    | 3    |
| KGSO  | Large | 1    | 6    | 1       | 5    | 0    | 1    | 2     | 12   |
|       | Small | 6    | 1    | 5       | 0    | 8    | 3    | 19    | 4    |
| KIAD  | Large | 0    | 5    | 1       | 2    | 1    | 3    | 2     | 10   |
|       | Small | 8    | 2    | 7       | 3    | 5    | 1    | 20    | 6    |
| Total | Large | 3    | 18   | 5       | 15   | 1    | 7    | 9     | 40   |
|       | Small | 18   | 3    | 16      | 4    | 18   | 6    | 52    | 13   |

area (cold layer) or positive area (warm nose) greater than the MAE. These are distinguished from all other instances, meaning an error in positive or negative area less than the MAE. Eliminating duplicate dates (coincident across the three models or three locations) yielded 19 instances of large error in negative area (cold layer), in which the forecast is too small, contrasted with 30 instances of little error in forecast negative area. Similarly, composites relating to the positive area (warm nose) are based on 21 instances of large error, in which the forecast is too large, and 35 instances of little error in forecast positive area.

Composite omega fields are characterized by negative values across the study region for large error in both the cold layer (Fig. 10a) and warm nose (Fig. 10c), indicating the expected upward motion (decreasing pressure) associated with the precipitation events studied. Observed upward motion is considerably

stronger when there is large error in the cold-layer forecast (Fig. 10a) compared to large error in the warm-layer forecast (Fig. 10c). Reinforcing this are the negative differences in observed omega between large error and little error in the cold-layer forecast (Fig. 10b) that are in contrast to the positive differences associated with the warm layer (Fig. 10d). Considering the negative omega values of the observed vertical motion field across the region, the negative differences in omega for the cold layer (Fig. 10b) indicate a more strongly negative omega associated with large error. In other words, the models have a tendency toward a warm bias within the cold layer under stronger synoptic-scale upward motion. The positive differences in observed omega in the warm layer (Fig. 10d) indicate a weaker negative omega associated with large error, indicating that the models have a tendency toward a warm bias in the warm nose under weaker synoptic-

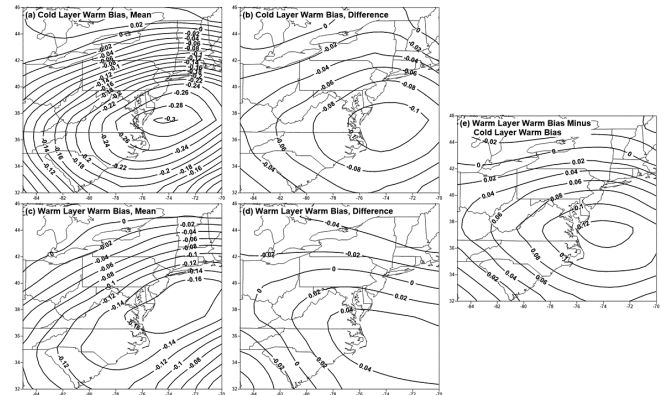


**Figure 9.** Box plots characterizing the population of forecast minus observed values for atmospheric thickness at (a) KRNK, (b) KGSO, and (c) KIAD for the +12-h model runs of the NAM and NAMnest and +3-h run of the RAP model. Thicknesses from 1000–500, 1000–700, 850–700, and 1000–800 hPa are ordered from left-to-right on the horizontal axis. Box plot whiskers represent the 10th and 90th percentiles.

scale upward motion. This dichotomy in the strength of observed upward motion corresponding with error in the cold- and warm-layer forecasts is underscored in the composite difference in omega associated with each. When subtracting the composite omega associated with large error in the cold layer from that for the warm layer (Fig. 10e), the positive values indicate the stronger observed upward motion in cases of cold-layer error.

