THESIS

# ENSEMBLE-BASED ANALYSES OF LIMINAL EXTREME RAINFALL EVENTS NEAR TAIWAN AND NORTHERN COLORADO

Submitted by
Alexandra S. Cole
Department of Atmospheric Science

In partial fulfillment of the requirements
For the Degree of Master of Science
Colorado State University
Fort Collins, Colorado
Summer 2022

Master's Committee:

Advisor: Michael M. Bell

Kristen Rasmussen
Peter Nelson

ABSTRACT

ENSEMBLE-BASED ANALYSES OF LIMINAL EXTREME RAINFALL EVENTS NEAR TAIWAN AND NORTHERN COLORADO

Heavy rainfall is a phenomenon that impacts a variety of climates around the world, from the moisture-rich, tropical northwestern Pacific to the drier Northern Colorado plains. Improvements over decades of numerical weather prediction have allowed for increased accuracy in simulations of heavy rainfall cases, but there are still improvements yet to be made. This thesis, in conjunction with the Prediction of Rainfall Extremes Campaign in the Pacific (PRECIP) field campaign, aims to study the mechanisms behind these heavy rainfall events to increase understanding of their underlying processes and improve the modeling of them. Two weather events are investigated in detail, one in which heavy rainfall was not forecast by global models but greater than 600 mm of rainfall accumulated, and a contrasting case in which heavy rainfall was forecast but little to no rainfall accumulated.

On 09 June 2020 near Taiwan, heavy rainfall was produced by quasi-stationary back-building mesoscale convective systems (MCS) associated with a mei-yu front. Peak rainfall amounts totaled over 600 mm with widespread rainfall totals greater than 100 mm. Global model forecast skill was poor in both location and intensity of rainfall. The mesoscale ensemble showed liminal conditions between heavy rainfall or little to no rainfall. The two most accurate and two least accurate ensemble members are selected for analysis via validation against radar-estimated rainfall observations. All members feature moisture-rich environments and moist neutral soundings with low levels of free convection (LFC) and sufficient instability for deep convection, and the synoptic setups do not suggest such different outcomes. Through our analysis, we find that stronger gradients in 100 m virtual potential temperature in the two most accurate members associated with a near-surface frontal boundary provide the primary lifting mechanism for enhanced rainfall. In the two heaviest rain-producing members, air moves north/northeastward and ascends the virtual potential temperature isentropes and rises above the LFC, producing back-building deep, moist convection. The near-surface gradients are weaker and more confined along Taiwan's coast in the two least accurate members, which leads to less rainfall that is misplaced from reality. The analyses suggest that subtle details in the simulation of frontal boundaries and meso-scale flow structures can lead to bifurcations in producing extreme or almost no rainfall.

A contrasting event occurred in Northern Colorado on 31 July 2021, where heavy rainfall was forecast and flood warnings were issued, but little to no rainfall and flooding took place in the forecast area. Synoptic and mesoscale conditions were ripe for heavy rainfall, with anomalously high precipitable water values and moderate values of CAPE. Similar to the 09 June 2020 case, the mesoscale ensemble showed a wide spread in rainfall totals, related in part to the variability of surface boundaries and forcing across the ensemble. Weak surface forcing led to very little rainfall in this case despite the high moisture, suggesting similar physical mechanisms and predictability challenges across both the analyzed cases. Implications for improved probabilistic forecasts, increased forecast accuracy, and thus increased public safety for heavy rainfall events are discussed.

ACKNOWLEDGMENTS

The research contained within this thesis would not have been possible without the support of a large number of people who guided and supported me. I would first like to thank my advisor, Michael Bell, for all of his guidance and insight through this process, particularly in the face of a pandemic, which changed the course of this research early on. His support and understanding were instrumental in completing this work, and I am appreciative of the environment he created. I have members of the Bell group to thank for this environment and support, as well, especially Jennifer DeHart, who acted as a close mentor throughout this work. I would also like to thank my committee members, Kristen Rasmussen and Peter Nelson, for their time and support in preparing this thesis.

I would like to thank my family for supporting me in my academic endeavors from day one. Specifically, I owe a lot to my dad who helped me with my math homework, until it became too challenging, and who has edited every essay I have written, no matter the subject. I could not have done it without my sister, Tori, and mom, Sandy, who provided support and encouragement in an unfamiliar process, nor my partner, Ben, who provided support and encouragement in a quite familiar process. All of the mentoring and advice I received from my undergraduate professors, especially my mentors in the ASU MCTP program who guided me in my undergraduate research, was invaluable in my journey to completing this thesis. This research was also supported by the National Science Foundation award AGS-1854559.

# DEDICATION

*To: my family*

*All of your endless encouragement, from near and far, will be forever appreciated. Thank you for never doubting me.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

# INTRODUCTION

## 1.1 MOTIVATION

Extreme precipitation and flooding are not unique to any one location in the world. The driving mechanisms behind these phenomena vary regionally but can occur in climates as dry as Colorado and as moist as the northwestern Pacific. Prediction of Rainfall Extremes Campaign in the Pacific (PRECIP) is a field campaign set to take place in Taiwan and Japan in the summer of 2022. PRECIP aims to investigate the mechanisms behind extreme rainfall events in a region with consistently high moisture, leading to highly favorable conditions for heavy precipitation and flooding. The impacts of extreme precipitation and flooding can be devastating no matter the location. In Northern Colorado, the Big Thompson River Flood of 1976 (notably on 31 July, the same date as our case study in Chapter 3) claimed the lives of 143 people when greater than 300 mm of rainfall accumulated in less than four hours due to a nearly stationary thunderstorm (**?**). On 13 May 2019, nearly 400 mm of rainfall fell on the small island of Yonaguni, Japan (a site selected for the PRECIP field campaign) within 24 hours, leading to evacuations and the declaration of the case as a 50-year flooding event by the Japan Meteorological Agency (JMA). There were no known fatalities in the Yonaguni event, but the island was inundated and it is likely that structural damage took place. Increasing the accuracy with which these events are forecast can increase public safety and reduce fatalities. This work is motivated by and completed in conjunction with the PRECIP field campaign.

## 1.2 HEAVY RAINFALL FORECASTING

One approach to heavy rainfall forecasting is through the use of **?**'s ingredients-based methodology for flash flood forecasting. For extreme precipitation to occur, there must be heavy rainfall over the same location for an extended period, or a high rainfall rate, $R$, and a high duration, $D$, in the following equation, where $P$ is total precipitation:

$$P = \bar{R}D. \tag{1.1}$$

Equation (1.1) provides a spectrum of possible rainfall events given combinations of $R$ and $D$, where the most extreme events take place when the largest rain rates occur for the longest durations (**?**). Despite the simplicity of Equation (1.1), multiple factors and processes can contribute to both the intensity and duration of rainfall, making forecasts of heavy rainfall challenging. Precipitation rate and duration are both easily discerned where observations are available, but prove difficult to forecast, particularly in the correct location. Duration can be expanded to $D = L_s/C_s$, where $C_s$ is the system motion vector, representing the propagation speed of the system, and $L_s$ is the length of the system in the direction of $C_s$. Rainfall rate can be further defined as

$$R = Ewq, \tag{1.2}$$

where $E$ is the precipitation efficiency ($E = m_p/m_i$, the ratio of the mass of water falling as precipitation to the influx of water vapor mass in the cloud), $w$ is the vertical velocity, and $q$ is the water vapor mixing ratio (**?**). We use **?**'s methodology as the basis for our approach in that we will evaluate the intensity and location of ascent and moisture in ensemble members.

Low-level shear (the variation of the wind vector within the 850 to 700 hPa vertical column) is a primary contributor to the organization of convection and can interact with cold pools to produce deeper and less inhibited lifting (**??**). Environmental convective available potential energy (the "maximum buoyancy of an undiluted air parcel" (AMS Glossary); CAPE) is generally required for strong vertical motions, but can be reduced by detrimental effects of dry air aloft, which can often weaken or entirely prevent deep convection from forming (**?**). High column water vapor (the amount of water vapor in the vertical column; CWV) is critical to reduce entrainment and support the transition from shallow to deep convection (**?**), and statistically provides a critical threshold for more frequent deep convection (**?**). **?** showed that entrainment should be included in buoyancy estimates in order for CAPE to be meaningful and related to vigorous deep convection. Strong low-level shear, high CAPE, and high CWV are all ingredients that can combine together to produce extreme rainfall events, but the specific combinations and quantitative values can vary across different rainfall events.

### 1.3 Model Accuracy

Quantitative precipitation forecasts (QPF) have improved over time but substantial room for improvement in their accuracy remains. Forecasts for the warm season, when many extreme rainfall events take place, have lower skill and more error because of the scales and types of these events, which is compounded by extreme events having lower skill particularly at larger lead times (**?**). Recent studies have examined the spatial resolution, domain size, and lead time of models and have shown that these factors can be equally as important in model setup as the environmental variables represented in the initial conditions (**???**). Wang found that a cloud-resolving model performed better on days with higher rainfall totals, which has positive implications for forecasting and modeling going forward (**?**). Studies have also been conducted on the limits of predictability, both intrinsic and practical, suggesting that there is likely a continuum between the two that dictates how well a model can perform (**?**).

Models that are convection-permitting with high enough resolution to resolve the scales needed to explicitly simulate heavy rainfall are often quite computationally expensive to run. Bryan et al. (2003) suggested that the general "rule of thumb" of 4-km resolution for resolving convection is marginal because deep moist convection can be smaller, sometimes on the scale of 1–2 km. They suggest resolution on the order of 100 m for the research community, largely so that models can resolve entrainment, which is not practical for the operational forecasting community due to the high computational power needed to achieve this resolution. However, small errors in the initial conditions of a model can lead to entirely different model outcomes because errors can propagate both up- and downscale (**???**). Ensemble approaches can take the uncertainties of initial conditions, and sometimes model physics as well, into account to produce a spectrum of possible solutions. Running multiple ensemble members can greatly increase computational cost, so in practice a compromise between lower and higher resolution is needed to provide better QPF guidance (**?**).

The need for more frequent usage of ensemble forecasts has become more obvious in recent years, as ensembles outperform deterministic forecasts (**??**). Ensembles provide a broad set of the potential realities that allow forecasters to see the different ways in which a storm or event might evolve with a variety of different slightly perturbed initial conditions, which can mimic the small inaccuracies that can arise from a lack of observations. Ensembles are becoming more widely used in numerical weather prediction. While ensembles are more computationally expensive to run than deterministic models, their increased usage can lead to better parameterizations and parallelizations, ultimately increasing the benefit of the computational expense.

### 1.4 Research Objectives

The main motivation of this thesis is to further understand the meteorological processes behind two extreme rainfall events, which took place in two different regions, and highlight the importance of ensemble forecasts in numerical weather prediction. In order to accomplish these goals, we evaluated ensemble model forecasts and various forms of observational data over the northwestern Pacific and

Northern Colorado for two separate case studies. Both case studies employ an ensemble-based analysis and an ingredients-based approach to meteorological analysis. We first examine a mid-season mei-yu heavy rainfall event that took place over Yonaguni, Japan and the surrounding ocean on 9 June 2020. We use output from a WRF ensemble and observational data to examine the ensemble spread in results as well as the meteorological variables in order to highlight the value of ensemble forecasts and to shed light on the meteorological drivers behind this case. For a contrasting case, we examine an event that occurred in Northern Colorado on 31 July 2021 in which heavy rainfall was forecast and flash flood warnings were issued, but little to no rain and flooding took place. Similar to the first case study, we use WRF ensemble output and observational data, including in-situ soundings, to investigate the meteorological processes at play. Both events were poorly forecast by multiple numerical weather models, so we aim to understand why to help forecasters and modelers alike in predicting and modeling events such as these in the future.

Our research objectives are as follows:

(1) Evaluate ensemble accuracy in forecasting two heavy rainfall events

(2) Explore the meteorological drivers behind each case

Chapter 2 will present the case study of 9 June 2020. Chapter 3 will present the case study of 31 July 2021. Chapter 4 will provide conclusions and suggestions for future work.

CHAPTER 2

## An ensemble-based analysis of a liminal extreme rainfall event near Taiwan

### 2.1 Introduction

[1]

#### *2.1.1 Motivation*

A number of weather phenomena contribute to heavy (and oftentimes extreme) rainfall during early summertime in tropical east Asia, including the southwest monsoon, the mei-yu, and typhoons. The area is extremely moisture-rich during this time due to the southwest monsoon providing a consistent supply of warm, moist air. The mei-yu (known by this name in China and Taiwan, but as "baiu" in Japan and "changma" in Korea) is the dominant rainfall regime over Taiwan from mid-May to mid-June, characterized by periods of heavy rainfall with afternoon rainfall maxima (**?**). The mei-yu is a nearly continuous, quasi-stationary rainband, sometimes referred to as a summer monsoon trough, that progresses northward through time, marked by low-level baroclinicity around China and Japan early in the season, and evolves into a trough of sea-level pressure and sharp moisture gradients near the surface later in the season (**???**). The mei-yu is primarily a moisture front in that it is characterized by a strong horizontal moisture gradient and a modest horizontal temperature gradient (**???**).

Mesoscale fronts eject southward off of the rainband toward Taiwan and the southernmost Japanese islands. These fronts are estimated to exist for anywhere from 4 to 10 days, depending on the way they are defined, and are often embedded with mesoscale convective systems (MCS) (**??**). These mesoscale mei-yu fronts and embedded MCSs positively interact and intensify each other (**?**). Low-level jets (LLJ) that form on the south or southwest side of the mei-yu also play an important role in the development and intensification of these heavy rainfall events by transporting moist tropical air into the area (**??**).

Extreme rainfall events occur more often in the mei-yu than in Taiwan's other rainy season regimes, largely due to mei-yu fronts, and particularly in the second half of the regime because of the warmer southwesterly flow. These extreme rainfall events contribute to the regime having the highest rainfall frequencies and rain rates among all of the rainy season regimes in Taiwan (**????**). The extreme rainfall that the mei-yu and mei-yu fronts can produce still provides large forecasting challenges.

[1]This chapter has been published in the Atmosphere special issue 'Precipitation in Taiwan and Neighboring East Asia Areas: Observation, Analysis, and Forecast (**?**).

The mei-yu season in 2020 is recognized as an anomalously rainy season, the reasons for which have been investigated in recent literature. **?** examined the contribution of atmospheric rivers to this extreme mei-yu season and found that atmospheric-river related precipitation contributes 50–80% to total mei-yu precipitation. **?** discovered that the halted northward progression of the Eastern Asian Summer Monsoon Front was associated with anomalously high precipitation in the Yangtzhe River Basin. The anomalous 2020 season was also found to be impacted by persistent Madden–Julian oscillation phases 1–2 (**?**). Mei-yu seasons such as that in 2020 may become more frequent in the future (**?**), and Liang suggests that there are quantitative challenges in predicting extreme mei-yu seasons such as 2020 (**?**), further highlighting the importance of case studies such as this one.

One such example of an extreme rain-producing event associated with a mei-yu front took place on 9 June 2020, when the open ocean region east of Taiwan and near Yonaguni, Japan received rainfall totals greater than 100 mm spanning a large area with peak amounts surpassing 600 mm. This event was poorly forecast by global models, with most producing too much orographic rainfall or vastly underestimating the rainfall that accumulated east of Taiwan. At the time of this case, Weather Research and Forecasting (WRF) ensemble simulations were being conducted in association with the Prediction of Rainfall Extremes Campaign in the Pacific (PRECIP) 2020 dry run. The event had model output on both synoptic-scale (9-km grid spacing resolution) and mesoscale (3-km grid spacing resolution) domains. The overall synoptic conditions across the ensemble members were fairly similar, though the mesoscale details and rainfall totals varied substantially even in 24-h forecasts. This event illustrates that while the synoptic scale can be reasonably well forecast, mesoscale features still contribute heavily to the challenge of heavy rainfall forecasting.

## 2.2 Model setup and validation

### *2.2.1 Model description*

This study analyzes the model output from select members of a 40-member ensemble that used the Weather Research and Forecasting (WRF) model version 4.1.3 (**?**). The ensemble member selection process is described below. The 40-member ensemble was created using the Pennsylvania State University WRF ensemble Kalman filter (PSU WRF-EnKF) modeling system, which uses data assimilation and perturbations in the model initial conditions to generate the ensemble (**???**). The PSU WRF-EnKF system uses two two-way nested domains with spatial resolutions of 9 km and 3 km. We will focus our analysis on the 3-km domain, which contains $300 \times 300$ horizontal grid points and 50 vertical levels.

The data assimilation runs for 12 h preceding the beginning of the forecast period, at which point a 48-h forecast with hourly output is generated.

### 2.2.2 *Observational validation data*

In order to quantitatively and spatially examine ensemble members' rainfall, we compared model output with rainfall estimates provided by the Quantitative Precipitation Estimation and Segregation Using Multiple Sensors (QPESUMS) dataset, which is a gridded operational product with hourly output and 0.0125° horizontal resolution that uses radar reflectivity and polarimetric data, where available from S-band single- and dual-polarimetric and C-band dual-polarimetric radars, to estimate rainfall and high-density rain gauge observations to correct the estimations (**?**). QPESUMS was developed by the Central Weather Bureau, Water Resource Agency, and Soil and Water Conservation Bureau of Taiwan and the National Severe Storms Laboratory. We linearly interpolated the QPESUMS data (effectively coarsening the grid) onto the native WRF grid at every time step in order to provide a means by which point calculations, as in the following subsection, can be made. We chose to coarsen the QPESUMS data, rather than to refine the WRF data, so as to not create false rainfall overestimations at any grid point in the WRF output.

TABLE 2.1. The contingency table used to define variables used in calculating equitable threat scores.

| | **forecast "yes"** | **forecast "no"** |
|---|---|---|
| **observed "yes"** | hit | miss |
| **observed "no"** | false alarm | correct negative |

### 2.2.3 *Ensemble member classification*

For the purposes of this study, we selected four ensemble members for analysis: a "best" member, a "better" (or second best) member, a "worst" member, and a "worse" (or second worst) member, representing the two most and two least accurate members. In order to select these members, we used a combination of equitable threat scores (ETS, commonly called Gilbert skill scores) and relative operating characteristic (ROC) curves. The ETS metric is similar to a threat score, a commonly used statistic in deterministic forecast evaluation which evaluates how well a forecast of an event happening corresponds to an observation of the same event happening, but the ETS also accounts for hits due to random chance. Equitable threat scores can be calculated via the following equation:

$$ETS = \frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{misses} + \text{false alarms} - \text{hits}_{\text{random}}}, \tag{2.1}$$

where hits, misses, and false alarms are defined in the contingency table shown in Table 2.1 (correct negatives will not be used in this study). Hits due to random chance are calculated via the following equation:

$$\text{hits}_{\text{random}} = \frac{(\text{hits} + \text{misses}) * (\text{hits} + \text{false alarms})}{\text{total}}. \tag{2.2}$$

Equitable threat scores lie in the range from -1/3 to 1, where the range from 0 to 1 is “skillful” and the range from -1/3 to 0 is “unskillful”. Relative operating characteristic curves plot the probability of detection (POD) and the false alarm rate (FAR), equations shown below, against each other.

$$POD = \frac{\text{hits}}{\text{hits} + \text{misses}} \tag{2.3}$$

$$FAR = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \tag{2.4}$$

On a ROC curve plot, a curve that reaches the top left corner (point (0,1)) has a perfect FAR of 0 and POD of 1 and is considered a perfect forecast, whereas a curve that reaches the bottom right corner (point (1,0)) has the worst possible FAR of 1 and POD of 0 and is considered a perfectly wrong forecast.

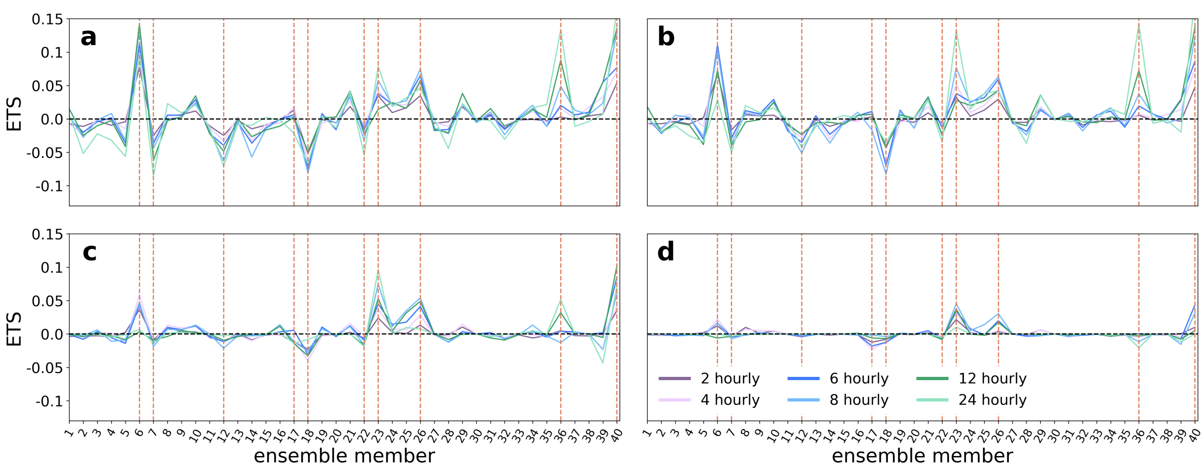


FIG. 2.1. (**a**) Equitable threat scores for a rainfall rate threshold of 1 mm/hr. Hourly periods are differentiated by different colors, which are defined in the legend in panel (d). The three best performing three worst performing ensemble members are highlighted by vertical dashed orange lines. **(b)** As in (a) for a rainfall rate threshold of 2 mm/hr. **(c)** As in (a) for a rainfall rate threshold of 5 mm/hr. **(d)** As in (a) for a rainfall rate threshold of 10 mm/hr.

As a first pass, we calculated ETS for all 40 members at a 24-h lead time within the region of interest (outlined by the red box in Fig. 2.4) at different hourly thresholds (1, 2, 5, and 10 mm of rainfall per hour) and chunks of varying duration within the 24-h evaluation period. We separated the 24-h period into chunks with different durations (of 2, 4, 6, 8, 12, and 24 hrs) so that we have a different number of chunks for each duration; e.g., we have 12 chunks with 2-hr duration and 3 chunks with 8-hr duration. The thresholds used in ETS calculation reflect the same average rain rate for all chunks, regardless of the duration (e.g., for a 2-hr chunk using a 1 mm/hr average rain rate, the threshold would be 2 mm, and for a 4-hr chunk using a 1 mm/hr average rain rate, the threshold would be 4 mm). We calculated ETS across each chunk, and, to fairly compare these chunks, we removed the time component by summing across the numerator and denominator of Equation 2.2 to achieve an ETS score for each duration. We calculate six ETS for each member (one for each duration), shown in Fig 2.3.

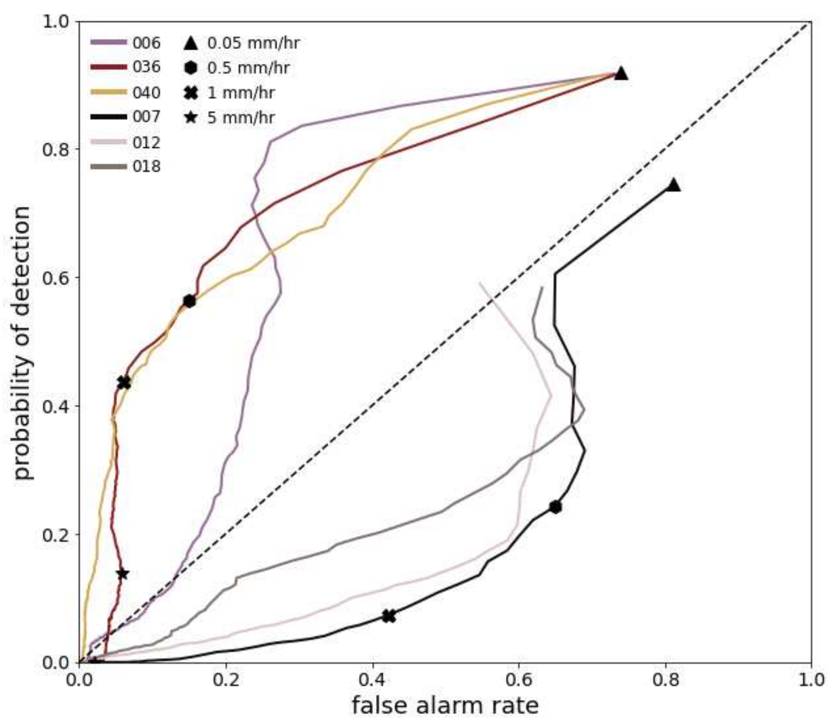


FIG. 2.2. Relative operating characteristic curves for the three best and three worst performing ensemble members. Different rainfall rate thresholds are marked on the single best and worst members only.

From Fig. 2.1, we select the members whose scores are consistently highest and lowest across the selected thresholds, members 006, 007, 012, 018, 036, and 040 highlighted by dashed orange lines. As a second pass, we plot the ROC curves for all six members (across a threshold ranging from 0 to 10 mm/hr of rainfall for smoothness), shown in Fig. 2.2. From Fig. 2.2, we determine that member 036 reaches closest to the top left corner and is "best", followed by member 040 as the "better" member, and member 007 reaches closest to the bottom right corner and is "worst", followed by member 012 as the "worse" member. Note that member 006 appears to do similarly well, if not better, in terms of the POD metric, however the increased FAR metric across the majority of thresholds led us to exclude it from the "best" or "better" classification. In some statistical and machine learning applications, ROC curves are evaluated by calculating the area under the curve (AUC). Here, the ROC curves do not extend to the top right corner at the lowest threshold point (0.0 mm/hr) and do not align in the upper right of the plot (at high FAR and POD values), so the AUC comparison is not viable. Going forward, we will primarily focus on the "best" and "worst" members, with some repeated comparison of "better" and "worse" to increase confidence in the validity of our conclusions. We will also examine composite plots showing the difference between the means of the five most and the five least accurate members, determined using both ETS scores and ROC curves, for a number of variables.

## 2.3 Analysis

### *2.3.1 Case Summary*

On 9 June 2020, the area east of Taiwan near Yonaguni, Japan received extreme amounts of rainfall associated with what we identify to be a mei-yu frontal event. Figure 2.3a shows the observed rainfall estimated by the QPESUMS dataset over the 24-h period starting at 00 UTC on 9 June 2020 (**?**). QPESUMS is described in Section 2.2.2. In this 24-h period, the area to the southeast of Yonaguni island, near 123 E, 24 N received peak rainfall amounts greater than 600 mm. Yonaguni received greater than 150 mm of rainfall, and a region off the east coast of Taiwan that runs diagonally from near 121.8 E, 23 N to 123.2 E, 25.5 N received rainfall amounts up to 300 mm. Hourly rainfall plots show that this event was characterized by a quasi-stationary, back-building rainband with two distinct periods of increased rainfall rates, the first starting around 00 UTC and the second around 13UTC (not shown here). The data showed peak rainfall rates of greater than 130 mm per h in the QPESUMS data. While other areas of the domain received total rainfall amounts in the 0–100 mm range, the primary focus of this study will be on the rainfall in and around the peak located southeast of Yonaguni so as to keep our focus on the maximum rainfall and processes occurring over the open ocean, as opposed to orographic processes.