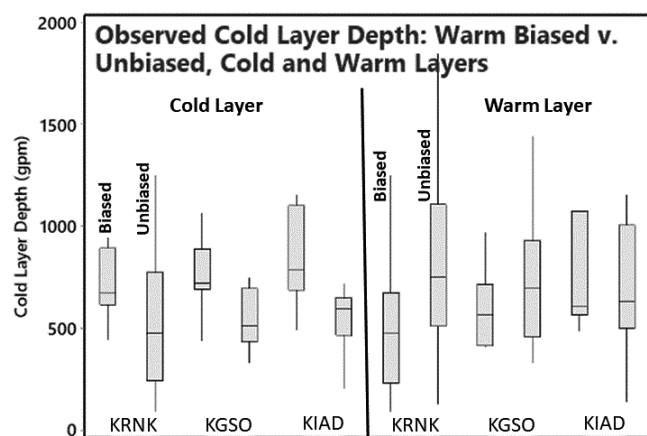
That model errors, regionally, tend to correspond with a different strength of observed upward motion depending upon whether the error is in the cold or warm layer reinforces the finding that warm bias in the two layers is usually not coincident. Regionally and across all models, of the 92 instances of large error in a positive area too large ( $n = 52$ ) and a negative area too small ( $n = 40$ ; Table 7), only 12 (13%) coincided. In other words, 40 of 52 (77%) large errors in the warm layer did not coincide with a large error in the cold layer, and 28 of 40 (70%) large errors in the cold layer did not coincide with a large error in the warm layer.

The difference in the strength of upward motion per layer of bias seems to align with the stratification of large model error by synoptic pattern, particularly for the Miller-B type cyclone evolution and for overrunning. The Miller-B pattern was associated with many more cases of a large warm bias in the cold layer (16) than in the warm layer (4; Table 8). The mid-latitude cyclone of the Miller-B pattern should yield stronger synoptic-scale upward motion (symptomatic of a cold-layer bias)



**Table 8.** Segregated by synoptic pattern, the frequency of large absolute error in negative area (too warm/shallow cold layer) and positive area (too warm/deep warm layer) for the 12-h lead-time runs of the NAM and NAMnest models and for the 3-h lead-time run of the RAP model at all locations combined (Table 7). Values are percentages of the total number of large errors of each synoptic pattern, and parenthetical numbers are raw counts.

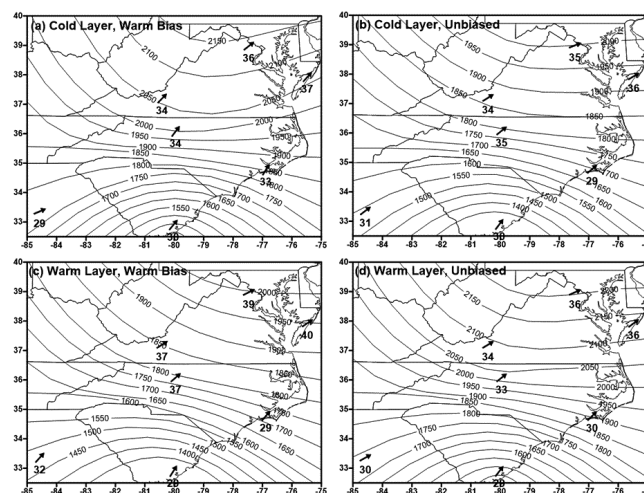
| Synoptic Pattern | Negative Area, Small Bias | Positive Area, Large Bias |
|------------------|---------------------------|---------------------------|
| Overrunning      | 11.5 (6)                  | 25.0 (10)                 |
| Western L        | 30.8 (16)                 | 25.0 (10)                 |
| Miller-A         | 23.1 (12)                 | 35.0 (14)                 |
| Miller-B         | 30.8 (16)                 | 10.0 (4)                  |
| Unclassified     | 3.8 (2)                   | 5.0 (2)                   |
| <b>Total</b>     | <b>100 (52)</b>           | <b>100 (40)</b>           |



**Figure 11.** Box plots characterizing the population of observed surface-based cold-layer depth associated with warm forecast bias and little forecast bias in the cold and warm layers at KRNK, KGSO, and KIAD. Data correspond to biases within the +12-h model runs of the NAM and NAMnest and +3-h run of the RAP model, collectively. Box plot whiskers represent the 10th and 90th percentiles.

of little error—most obviously at KRNK, but also at KGSO, while only marginally so at KIAD (Fig. 11). The shallower observed cold-layer depth coinciding with warm bias in the warm nose corresponds to the weaker observed upward motion outlined earlier.

In summary, instances of a cold-layer forecast as too shallow generally coincide with observed stronger upward motion over a deeper cold layer, while instances of a warm-nose forecast as too thick generally coincide with observed weaker upward motion over a shallower cold layer. These findings are collectively expressed in synoptic composites of the 295K isentropic surface derived from sounding data from seven sounding sites across and to the south of the study area (Fig. 12). The same warm-bias and little-bias dates for the three



**Figure 12.** For instances of large warm forecast bias and little forecast bias in cold (a, b) and warm (c, d) atmospheric layers, composite means of the observed 295K isentropic surface elevation (gpm) and wind speed (knots)/directional vector. Elevation data are derived from the 295K elevation at the seven sounding locations for which wind speed and direction are shown. Data correspond to biases within the +12-h runs of the NAM and NAMnest models and +3-h model run of the RAP, collectively, for all models regionally.

models and three study locations (collectively) that were used to create composite omega maps (Fig. 12) were used to create composite maps of the observed 295K height and wind direction/speed (Fig. 12). There is little difference in observed wind direction and speed for the 295K isentropic surface between instances of large error and little error in either layer (Fig. 12). However, the observed isentropic surface across the mid-Atlantic study area is positioned higher in the atmosphere for instances of thin bias in surface-based cold-layer forecasts (Fig. 12a) compared to little error (Fig. 12b). The opposite is the case for the warm nose;

the observed isentropic surface across the mid-Atlantic is positioned lower in the atmosphere for instances of a thick forecast bias (Fig. 12c) compared to little error (Fig. 12d). While not excluding other dynamical processes associated with vertical motion, it appears that the common warm model bias in the cold and warm layers of the WNPEs is related to the degree of isentropic lift occurring across the region. It is intuitive that model forecasts err toward a cold layer that is too thin for those study cases possessing the thicker cold layers, and that they err toward a warm layer that is too thick for those study cases possessing the thinner cold layers. This characterizes the models as too conservative in their forecasts of the thickest and thinnest cold layers for the events studied. But without deeper study of model physics, it is difficult to conclude whether the differences in apparent isentropic lift are an inherent coincidence with model error, or a cause of model error.

While forecast biases within the thermal profile during WNPEs are important, the accuracy with which the resulting precipitation type is forecast is ultimately most important for forecasters. Using the original Bourgoïn (2000) method, we estimated precipitation type for the observed and forecast soundings studied, but we re-rendered precipitation type as rain when the surface air temperature was above freezing. Consolidating the results from the three locations yields 52 forecast–observed precipitation-type comparisons for both the NAM and NAMnest models, and 45 for the RAP model. There is agreement in precipitation type for approximately two-thirds of the cases for each model (63.5% for the NAM, 67.3% for the NAMnest, and 64.4% for the RAP). In terms of the nature of bias in precipitation-type forecast, the NAM model is exclusively warm-biased with precipitation type (36.5% of all cases), while the NAMnest and RAP models are nearly so. For the NAMnest model, 28.9% of the 52 cases are warm-biased and 3.8% are cold-biased, while for the RAP model, 26.7% of the 45 cases are warm-biased and 8.9% are cold-biased.

#### 4. Concluding discussion

The NAM, NAMnest, and RAP models adequately forecasted the general thermal structure of 55 atmospheric profiles from 45 warm-nose winter precipitation events occurring within the mid-Atlantic study region. The minimum air temperature within a surface-based cold layer, and the maximum air temperature within an elevated warm layer (or warm

nose), were both well forecast by the models. However, each of the models exhibited a warm bias that is evident in layer thickness—a cold layer too thin for each of the models, and a warm nose too thick, particularly within NAM and NAMnest forecasts. A well-forecast thermal structure (but with a warm bias) matches the findings of Ikeda et al. (2017) in their study of the HRRR model.

Large warm biases in each of the cold and warm layers, in terms of negative and positive areas, infrequently coincided, indicating that model error is not simply a deep-layer warm bias, but rather that it commonly is specific to thermal layer. Differences between forecast and observed atmospheric-layer thickness suggest this as well. While differences in the broader 1000–500-hPa thickness depict the general warm bias within the models, the upper portion of the warm layer, represented by the thickness of the 850–700-hPa layer, is adequately forecast, particularly for the lower-elevation locations studied. Generally, the warm bias stems from the lower portion of the warm layer and below.