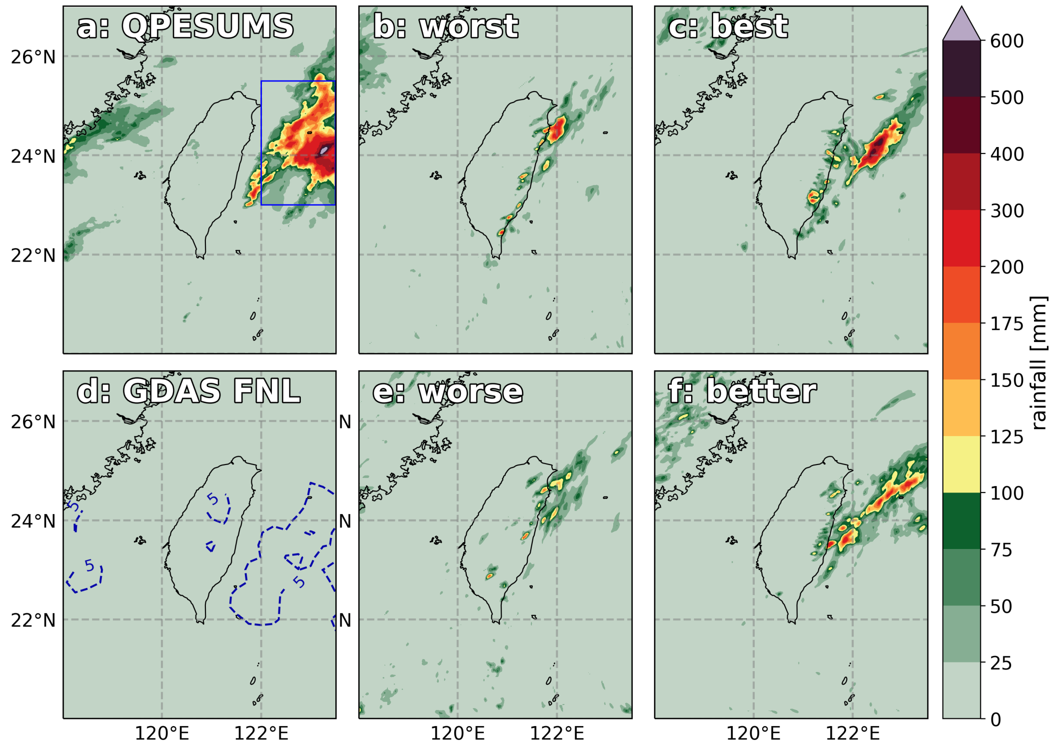


FIG. 2.3. (**a**) Map of rainfall totals for QPESUMS for the 24-h period beginning at 00 UTC 9 June 2020. (**b**) As in (**a**) for the worst WRF ensemble member. (**c**) As in (**a**) for the best WRF ensemble member. (**d**) As in (**a**) for GDAS FNL. The blue dashed contour lines outline areas that received greater than 5 mm of rainfall (**e**) As in (**a**) for the second worst (worse) ensemble member. (**f**) As in (**a**) for the second best (better) ensemble member.

Figure 2.4 shows the brightness temperature from NASA GES DISC satellite imagery. The mei-yu front is along the periphery of the 850hPa high shown in Figure 2.5 (described in further detail in Section 2.2.3). The rainband associated with the front extends behind the high southwest towards Taiwan, from approximately 140 E, 30 N to 105 E, 25 N over southern China, which is reflected in Figure 2.4, and the wind field is shown in Figure 2.5.

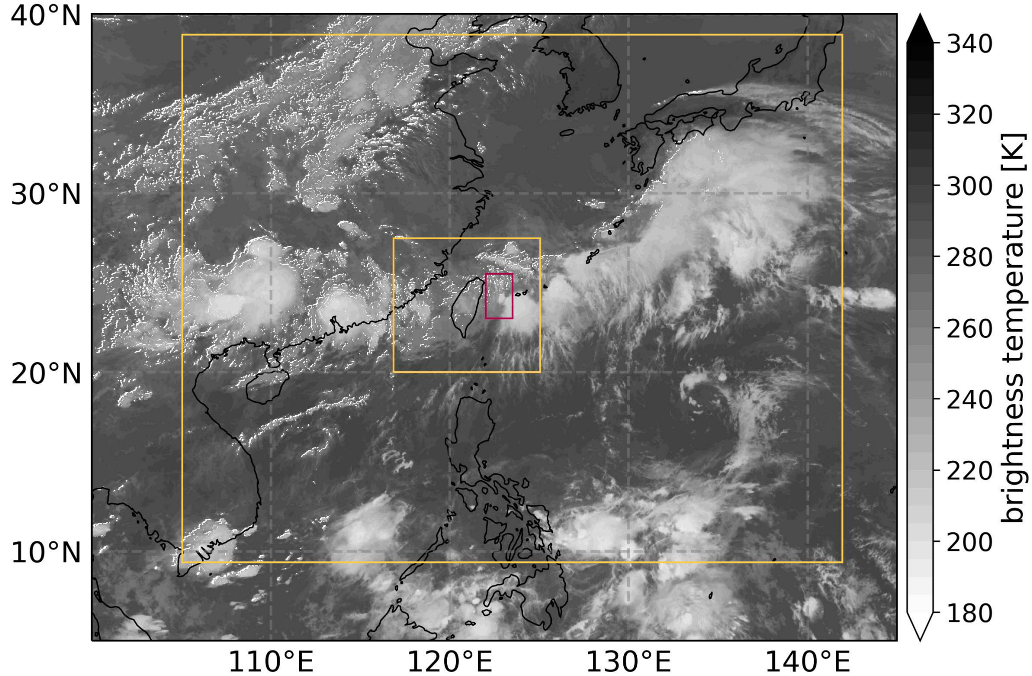


FIG. 2.4. Map of NASA GES DISC brightness temperature at 00 UTC 9 June 2020. The yellow boxes represent the 9-km (**outer**) and 3-km (**inner**) WRF domains used in the PSU WRF-EnKF simulations. The red box shows the region for ensemble statistical calculations and soundings, also called the "region of interest".

Central Weather Bureau (CWB) forecast discussions for the night of 8 June 2020 to the morning of 9 June 2020 mentioned heavy rainfall on the NE side of Taiwan associated with a stationary front extending from 28 N, 137 E to 26 N, 110 E, but CWB did not issue any warnings for this event. Global models, including National Centers for Environmental Prediction (NCEP) and European Centre for Medium-Range Weather Forecasts (ECMWF, not shown here), poorly forecast this event at 24–36 h lead times. These models typically produced excessive orographic rainfall or vastly underestimated the rainfall that accumulated east of Taiwan, or both. We aim to investigate what led to these model forecast inadequacies by examining the selected members of the WRF ensemble described in the Ensemble Member Selection section.

Panels b, c, e, and f in Figure 2.3 show the 24-h rainfall totals produced in the WRF ensemble run initialized at 12 UTC 8 June 2020 for the worst member (b), the best member (c), the worse member (e), and the better member (f), as determined by forecast verification metrics. Panel a shows the QPESUMS observed rainfall totals. Panel d shows the rainfall totals from the GDAS FNL 24-h forecast, with dashed blue contours outlining areas that received greater than 5 mm of rainfall as the accumulated rainfall does not surpass the first threshold on the colormap. The GDAS FNL 24-h forecast produced little to no rainfall. All members fail to accurately replicate rainfall in the location of the >600 mm peak shown in the observations, both in magnitude and location. The best member most accurately captures the intensity and orientation of the rainfall from the primary rainband, and the better member similarly captures the orientation but does not produce as much rainfall as the best member. The worst and worse members both fail to capture the intensity and location of the primary band of rainfall, where the rainfall is too weak, too far north, and too close to Taiwan's coast. The worst and worse members do, however, capture the orientation of the storm's movement (shown in the diagonal streaking/direction of the rainfall totals).

It is of note that the ensemble members appear to produce more rainfall over the Central Mountain Range (CMR) in Taiwan than we see in the observations. While orographic processes likely impacted the rainfall here, including them is outside of the scope of this study. We will instead focus on the rainfall to the east of Northern Taiwan, over and around Yonaguni, to better understand the thermodynamic and dynamic processes taking place over the open ocean through analysis of the WRF ensemble members. Moving forward, we will focus the majority of our analysis on the best and worst members. The best and better members' similarity as well as that of the worst and worse members' is relevant, though, as the results of the best and worst members prove to be quite representative of their similar counterparts. Section 2.3.3 will elaborate on further analysis using the better and worse members and the five most and the five least accurate members and their composite differences.

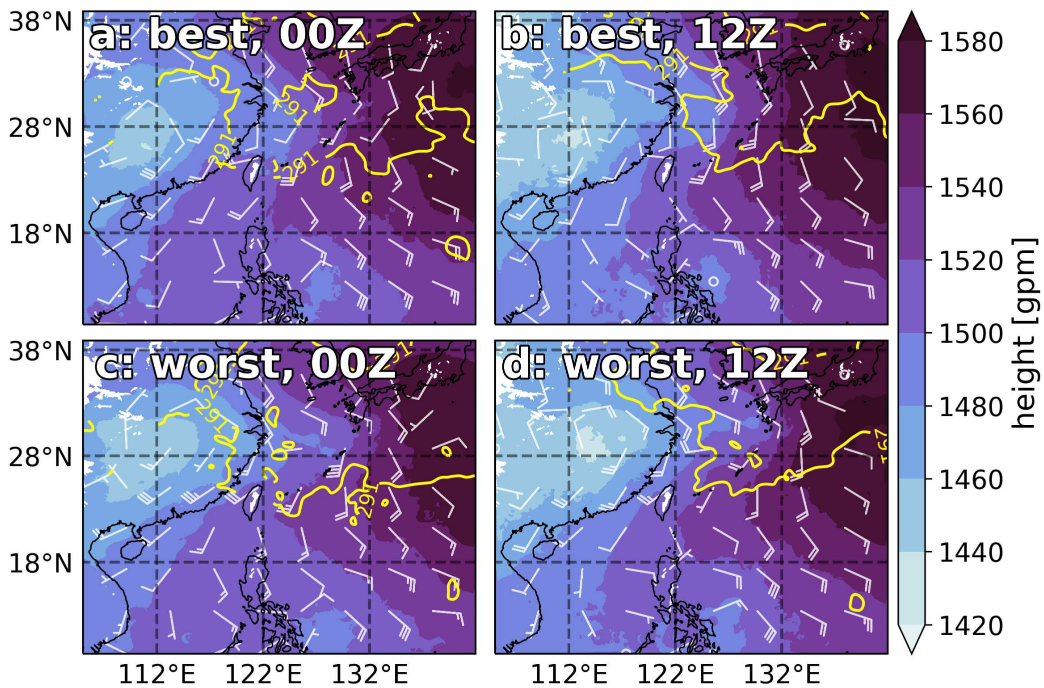


FIG. 2.5. **(a)** Map of 850 hPa heights (shading), zonal/meridional wind vectors in units of knots (white), and the 291 K temperature contour (yellow) on the 9 km domain for the best member at 00UTC 09 June 2020. The temperature has been filtered with a 2-dimensional Gaussian filter with a sigma of 3, using the SciPy package (**?**). **(b)** As in (a) at 12UTC. **(c)** As in (a) for the worst member at 00UTC. **(d)** As in (a) for the worst member at 12UTC.

### *2.3.2 Synoptic-Scale Analysis*

Fig. 2.4 shows the synoptic-scale cloud patterns from satellite brightness temperature at 00UTC on 09 June 2020. Fig. 2.5 shows the 850 hPa geopotential heights, winds, and 291K temperature contour for the best (top) and worst (bottom) members at 00UTC (left) and 12UTC (right) for the 9 km domain in the WRF ensemble. The synoptic features visible in (Fig 2.5) are representative of a classic, albeit weak, mei-yu setup, with east-west gradients in geopotential height driving southerly monsoonal flow. The east-west gradients are apparent in the contrast between the 850 hPa subtropical high pressure system in the NE corner of the domain and the 850 hPa low pressure system over China. As the simulation progresses from 00UTC to 12UTC, the 850 hPa low over China strengthens slightly. The differences between the best member (top) and worst member (bottom) in Fig. 2.5 are small, but these small differences in the synoptic setup are enough to motivate differences in the mesoscale environment that significantly change the rainfall produced in each member. In the best member, the 850 hPa low over China is smaller and does not extend as far east as in the worst member; winds near the area of interest, just east of Taiwan, are stronger than in the worst member; and the 291K temperature contour extends further southwest near the Japanese islands just east of Taiwan, whereas the 291 K contour is further north in the worst member. The key differences in the wind-fields between the two members are in the meridional direction, where the worst member has slightly more westerly winds than the best member.

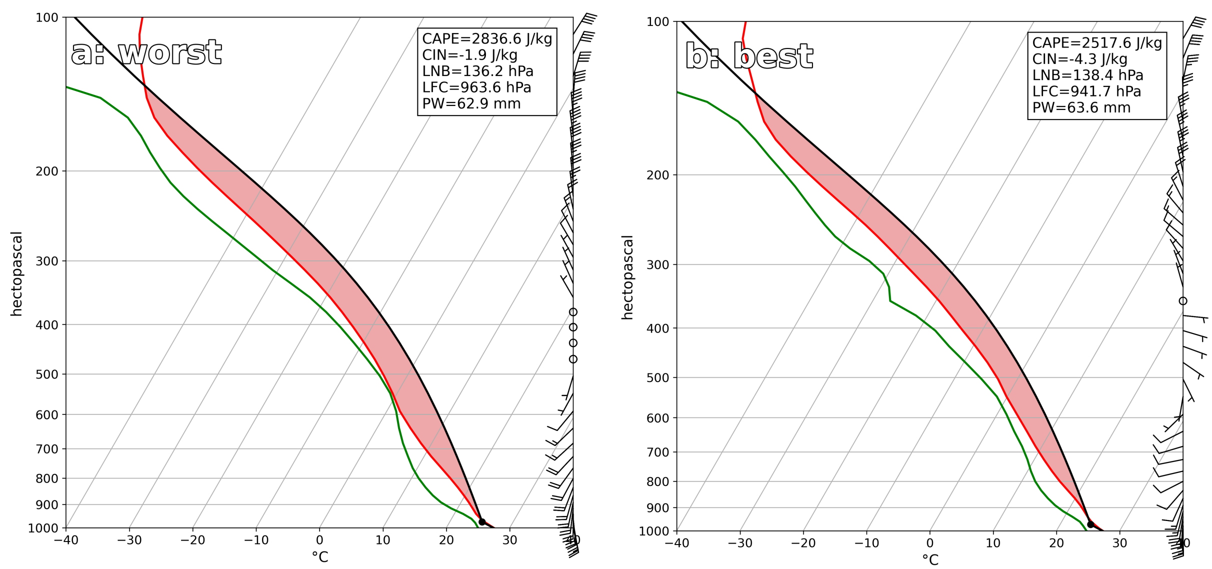


FIG. 2.6. **(a)** Box-average sounding calculated over the red box in Fig. 2.4 for the worst member at 12UTC. The lifted condensation level (LCL) is marked by a black dot. Convective available potential energy (CAPE), convective inhibition (CIN), the level of neutral buoyancy (LNB), the level of free convection (LFC), and precipitable water (PW) are listed on the sounding. Wind barbs along the right side of the plot show zonal/meridional winds in units of knots. **(b)** As in (a) for the best member.

We calculated a representative sounding using a box average over the area of interest (outlined by the red box in Fig. 2.4) at 12UTC (2.6). We chose 12UTC as the time of interest for the soundings (and for some future plots which focus on one time only) because it is when the storms have fully developed in the model and thus, the environments are more individually representative and comparable. The soundings are overall remarkably similar. Both members have veering winds from the surface up to about 700 hPa, indicating that there is warm air advection taking place in the lower atmosphere in both members. The best member's flow is more westerly from the surface up to 700 hPa, while the worst member has a slightly stronger (around 5 kt) and more predominantly southerly flow, which will be important in the analysis going forward as we focus on sources of moisture within the domains.

Both members show moist conditions, with remarkably similar levels of neutral buoyancy (LNB) and free convection (LFC) and amounts of precipitable water (PW), suggesting that the overall thermodynamic profiles throughout the atmosphere are not significantly different. The PW values are both high and show that there is ample moisture in both environments for heavy rainfall to take place. Both members have low levels of free convection (LFC), which indicate thermodynamic environments which support uplift in regions where boundaries exist. We will examine the frontal interactions in upcoming analysis.

The differences between the best and worst members are most notable in the convective available potential energy (CAPE), convective inhibition (CIN), and moisture in the lower atmosphere. The best member has around 2 J/kg more CIN and around 300 J/kg less CAPE than the worst member. The lower CAPE is associated with a profile closer to moist adiabatic in the best member, with greater moisture in the lower atmosphere from the surface up to around 700hPa. The worst member's sounding features a drier profile than the best member in the lower atmosphere below 600hPa. The difference in low-level moisture also shows up in relative humidity (RH) and column-integrated water vapor (CWV) analysis to be described later (Fig. 2.8-2.10).

Soundings calculated for entraining CAPE (not shown here) indicate that entrainment reduces but does not eliminate CAPE, though it reduces CAPE more in the worst member than the best. Even at entrainment rates of up to 2% per 100 m, the soundings still show some instability. Thus, entrainment is not likely to be the primary factor resulting in differences in the two members' rainfall production.

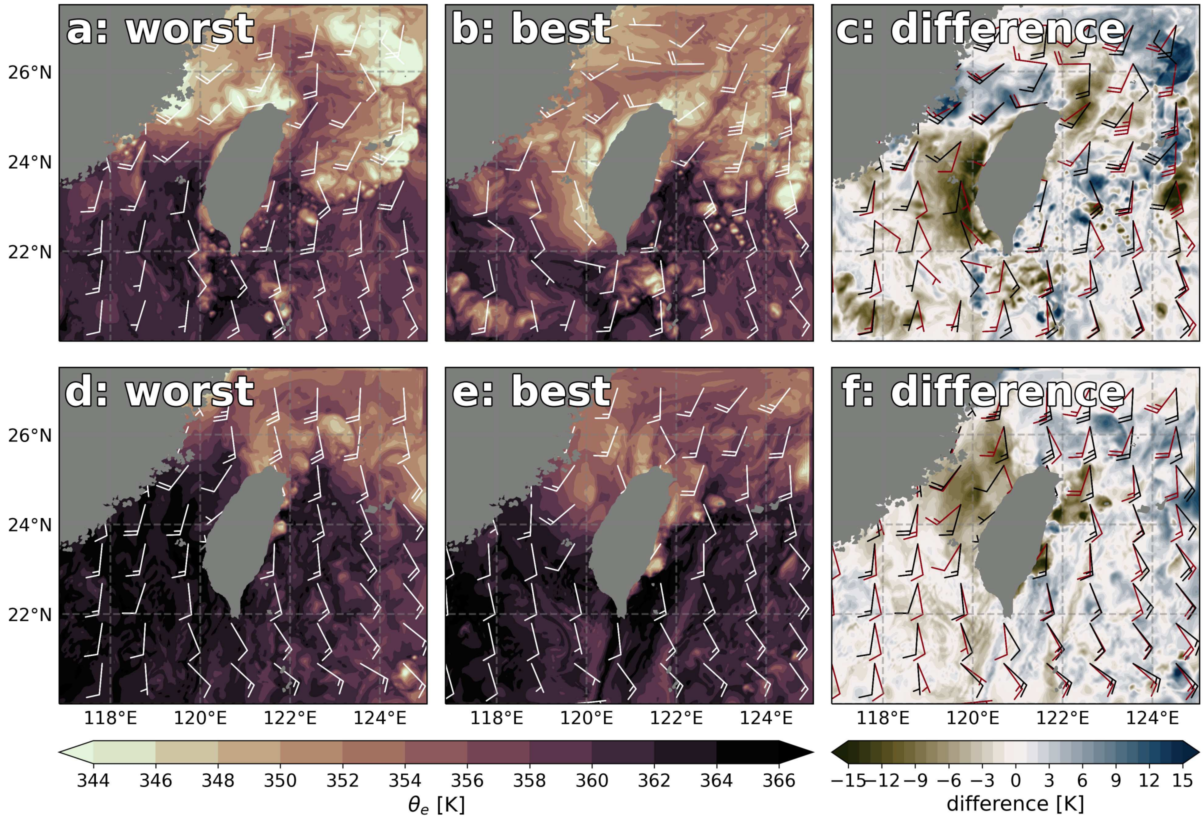


FIG. 2.7. **(a)** Map of 100 m $\theta_e$ (shading) and zonal/meridional winds in units of knots (white) on the 3 km domain for the worst member at 00UTC. **(b)** As in (a) for the best member. **(c)** Map of the difference between panel (a) and (b) (worst subtracted from best) and zonal/meridional winds (worst in black barbs, best in red barbs). **(d)** As in (a) for the worst member at 12UTC. **(e)** As in (a) for the best member at 12UTC. **(f)** As in (c) at 12UTC.

### *2.3.3 Mesoscale analysis*

Equivalent potential temperature ($\theta_e$) acts as a valuable measure of temperature and moisture, as well as unstable, buoyant air. $\theta_e$ is calculated using Equation 2.5 according to **?**, using the metpy package (**?**).

$$\theta_e = \theta_{DL} \exp\left[\left(\frac{3036.}{T_L} - 1.78\right) * r(1 + .448r)\right], \tag{2.5}$$

where $\theta_{DL}$ is the potential temperature at the lifted condensation level (LCL), $T_L$ is the temperature at the LCL, and $r$ is the mixing ratio. Plots of 100-m temperature and moisture (not shown here) suggest that the moisture contribution to $\theta_e$ variability dominates so we will focus on $\theta_e$ as a measure for moisture going forward. We choose 100 m to show plots of $\theta_e$ (and later $\theta_v$ and frontogenesis) as it is a representative layer of air that is typically lifted in buoyant plumes of convection (**?**). Gradients in $\theta_e$ at

100 m above the surface can thus be used to identify near-surface boundaries in moisture and buoyancy but are not as valuable in identification of near-surface fronts (we will employ virtual potential temperature ($\theta_v$) for that purpose later on) (**?**).

There is ample moisture in the area of interest for strong uplift mechanisms to trigger deep, moist convection, particularly when interacting with a low LFC as shown in the soundings in Fig. 2.6. At 00UTC, there is an intrusion of warm, moist air near 122E, 24N in the best member, which creates a strong $\theta_e$ gradient (on the order of 15K) in the area of interest between the intrusion and the colder, drier air mass just northeast of Taiwan (Fig. 2.7). In the worst member at 00UTC, the $\theta_e$ gradient is not as strong and the colder, drier air mass hugs the coast of Taiwan rather than extending eastward, as it does in the best member. The combination of the weaker warm, moist intrusion and the less extended cold, dry air mass in the worst member creates a weaker north-south $\theta_e$ gradient in the area of interest. By 12UTC, the gradient in the worst member has been pushed north/northwest out of the area of interest by the stronger southerly winds and becomes confined along Taiwan's coast, while the gradient in the best member has remained in the area of interest and has retained its strength, though its spatial spread is smaller. These differences are highlighted in panels c and f, which show the differences between the $\theta_e$ fields at 00UTC and 12UTC. The cooler air in panel f shows that the $\theta_e$ gradient in the best member is more prominent than in the worst member at 12UTC, suggesting that the air mass is either significantly cooler, closer to the surface, or both. Notably, the 100 m winds in panel f are quite similar between the best and worst member, but this does not have a large impact on the air mass and storm progression because the system has already set up and is producing heavy rainfall at this point in time in reality.

It is not immediately clear whether the convection taking place is reinforcing the $\theta_e$ gradient or whether the $\theta_e$ gradient is helping to drive the convection—rather, it is likely a combination of the two processes contributing to a positive feedback mechanism. In any case, this difference in the locations of the near-surface gradients helps to explain the presence of stronger rainfall in the area of interest in the best member and the lack thereof in the worst member.

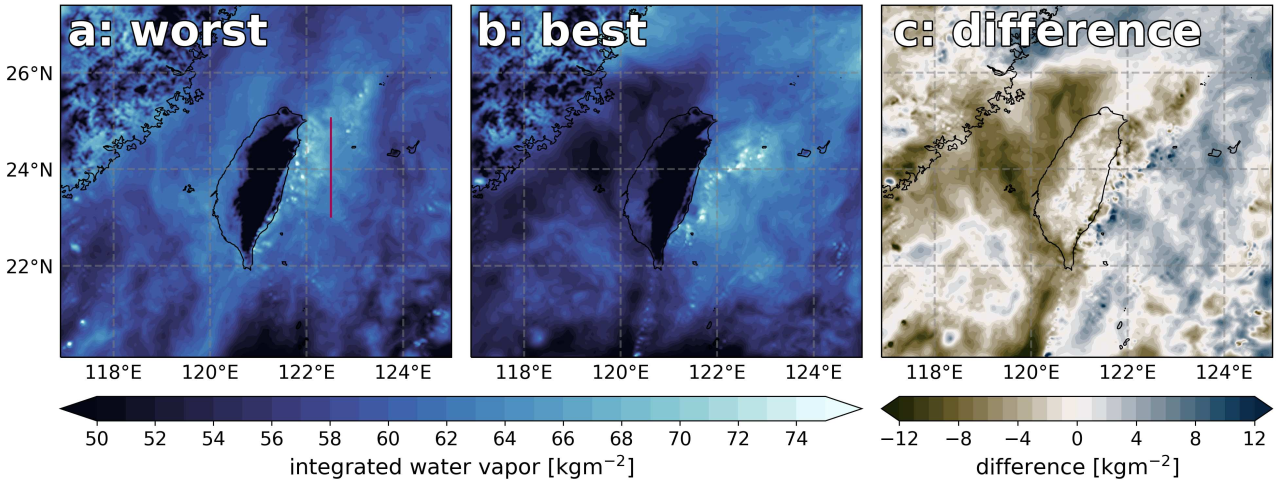


FIG. 2.8. **(a)** Map of the column-integrated water vapor for the worst member at 12UTC on the 3 km domain. The red line shows the line across which vertical cross-sections are taken for Fig. 2.12. **(b)** As in (a) for the best member. **(c)** Map of the difference between (a) and (b) (worst subtracted from best).