Comparison of observed and forecast wind velocity and direction revealed no obvious error for either of the models. However, synoptic-scale composites of observed vertical-velocity fields indicate that the instances of large warm model bias within the cold layer (too thin) tended to coincide with an observed atmosphere characterized by stronger upward motion across the region, while instances of large warm bias within the warm nose (too thick) tended to coincide with weaker observed upward motion across the region. This provides a basis for some speculation of the source of model error, with the degree of isentropic lift representing one possible mechanism by which model errors emerge. The stronger synoptic-scale vertical motion that coincides with the warm bias in the cold surface layer generally occurs when a deeper cold layer and higher isentropic surface are in place. In contrast, the weaker vertical motion that coincides with the warm bias in the elevated warm layer generally occurs when a shallower cold layer and lower isentropic surface are in place across the mid-Atlantic region.

Whether the association of model bias and vertical velocity stems from isentropic lift or dynamic forcing above, it is possible that model physics related to precipitation plays a role in layer-depth error. The varying degree of upward motion associated with error in the cold and warm portions of the atmosphere is likely associated with varying precipitation rates, from which the models may overestimate or underestimate

the degree to which warm air is brought downward from the elevated warm layer and/or the degree of evaporative cooling that is occurring. The latter is an error source suggested by Ikeda et al. (2017) in their study of the performance of the HRRR model. That large warm model bias in the cold layer (too thin) coincides with stronger upward motion across the mid-Atlantic region may mean a greater precipitation rate and larger degree of evaporative cooling that the models inadequately represent, yielding an insufficiently thick cold layer. As large warm model bias in the warm layer (too thick) coincides with weaker upward motion, a lesser precipitation rate may be the result, in which case the models may overestimate precipitation intensity and the degree to which warm air aloft is brought downward to yield a warm nose that is too thick. A definitive answer requires testing the model physics, but the results of this research can inform model-refinement efforts while offering guidance to forecasters within the mid-Atlantic region.

Within a few years, the NAM, NAMnest, and RAP models will be replaced by a different suite of high-resolution models as part of the NWS's Finite Volume Cubed-Sphere (FV3) dynamical core. To what extent the conclusions of this research will still apply is unknown. But, as the models will remain frozen in the interim, forecasters in the mid-Atlantic region may consider the inherent warm bias in model forecasts of thermal profile and precipitation type. Future work should investigate whether this warm bias persists in models with the FV3 core and, if so, whether the mechanisms might be similar to those identified in this study.

*Acknowledgments.* This work was supported by NOAA under award number NA16NWS4670042 and the University Corporation for Atmospheric Research (UCAR) Cooperative Program for Operational Meteorology, Education, and Training (COMET) program under sub-award number SUBAWD001737. The authors thank Ms. Katie Slusher and Mr. Howard Harris for their help with data processing at the outset of this research, and Mr. Jeff Waldstreicher (NWS Eastern Region Headquarters) for guidance in improving the initial manuscript.