Integrated water vapor is calculated by vertically integrating the mixing ratio ($r$) through the depth of a sounding ($p_{bottom}$ to $p_{top}$), using the metpy package, as follows:

$$IWV = -\frac{1}{g}\int_{p_{bottom}}^{p_{top}} r\,dp \tag{2.6}$$

At 12UTC, the worst member and the best member have similar maximum values of integrated water vapor, but these maximums are displaced (Fig. 2.8). The best member's maximums are within the region of interest and peak above 75 kg m$^{-2}$, while the maximum values in the worst member are closer to 70 kg m$^{-2}$ and are confined along Taiwan's coast. Panel c in Fig. 2.8 shows the difference between the best and worst member. The negative values north of and positive values within the region of interest highlight this displacement of moisture well. North of the region of interest, the worst member is more moist than the best member, but within the region of interest the best member is more moist than the worst member. Integrated water vapor values in the high 60s and low 70s kg m$^{-2}$ indicate large quantities of moisture to begin with but differences on the order of 8-10 kg m$^{-2}$ are non-negligible. In both members, the cells move eastward and northeastward off of the coast of Taiwan, so the drier air in the worst member over the region of interest inhibits the formation of deep, moist convection, while the moisture in the best member promotes the formation of deep, moist convection and reflects the presence of clouds that have already developed.

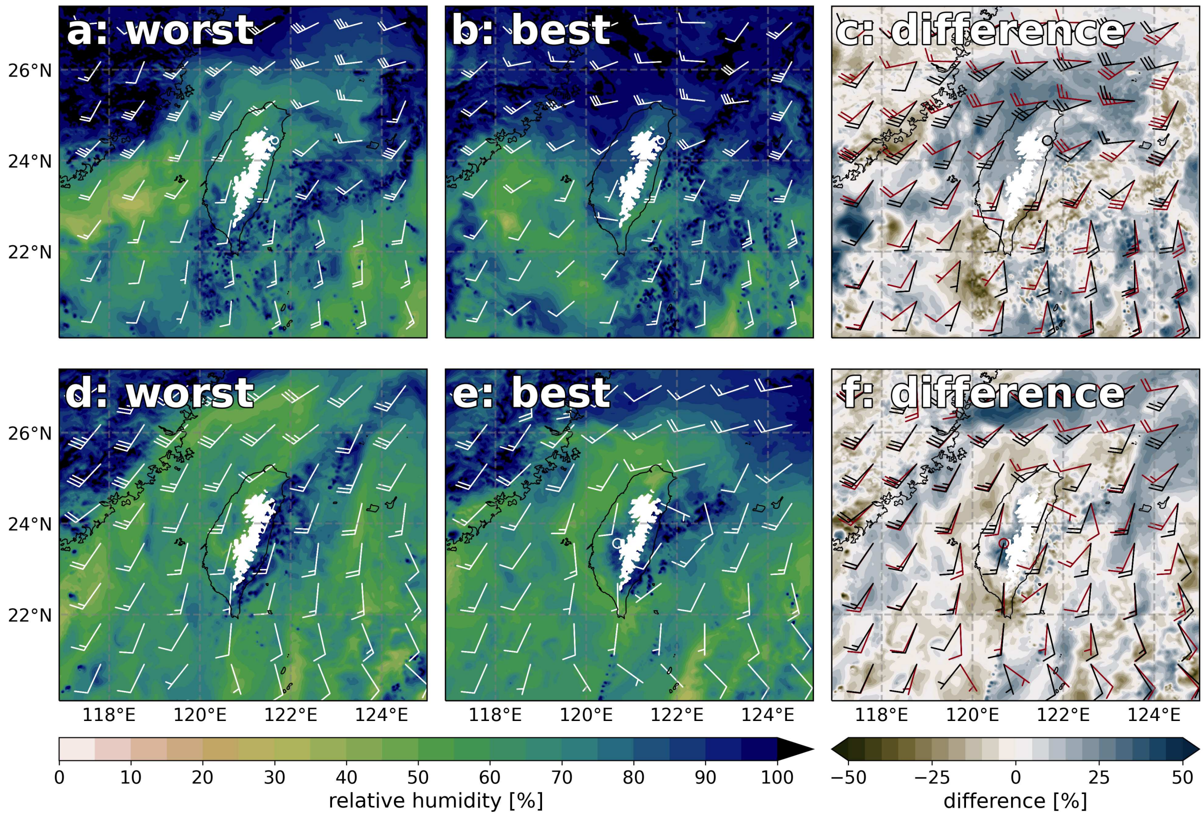


FIG. 2.9. **(a)** Map of the 850 hPa relative humidity (shading) and zonal/meridional winds in units of knots (white) for the worst member at 00UTC on the 3 km domain. **(b)** As in (a) for the best member. **(c)** Map of the difference between panel (a) and (b) (worst subtracted from best) and zonal/meridional winds in units of knots (worst in black barbs, best in red barbs). **(d)** As in (a) at 12UTC. **(e)** As in (b) at 12UTC. **(f)** As in (c) at 12UTC.

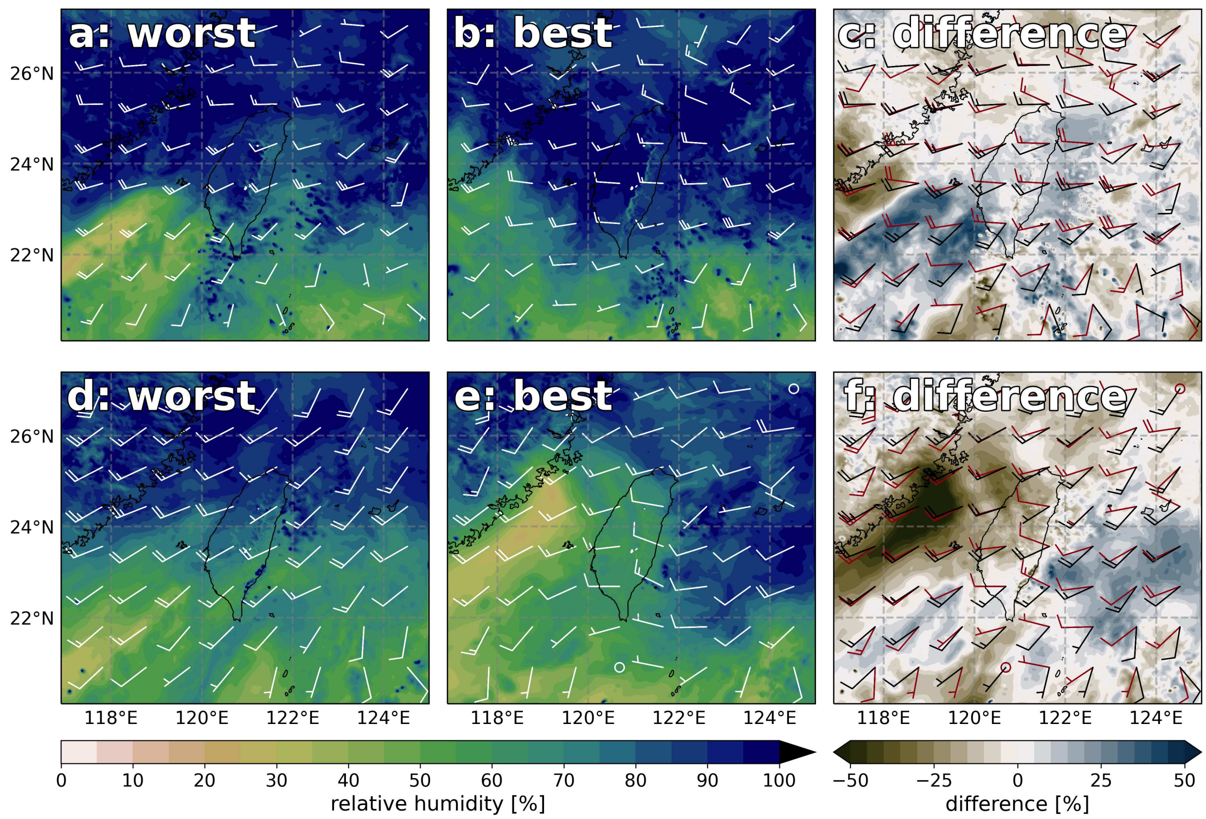


FIG. 2.10. As in Fig. 2.9 at 700 hPa.

At 850 hPa, the winds in the worst member remain organized and southerly from 00UTC (panel a) to 12UTC (panel d), whereas in the best member, the winds weaken and become less organized over that period (Fig. 2.9). The worst member has greater relative humidity values than the best member in the area of interest at 00UTC, but drier air is advected into the region in the worst member and the more humid air is pushed north. Focusing on panel f, we see that the best member has relative humidity values 25% greater than the worst member south of and in the area of interest, and the flow is weak and has a stronger westerly component, while the worst member has a predominantly southwesterly flow. The stronger, southwesterly winds in the worst member are advecting the drier air into the area of interest, while the weaker westerly winds in the best member are not significantly changing the moisture profile in the region.

At 700hPa at 00 UTC, the best member has RH values that are up to 20% greater just below the region of interest and more predominantly southwesterly flow, allowing moister air to be advected into the region (Figure 2.10c). The worst member, however, has slightly stronger (5 kt), more southerly flow and drier air south of the region of interest, which causes the drier air to be more quickly advected into the region of interest. At 12 UTC (2.10f), there is slightly drier (10% less relative humidity) air in the best member in the area of interest closer to Taiwan's west coast but moister air near 122.5 E near Yonaguni, which could be indicative of the location of existing storms at this time.

The moisture differences between pressure levels in the members are also noteworthy. The worst member shows more dryness moving from 850 upwards in the atmosphere. This drier air higher in the atmosphere, especially at 00UTC, could be contributing to the lack of deep, moist convection in the worst member by inhibiting storm formation with a dry layer near 700 hPa. If air is able to be lifted enough to rise above the low LFC, storm formation could be prevented when that air encounters dryness above the LFC.

Fig. 2.11 shows the $\theta_e$, $\theta_v$, and frontogenesis parameter for both members at 12UTC. We calculated $\theta_e$ as previously described (Equation 2.3). We calculated $\theta_v$ using Equation 2.7 (**?**) via the metpy package, where $\theta$ is the potential temperature, $\varepsilon$ is the ratio of water vapor to dry air, and $r_v$ is the water vapor mixing ratio. We calculated frontogenesis (F) with $\theta_v$ using Equation 2.3.21 from **?** and neglecting the diabatic heating terms (both for simplicity and due to the lack of heating output by the model; Equation 2.8), where $u$, $v$, and $w$ are the velocities in the $x$, $y$, and $z$ directions, respectively.

$$\theta_v = \theta \frac{1 + r_v/\varepsilon}{1 + r_v} \approx \theta \left(1 + 0.61 r_v\right) \tag{2.7}$$

$$\begin{aligned} F = & \frac{1}{|\nabla \theta_v|} \left( \frac{\partial \theta_v}{\partial x} \left\{ - \left( \frac{\partial u}{\partial x} \frac{\partial \theta_v}{\partial x} \right) - \left( \frac{\partial v}{\partial x} \frac{\partial \theta_v}{\partial y} \right) - \left( \frac{\partial w}{\partial x} \frac{\partial \theta_v}{\partial z} \right) \right\} \right. \\ & + \frac{\partial \theta}{\partial y} \left\{ - \left( \frac{\partial u}{\partial y} \frac{\partial \theta_v}{\partial x} \right) - \left( \frac{\partial v}{\partial y} \frac{\partial \theta_v}{\partial y} \right) - \left( \frac{\partial w}{\partial y} \frac{\partial \theta_v}{\partial z} \right) \right\} \\ & \left. + \frac{\partial \theta}{\partial z} \left\{ - \left( \frac{\partial u}{\partial z} \frac{\partial \theta_v}{\partial x} \right) - \left( \frac{\partial v}{\partial z} \frac{\partial \theta_v}{\partial y} \right) - \left( \frac{\partial w}{\partial z} \frac{\partial \theta_v}{\partial z} \right) \right\} \right) \end{aligned} \tag{2.8}$$

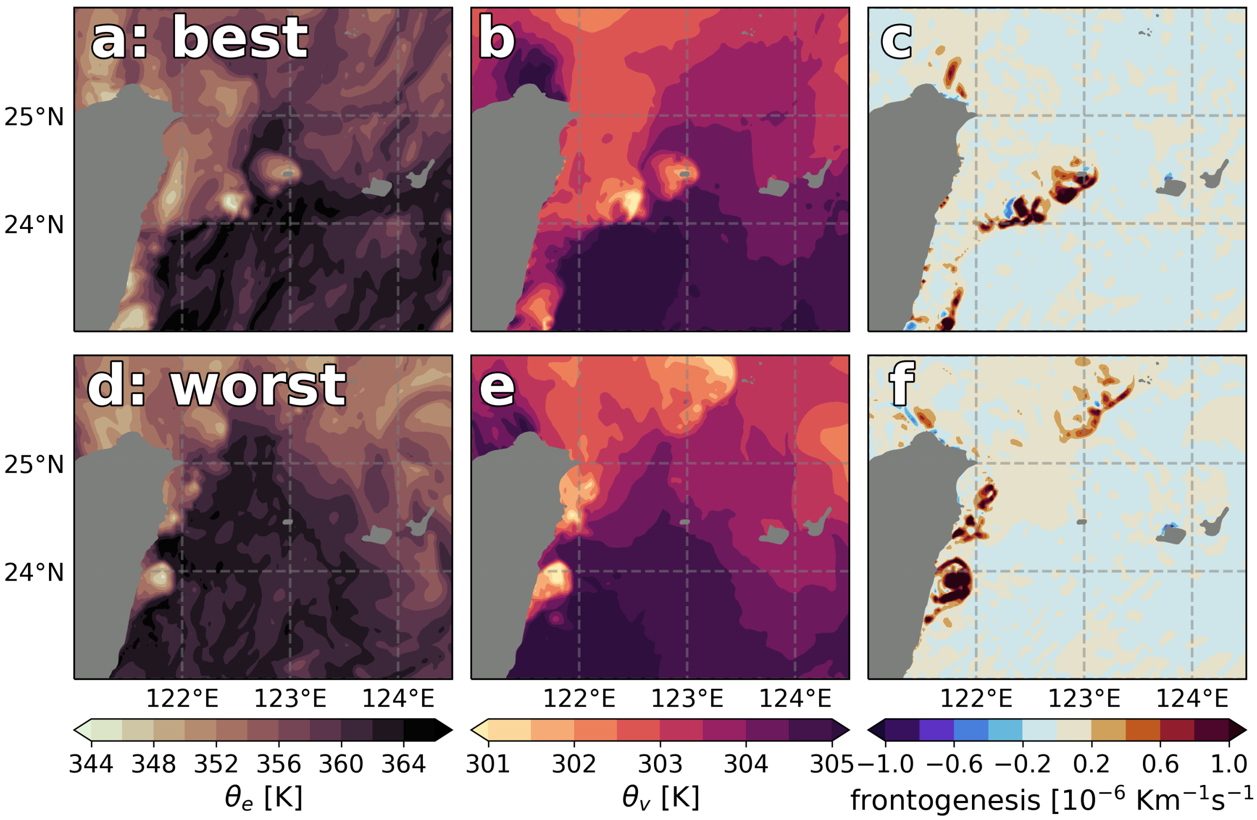


FIG. 2.11. **(a)** Map of 100 m $\theta_e$ for the best member at 12UTC on the 3 km domain. **(b)** As in (a) for $\theta_v$. **(c)** As in (a) for the frontogenesis parameter. **(d)** As in (a) for the worst member. **(e)** As in (a) for the frontogenesis parameter in the worst member. **(f)** (left), $\theta_v$ (middle), and the frontogenesis parameter (right) for the best member (top row) and the worst member (bottom row) on the 3 km domain at 12UTC.

Frontogenesis is the increase of the horizontal thermal gradient with time. The frontogenesis parameter is useful for identifying the formation of fronts or frontal zones. Virtual potential temperature ($\theta_v$) is a good proxy for density and is also useful in frontal identification, as it highlights regions in which parcels can become positively buoyant through lateral movement. In both $\theta_e$ and $\theta_v$, the gradients are the important components in Fig. 2.11. Panels (a) and (b) and (d) and (e) highlight the similarities between the $\theta_e$ gradients and $\theta_v$ fronts, with the primary difference being that the differences across the $\theta_e$ gradients are on the order of 15-20 K while the differences across the gradients associated with $\theta_v$ fronts are closer to 4-5 K, indicating that moisture aspect of the gradients is stronger than the temperature aspect. Together, these two variables both highlight the existence of a weak mei-yu front in the region that can provide an uplift mechanism for the moist air moving northward/northeastward.

By zooming in on the area of interest, we can highlight what is taking place more specifically. In both $\theta_e$ and $\theta_v$, we see the differences in both location and strength of the frontal boundary. In the best member, both variables highlight the southeastward extension of the boundary, as far east as Yonaguni at 123E and as far south as 24N. The frontal boundary in the worst member is more confined to the coast of Taiwan, not extending much past 122E and only dropping south of 24N along the coast. This difference in frontal boundary location, combined with aforementioned moisture differences, strongly contributes to the difference in intensity and location of deep, moist convection between the two members. The frontogenesis parameter shows that frontogenesis is indeed taking place along the gradients shown in $\theta_e$ and front identified by $\theta_v$. The weak frontal forcing evident by the frontal boundaries and frontogenesis parameter is necessary to release the conditional instability contained in the parcels and shown in the soundings. Once the conditional instability is released, the deep, moist convection is able to grow and generate the storms responsible for the extreme rainfall seen in the best member. The confinement of the strong regions of frontogenesis in the worst member to the coast of Taiwan suggests that storms were not able to make it far off the coast of Taiwan, which accounts for the majority of the total rainfall in the worst member remaining close to the coast.

It is important to acknowledge that frontogenesis and convection often occur in a positive feedback cycle, such that the frontogenesis parameter in Fig. 2.11 may be highlighting the existence of convection where stronger positive values are seen. However, whether the frontogenesis is occurring as pure frontogenesis or as an artifact of the convection does not negate the conclusion that the positioning of the $\theta_e$ gradients and $\theta_v$ fronts are contributing to the deep, moist convection in the best member in the area of interest (or lack thereof in the worst member). If the frontogenesis is occurring as an artifact of the convection, its placement still cements that the $\theta_e$ gradients and $\theta_v$ fronts are providing strong uplift mechanisms contributing to that convection due to the co-location between the variables.

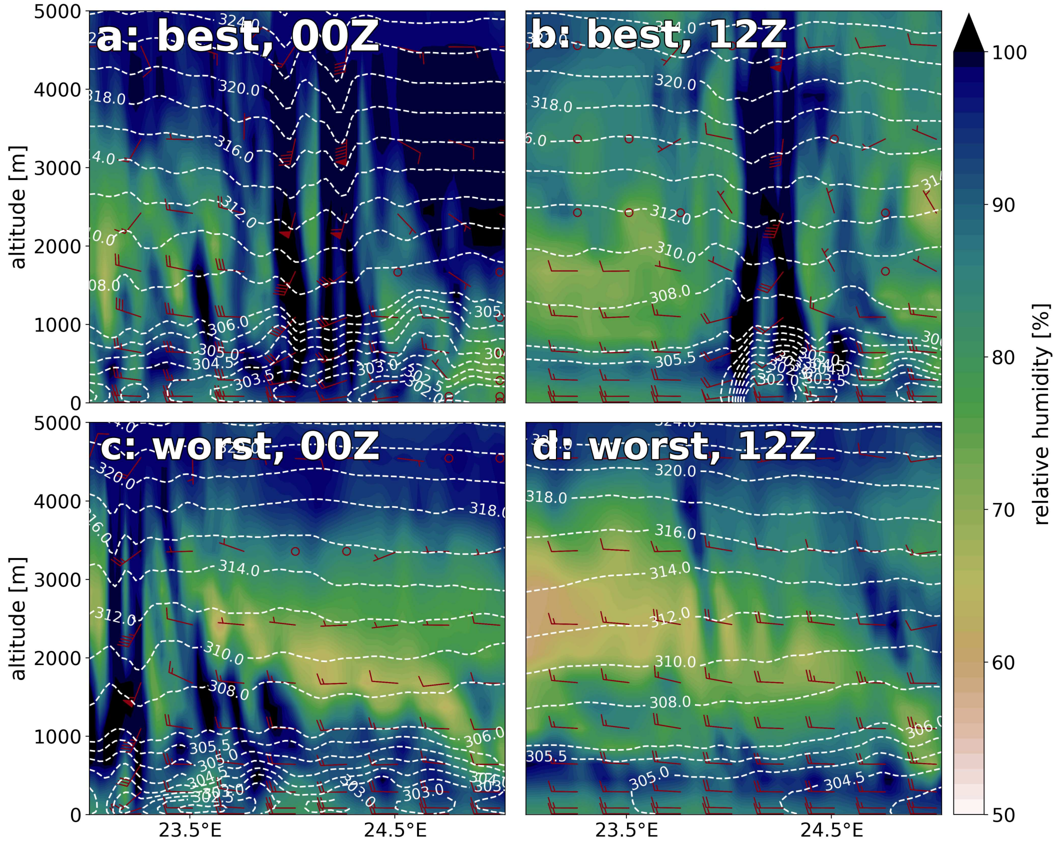


FIG. 2.12. **(a)** Vertical cross section of relative humidity (shading) and $\theta_v$ (white dashed contours) taken along the red line in Fig. 2.8 for the best member at 00UTC. Meridional/vertical wind vectors in units of knots are plotted in red wind barbs, with the vertical component multiplied by a factor of 10 to highlight vertical motion. **(b)** As in (a) at 12UTC. **(c)** As in (a) for the worst member at 00UTC. **(d)** As in (a) for the worst member at 12UTC.

Fig. 2.12 shows a cross-section taken along the red line in Fig. 2.8a with the relative humidity shown in shading, $\theta_v$ isentropes shown in white dashed contours, and wind shown in the red wind barbs in the lower 5 km of the atmosphere. $\theta_v$ is filtered using a 2-dimensional Gaussian filter with $\sigma$=1, using the SciPy package (**?**). Note that wind vectors are plotted as the meridional wind and 10 times the vertical wind to highlight the vertical wind velocities due to the relative difference in magnitude of the two variables.

In the RH field in Figure 2.12, we see much drier, up to 40% lower RH, air in the worst member as compared to the best member. We also see more vertical cores of moisture coupled with strong updrafts in the best member, representative of clouds. The worst member does not have as many vertical moisture cores and where they do exist, we see weaker downdrafts and an absence of updrafts. Similar to the $\theta_e$ gradient in Figure 2.8, in Figure 2.12 we see a $\theta_v$ front in a similar location. The shift of the $\theta_v$

isentropes from 00 to 12 UTC clearly outlines this front and its evolution. In the best member, we see the isentropes with a strong horizontal gradient shift northward and the gradient strengthens from 00 to 12 UTC, highlighting the intensification of the $\theta_v$ front. As warm, moist air moves north/northeastward with the southerly/southwesterly winds, it enters a region of cooler, drier air, where it is positively buoyant, such that this movement of the warmer, moister air along the $\theta_v$ isentropes acts as the primary lifting mechanism. However, in the worst member the tightening and shifting of the $\theta_v$ isentropes is not present, so warmer, moister air moving into the region is unaffected by the same uplift mechanism as in the best member. The lack of an uplift mechanism in the worst member due to the $\theta_v$ front is also highlighted by the lower number of vertical moisture cores and the lack of strong updrafts, or updrafts at all along the cross-section. Figure 2.12 suggests that it is a combination of the $\theta_v$ front and the moisture that leads to the moist, deep convection in the best member, not necessarily one factor or the other. Both members exhibit plentiful warm, moist air in the region of interest, but the drier air in the worst member, coupled with the lack of an uplift mechanism, contributes to the reduced production of heavy rainfall.

#### *2.3.4 Comparison to other members*

In order to confirm that the conclusions drawn from the best and worst member were not only applicable to those members, we further compared the members deemed to be second best (“better”) and second worst (“worse”). As was discussed previously, the rainfall plots in Fig. 2.3 show that the best and better members had similar rainfall patterns both in orientation and quantity, as did the worst and worse members. If the best and better members also showed similar patterns in meteorological fields, as did the worst and worse, it would be reasonable to apply the physical interpretations from the best and worst members more broadly. This extension of analysis to the better and worse members allows us to come closer to generalizing our conclusions, short of running further simulations that change the moisture profile and location of the $\theta_v$ front, which are out of the scope of this study.

In repeated analysis of the second best ("better") and second worst ("worse") members, whose rainfall patterns are shown in Fig. 2.3, we find that the better and worse member are qualitatively similar in all fields examined to the best and worst members (not shown here). We find some quantitative differences between better and best and between worst and worse, but the differences are not remarkable.

In plots of relative humidity and differences for the second best ("better") and second worst ("worse") members (as in Fig. 2.9 and 2.10; not shown here), patterns for the better and worse members reflect those in the best and worst members. At 700 hPa, the better member has 20-25% greater relative humidity values than the worse member in the region of interest, and the worse member has more organized, more southerly flow, which allows it to advect more dry air into the region of interest. At 850 hPa, we find that the better member has a large region of higher moisture than the worse member, and the winds in the better member are predominantly westerly/southwesterly, while the winds in the worse member are predominantly southerly/southwesterly. These differences highlight the same patterns shown in panels c and f of Fig. 2.9 and 2.10: the best/better members are not advecting as much dry air into the region of interest as the worst/worse members at both 700 and 850 hPa. The similarities shown between the top two most accurate members and the bottom two least accurate members, and their respective differences, gives us confidence in our conclusions surrounding what lead to the bifurcation in rainfall totals among ensemble members.