## REFERENCES

- Birk, K., E. Lenning, K. Donofrio, and M. T. Friedlein, 2021: A revised Bourgoquin precipitation-type algorithm. *Wea. Forecasting*, **36**, 425–438, [Crossref](#).
- Bourgoquin, P., 2000: A method to determine precipitation types. *Wea. Forecasting*, **15**, 583–592, [Crossref](#).
- Changnon, S. A., 2003: Characteristics of ice storms in the United States. *J. Appl. Meteor.*, **42**, 630–639, [Crossref](#).
- \_\_\_\_\_, and T. R. Karl, 2003: Temporal and spatial variation in freezing rain in the contiguous United States: 1948–2000. *J. Appl. Meteor.*, **42**, 1302–1315, [Crossref](#).
- Craven, J. P., D. E. Rudack, and P. E. Shafer, 2020: National Blend of Models: A statistically post-processed multi-model ensemble. *J. Operational Meteor.*, **8**, 1–14, [Crossref](#).
- 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. *Mon. Wea. Rev.*, **115**, 564–591, [Crossref](#).
- Gendreau, H., cited 2021: The ice storm cometh, and leaves darkness in its wake. *Roanoke Times*, 15 November 2018. [Available online at [https://www.roanoke.com/news/local/the-ice-storm-cometh-and-leaves-darkness-in-its-wake/article\\_4b2a9442-e22d-58ac-ab96-9d2fba9e0874.html](https://www.roanoke.com/news/local/the-ice-storm-cometh-and-leaves-darkness-in-its-wake/article_4b2a9442-e22d-58ac-ab96-9d2fba9e0874.html).]
- Hux, J. D., P. C. Knappenberger, P. J. Michaels, and P. J. Stenger, 2001: Development of a discriminant analysis mixed precipitation (DAMP) forecast model for mid-Atlantic winter storms. *Wea. Forecasting*, **16**, 248–259, [Crossref](#).
- Ikeda, K., M. Steiner, J. Pinto, and C. Alexander, 2013: Evaluation of cold-season precipitation forecasts generated by the hourly updating High-Resolution Rapid Refresh model. *Wea. Forecasting*, **28**, 921–939, [Crossref](#).
- \_\_\_\_\_, \_\_\_\_\_, and G. Thompson, 2017: Examination of mixed-phase precipitation forecasts from the High-Resolution Rapid Refresh model using surface observations and sounding data. *Wea. Forecasting*, **32**, 949–967, [Crossref](#).
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [Crossref](#).
- Keeter, K. 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. *Wea. Forecasting*, **10**, 42–60, [Crossref](#).
- Landolt, S. D., J. S. Lave, D. Johnson, and A. Gaydos, 2019: The impacts of automation on present weather-type observing capabilities across the conterminous United States. *J. Appl. Meteor. Climatol.*, **58**, 2699–2715, [Crossref](#).



- Mahoney, E. A., T. A. Niziol, 1997: BUFKIT: A software application toolkit for predicting lake effect snow. Preprints, *13th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*. Long Beach, CA, Amer. Meteor. Soc., 388–391.
- McCray, C. D., E. H. Atallah, and J. R. Gyakum, 2019: Long-duration freezing rain events over North America: Regional climatology and thermodynamic evolution. *Wea. Forecasting*, **34**, 665–681, [Crossref](#).
- \_\_\_\_\_, J. R. Gyakum, and E. H. Atallah, 2020: Regional thermodynamic characteristics distinguishing long- and short-duration freezing rain events over North America. *Wea. Forecasting*, **35**, 657–671, [Crossref](#).
- Miller, J. E., 1946: Cyclogenesis in the Atlantic coastal region of the United States. *J. Meteor.*, **3**, 31–44, [Crossref](#).
- NCEP, 2020: Product description document: NCEP Model Analyses & Guidance. September 2020. 18 pp. [Available online at [https://mageval.ncep.noaa.gov/docs/NCEP\\_PDD\\_MAG.pdf](https://mageval.ncep.noaa.gov/docs/NCEP_PDD_MAG.pdf).]
- Ralph, F. M., and Coauthors, 2005: Improving short-term (0–48 h) cool-season quantitative precipitation forecasting. *Bull. Amer. Meteor. Soc.*, **86**, 1619–1632, [Crossref](#).
- Rauber, R. M., L. S. Olthoff, M. K. Ramamurthy, D. Miller, and K. E. Kunkel, 2001: A synoptic weather pattern and sounding-based climatology of freezing precipitation in the United States east of the Rocky Mountains. *J. Appl. Meteor.*, **40**, 1724–1747, [Crossref](#).
- Reeves, H. D., K. L. Elmore, A. Ryzhkov, T. Schmur, and J. Krause, 2014: Sources of uncertainty in precipitation-type forecasting. *Wea. Forecasting*, **29**, 936–953, [Crossref](#).
- Robins, C. C., and J. V. Cortinas Jr., 2002: Local and synoptic environments associated with freezing rain in the contiguous United States. *Wea. Forecasting*, **17**, 47–65, [Crossref](#).
- Wandishin, M. S., M. E. Baldwin, S. L. Mullen, and J. V. Cortinas Jr., 2005: Short-range ensemble forecasts of precipitation type. *Wea. Forecasting*, **20**, 609–626, [Crossref](#).