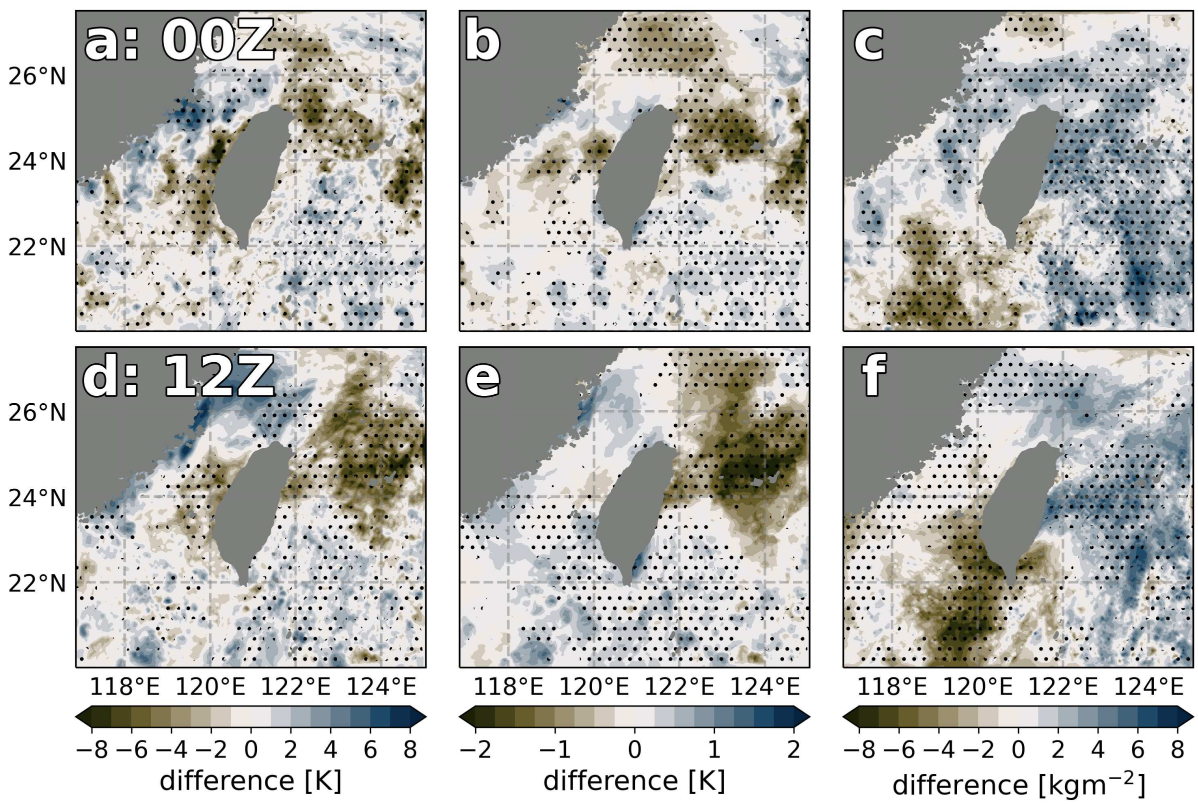


FIG. 2.13. **(a)** Map of the difference in 100 m $\theta_e$ between the mean of the top five and mean of the bottom five ensemble members at 00UTC. **(b)** As in (a) for the difference in 100 m $\theta_v$ at 00UTC. **(c)** As in (a) for the difference in integrated water vapor at 00UTC. **(d)** As in (a) at 12UTC. **(e)** As in (b) at 12UTC. **(f)** As in (c) at 12UTC.

In order to further generalize our conclusions and have a broader view of the most accurate and least accurate members in the ensemble, we analyzed plots showing the differences between the mean of the five most accurate members and the five least accurate members (evaluated on how well they reproduce rainfall totals and patterns, selected using the method outlined in Section 2.2.3) for different meteorological fields and the statistical significance of those differences. We calculated statistical significance using a paired t-test as shown in Equation 2.9,

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_{1,2}}{\hat{\sigma}\sqrt{\frac{1}{N_1^2} + \frac{1}{N_2^2}}}, \hat{\sigma} = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}} \tag{2.9}$$

where $x_1$ and $x_2$ are the means, $N_1$ and $N_2$ are the sample sizes, and $s_1$ and $s_2$ are the standard deviations of the top five and bottom five members, respectively. $\Delta_{1,2}$ is the difference between the samples, set to 0 to test the null hypothesis that the samples come from the same population. On Fig. 2.13, any area covered by black stippling represents an area where the difference between the mean of the top five members and the bottom five members is statistically significant at the 99.95% confidence level. These areas indicate regions of high confidence in the differences between the meteorological fields of most accurate and least accurate members.

The stippling for all three variables tends to be strongest over the regions with large differences, both positive and negative (Figure 2.13). Since our chosen test was the paired *t*-test, the statistical significance indicates the most and least accurate members likely come from different populations. The likelihood that these most and least accurate members come from different populations provides additional confidence that there is a bifurcation occurring within the ensemble.

The difference plots shown in Fig. 2.13 also confirm what we have concluded in previous analysis: for both $\theta_e$ and $\theta_v$ (panels a/d and b/e, respectively), the negative differences in temperatures highlight that the near-surface fronts are stronger and more contained within the region of interest in better-performing members; and in integrated water vapor (panels c and f), the members that most accurately reproduce the rainfall pattern and totals are more moist throughout the vertical column at 0 and 12UTC than those that least accurately reproduce the rainfall pattern and totals.

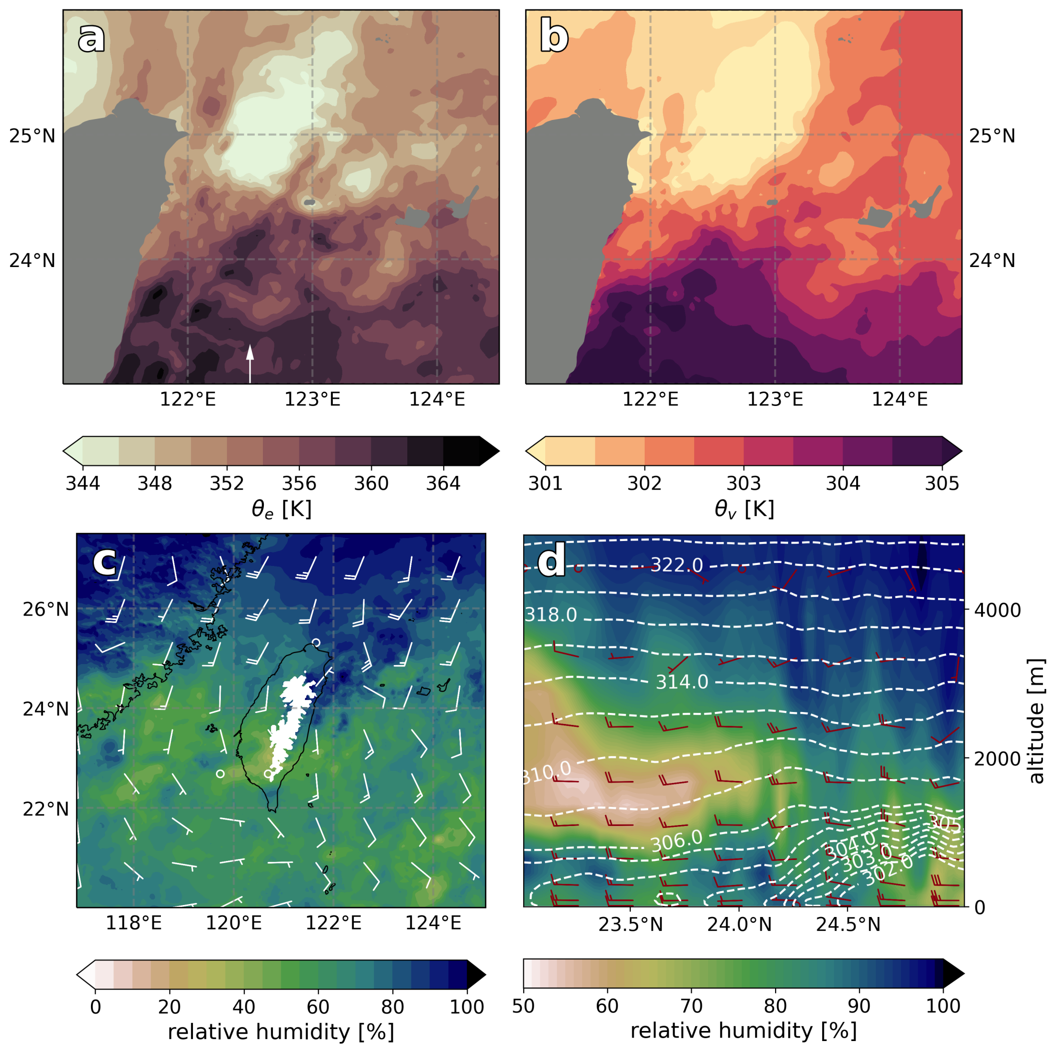


FIG. 2.14. **(a)** Map of 100-m $\theta_e$ for the verification ensemble mean at 12UTC 09 June 2020. **(b)** As in (a) for 100 m $\theta_v$. **(c)** Map of 850 hPa relative humidity for the verification ensemble mean at 12UTC 09 June 2020. **(d)** Vertical cross section of relative humidity (shading) and $\theta_v$ (white dashed contours) taken along the red line in Fig. 2.8 for the verification ensemble mean at 12UTC 09 June 2020. The latitude along which the cross-section is taken is highlighted in (a) by a white arrow along the x-axis. Meridional/vertical wind vectors in units of knots are plotted in red wind barbs, with the vertical component multiplied by a factor of 10 to highlight vertical motion.

While examining 10 out of 40 members in the ensemble does not cover all the possible solutions, it allows us to see the features that are occurring most often in the most extreme ends of the ensemble and leave out analysis of those that were in the middle of the two. This approach also highlights the importance of analyzing different extremes in an ensemble, as analysis of the mean alone can sometimes obscure meteorological features, especially on the mesoscale as seen in this analysis.

Given the prior analysis, we aim to understand how well the best members in the ensemble verified. To accomplish this verification, we analyze the ensemble mean output from the PSU WRF-EnKF ensemble whose forecast initialized at 12 UTC 9 June 2020. The analysis comes immediately after 12 h of data assimilation spin up and can be considered to be the best estimate of the atmospheric state and used as a verification. We focus here on the mean at the 12 UTC analysis time (hereafter the "verification ensemble mean"). We choose to analyze 100-m $\theta_e$ and $\theta_v$, as in Figure 2.11, and the vertical cross-section of relative humidity, $\theta_v$, and winds, as in Figure 2.9, in order to focus on the near-surface frontal boundary and the low- to mid-level vertical profile of moisture in the region of interest.

Panels a and b in Figure 2.14 show the existence of a sharp near-surface front concentrated near Yonaguni in the verification ensemble mean at 12 UTC. The gradients are apparent in $\theta_e$ and $\theta_v$ with similar positioning, and the front also shows up clearly in panel d in the $\theta_v$ isentropes from approximately 24.0° N to 24.5° N. The frontal strength and positioning in the verification ensemble mean at 12 UTC are stronger on the mesoscale than that of the best member and the mean of the top five members (from the forecast) shown in Figures 2.7, 2.11 and 2.14. While the absolute gradients are similar to the best member, they are confined to a much smaller scale consistent with their production by convection. In the verification, the front is stronger over a larger area.

The moisture profile from the surface to the mid levels differs from the well-performing members and is actually more visually similar to the worst and poorly-performing members. From approximately 23.0° N to 24.0° N, there is an intrusion of dry air, as dry as ∼50% relative humidity, from the south that spans 1000 m above the surface up to 2000–3000 m above the surface. This dry air tongue appears in the worst member at 12 UTC (Figure 2.9d) at a similar magnitude as the verification ensemble mean at 12 UTC, but is much weaker or not present at 12 UTC in the best member (Figure 2.9b).

Panel c in Fig. 2.14 shows the 850 hPa relative humidity in the verification ensemble mean at 12UTC. Note the difference in colorbar scale between panels c and d. The best and worst member (Fig. 2.9 panels d and e) both capture the drier air south of ∼23°N that is visible in the verification ensemble mean well, in both magnitude and location. Focusing on the areas of higher relative humidity in the verification ensemble mean, near and northeast of the region of interest (near 24°N, 122°E), the worst member's RH magnitude and pattern more closely matches those of the verification ensemble mean. While the best member more accurately captures the moister air north of ∼26°N, it places the moister air in and near the region of interest slightly too far south. The closeness of the worst member's

850 hPa RH profile to that of the verification ensemble mean suggests that more accurate simulation of relative humidity profiles may not be critical to accurate reproduction of rainfall totals and patterns.

The presence of the mid-level dry air tongue in the worst member and verification ensemble mean, and lack thereof in the best member, suggests that the vertical profile of moisture, particularly in the mid levels, is not, in fact, as important to the production of extreme rainfall as our analysis of RH differences in Fig. 2.9 and 2.10 suggested. Rather, the strength and location of the near-surface front, shown in $\theta_v$, is the critical component leading to the production of deep, moist convection and henceforth widespread extreme rainfall. It is likely that the vertical profile of moisture throughout the entire column remains a necessary factor for deep, moist convection, but that it is not sufficient to produce such convection without a lifting mechanism.

## 2.4 Conclusions

This study analyzed output from a mesoscale WRF ensemble simulation of an extreme rainfall event associated with a relatively weak mei-yu front on 09 June 2020. Areas east of Taiwan, including Yonaguni, Japan, saw 24-h widespread accumulation over 100 mm and peak rainfall accumulation greater than 600 mm caused by quasi-stationary, back-building MCSs. The nsemble suggested a bifurcation in possible outcomes, with some members producing heavy rain in the correct location and others with little to no rain in the wrong location. We assessed this event as "liminal" with conditions near a threshold to produce extreme rain. We selected the best- and worst-performing ensemble members according to rainfall totals and patterns by using a combination of equitable threat score and relative operating characteristics. We performed mesoscale meteorological analysis focused on the singular best and worst performing members and extended this analysis to the top five and bottom five performing members to find similarities and emerging patterns. We identified consistent patterns amongst the best-performing members, worst-performing members, and differences between them in rainfall intensity and patterns, relative humidity, and near-surface virtual and equivalent potential temperatures.

Differences between the better-performing and worse-performing members are largely apparent on the mesoscale, with minimal differences on the synoptic scale. Ensemble members on both ends of the spectrum feature moisture rich environments, surface lows south of the region of interest, and moist soundings with low altitude levels of free convection. Our analysis reveals that the location and intensity of the shallow near-surface $\theta_v$ frontal boundary are key to ensemble member accuracy. In the two best members, south-southwesterly flow ascends along the sloped isentropes with a stronger horizontal gradient in $\theta_v$, rises above the LFC, releasing conditional instability. In the two worst members, the near-surface gradients are weaker and more confined along Taiwan's coast, which leads to less rainfall displaced west of the observed rainfall. In addition, as the less accurate simulations progress, stronger southerly winds advect drier mid-level air into the region of interest and shift the near-surface boundary further north and west. Winds in the best-performing members are weaker and more southwesterly with less dry air advection and frontal displacement.

Verification using the analysis ensemble mean from the subsequent model run (12 UTC 9 June 2020, after data assimilation) reveals the importance of certain mesoscale features over others. The verification ensemble mean at 12 UTC has a strong, sharp near-surface $\theta_v$ frontal boundary, with a larger spatial scale than that seen in the well-performing members. However, the verification ensemble mean at 12 UTC also features drier air aloft, between 1000 m and 3000 m above the surface, which is more similar to the dry air patterns found in the poorly-performing members. Given the accuracy in reproducing rainfall of the well-performing members and the similar frontal placement between those members and the verification ensemble mean at 12 UTC, this analysis of the verification ensemble mean suggests that the positioning of near-surface fronts is more critical to the production of deep, moist convection than the mid-level dry air is detrimental to that production for this liminal case.

Fortunately, the environmental impacts of this case were not severe due to the location of the heaviest rainfall. A slight displacement in the peaks of the heavy rain could have led to major flooding on the islands of Yonaguni and Ishigaki or Taiwan. A similar case occurred on 13 May 2019, in which nearly 400 mm of rainfall accumulated on Yonaguni in a 24-h period, leading to widespread flooding across the island and the labeling of the case as a 50-year flooding event by the Japanese Meteorological Agency. The case in May 2019 serves as a good example of the possible impacts of the case in June 2020 had the majority of the heavy rain been over land.

As extreme mei-yu seasons are expected to become more frequent (**?**), the use of model ensembles, such as the PSU WRF-EnKF ensemble used in this study, will become more important for forecasters in realizing the potential heavy rainfall outcomes associated with extreme mei-yu seasons. Future work should focus on analysis of other cases with similar liminal synoptic and mesoscale conditions in order to be able to further generalize these results for bifurcating forecasts. A better understanding of the mesoscale features that led to this case and similar cases is crucial in improving heavy rainfall modeling and predictability. When there is weak synoptic forcing present, it is especially important to understand the mesoscale features present and how they may interact. Our study suggests that increased accuracy in simulating the location and movement of near-surface boundaries is critical to improved accuracy in forecasts. Increased spatial and temporal frequency of observations, especially of near-surface moisture, and data assimilation of those observations would also contribute to improving forecasts. Especially given the findings of **?**, which suggest the quantitative difficulties of forecasting extreme mei-yu seasons, modeling improvements will be critical. Special observations will be conducted in this region during the Prediction of Rainfall Extremes Campaign in the Pacific in spring and summer 2022 to contribute to these goals and future forecast= improvements.

# CHAPTER 3

## A case study of a high-moisture, low rainfall event in Northern Colorado

### 3.1 Introduction

The Prediction of Rainfall Extremes Campaign in the Pacific (PRECIP) is a field campaign set to occur in May-August 2022 in the northwestern Pacific and will be based in Taiwan and Japan. PRECIP was delayed due to the COVID-19 pandemic. In both 2019 and 2020, the PRECIP PIs and science teams conducted preliminary/dry run field campaigns in order to test modeling, observational, and analysis methods to be best prepared for the in-person PRECIP campaign. The PRECIP 2020 Pilot Study focused on modeling over the selected domains for PRECIP using both PSU WRF-EnKF and MPAS as its primary models. Preparatory Rockies Experiment for the Campaign in the Pacific (PRE-CIP) 2021 focused on a combination of modeling and observational methods, employing the PSU WRF-EnKF system over the continental US and northern Colorado, where field teams launched weather balloons with Vaisala soundings on days of interest (those with high moisture, rain forecast, or both) and monitored between two and three radars covering the area, Colorado State University's CHIVO and CHILL (CSU-CHIVO and CSU-CHILL) and NCAR's S-Pol. Northern Colorado's dry environment provided a valuable contrast to the moist environment found in the northwestern Pacific and allowed for comparison between heavy rainfall in both environments.

This chapter focuses on preliminary analysis of a high-moisture, low rainfall case that took place on 31 July 2021 over Northern Colorado. As a result of the PRE-CIP 2021 campaign, this case has a large amount of observational and modeling data available for analysis, including three soundings, 9 km and 3 km nested PSU WRF-EnKF ensemble simulations, and radar data. We selected this case due to its contrast with the case examined in Chapter 2. In Chapter 2, we focused on a case in which the models forecast minimal rain and largely failed to accurately capture the intensity and location of extreme rainfall that fell in reality. On 31 July 2021, models and forecasters expected extreme rainfall to accumulate over our area of interest and, despite anomalous moisture in the area, little to no rainfall accumulated. Section 2 will outline the data and methodology used, Section 3 will provide a case summary and preliminary analysis of meteorological variables, and Section 4 will provide brief conclusions and suggestions for future work.

## 3.2 DATA AND METHODOLOGY

### *3.2.1 Model description*

This study analyzes the model output from select members of a 40-member ensemble that used the Weather Research and Forecasting (WRF) model version 4.1.3 (**?**). The ensemble member selection process is explained in subsection 3.2.3 below. The 40-member ensemble was created using the Penn State WRF ensemble Kalman filter (PSU WRF-EnKF) modeling system, which uses data assimilation and perturbations in the model initial conditions to generate the ensemble (**???**). The PSU WRF-EnKF system uses two one-way nested domains with spatial resolutions of 9 km and 3 km. We will focus our analysis on the 3 km domain, which contains 330x390 horizontal grid points and 50 vertical levels. The data assimilation runs for 12 hours preceding the beginning of the forecast period, at which point a 12-h forecast with hourly output is generated for all 40 members and a 48-h forecast with hourly output is generated for 20 members. Our analysis will focus only on the 20 members whose forecasts extend to 48 hours. The system assimilates conventional meteorological observations, as well as satellite brightness temperatures for the inner 3 km-resolution domain, and employs Thompson microphysics (**?**), MYNN surface and planetary boundary layer physics (**?**), RRTMG shortwave and longwave radiation (**?**), and the Noah surface model (**?**). Fig. 3.1 shows the WRF 9 km and 3 km domain boxes (in red) and the region of interest for this case study (in blue).

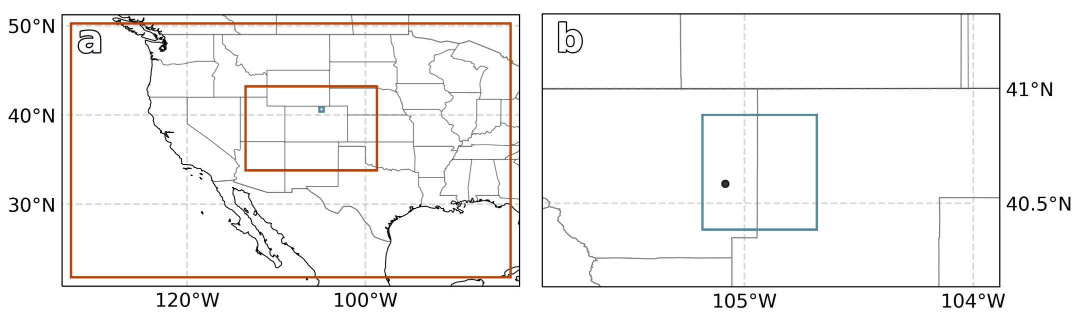


FIG. 3.1. **(a)** The red boxes represent the 9 km (outer) and 3 km (inner) WRF domains used in the PSU WRF-EnKF simulations. The blue box shows the region for ensemble statistical calculations and soundings, also called the "region of interest". **(b)** This panel zooms within the 3 km domain to better highlight the region of interest. The blue box is as in (a), and Fort Collins, CO is marked by a solid black dot.

### *3.2.2 Observational and validation data*

On 31 July 2021, the PRE-CIP field team launched three weather balloons from Christman Airfield in Fort Collins, CO with two types of vertical profilers attached: a Vaisala sounding system and a National

Taiwan University (NTU) storm tracker (**?**). Analysis in this chapter focuses on the soundings generated from the Vaisala data, though future analysis should incorporate data from the NTU storm trackers. Balloons were launched at 1446, 1741, and 2047 UTC (8:47 am, 11:41 am, and 2:47 pm MST) and their associated soundings will be referred to as the 15, 18, and 21Z soundings, respectively.

We do not examine radar data from CSU-CHILL, CSU-CHIVO, or S-Pol here, but future work should include analysis of radar data, especially to track storm motion and evolution.

In order to quantitatively and spatially examine ensemble members' rainfall, we compare model output with rainfall estimates provided by the Multi-Radar/Multi-Sensor System (MRMS), a gridded, operational dataset with hourly output and 0.01° by 0.01° spatial resolution that uses a combination of radar data, surface and upper air observations, lightning detection, satellite observations, and forecast model data to estimate ground-based rainfall totals. For preliminary analysis, the MRMS data was not manipulated from its raw form, but future analysis should interpolate the MRMS data to match the native WRF grid in order to perform point calculations, following the methodology outlined in Section 2.2.3 of Chapter 2.

### *3.2.3 Ensemble member selection*

In this case, we are primarily interested in the region surrounding Fort Collins, where heavy rainfall was forecast but did not occur, approximately the box from 105.18°W to 104.68°W and 40.39°N to 40.89°. For our preliminary analysis, we selected the best performing and worst performing ensemble members by comparing the maximum rainfall totals over the 24-h period from 06Z 31 July 2021 to 06Z 1 August 2021 (12:00 am 31 July 2021 to 12:00 am 1 August 2021 MST) from the ensemble to the total rainfall over that period from MRMS. Fig. 3.2 shows the ensemble members' maximum rainfall within the region of interest in blue if they are below and yellow if they are above the MRMS maximum rainfall within the region of interest, shown in the dashed red line.

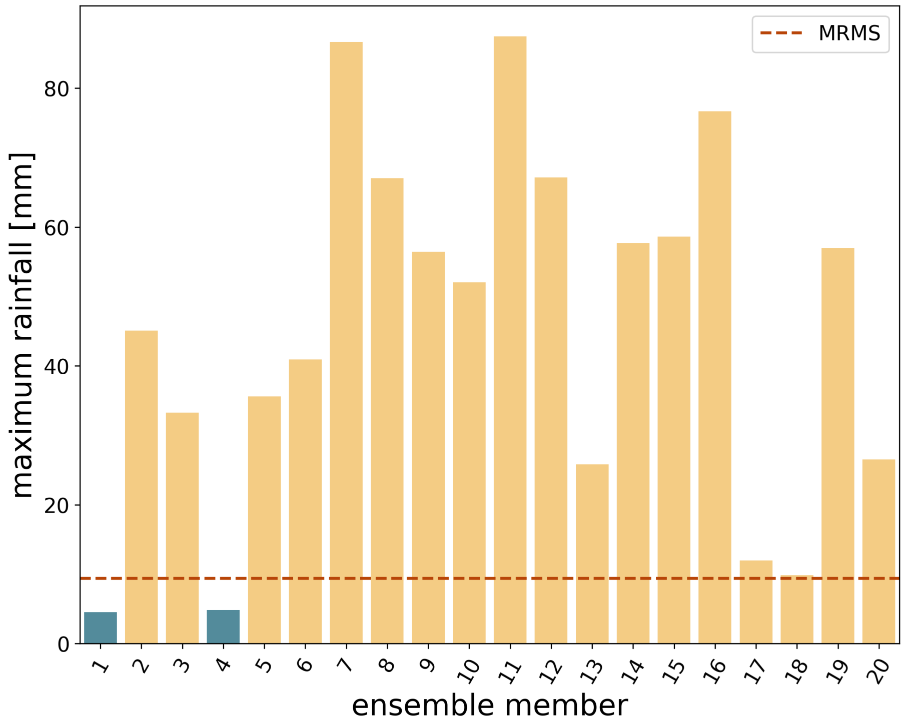


FIG. 3.2. Maximum accumulated rainfall totals over the 24h period beginning at 6Z 31 July 2021 calculated over the blue box in Fig. 3.1. Yellow shading denotes maximums greater than the MRMS maximum, blue shading denotes maximums lesser than the MRMS maximum, and the red dashed line denotes the MRMS maximum.

As a first pass, we selected those members that had maximums less than or close to that in MRMS as the best performing members and those that had the greatest maximums as the worst performing members. We visually compared the selected members' 24-h accumulated rainfall patterns over a broader domain (shown in Fig. 3.3, representative of the spatial span of the CHIVO radar) to that of MRMS in order to qualitatively select the best performing and worst performing member for further meteorological analysis. As shown in Fig. 3.3, we selected member 017 as the best member and member 007 as the worst member.

## 3.3 ANALYSIS

### *3.3.1 Case summary*

On 31 July 2021, the majority of Colorado was under a slight risk (10-20%) of rainfall exceeding flash flood guidance according to the National Oceanic and Atmospheric Administration's (NOAA) Day 1 Excessive Rainfall Guidance, issued at 1556Z (9:56 am MST) Sat Jul 31 2021. This area included our region of interest in Northern Colorado, and a marginal risk (5-10%) covered the areas not included under the slight risk. The National Weather Service office in Boulder (NWS Boulder) issued a Flash Flood Warning for the Cameron Peak Fire burn area, directly west of our region of interest by approximately 35 miles. Global models forecast heavy rainfall over the burn area and our region of interest, and forecast discussions agreed. One forecast discussion put out by the NWS Weather Prediction Center discussed high MUCAPE (most unstable convective available potential energy), "renewed convective development over the eastern CO/WY border", and "[s]low movement of cells... due to week deep mean-layer flow of 10-15 kt, which may support some 2-3 inch totals through 10Z... some localized flash flooding may result." The field team emailed NWS Boulder the 15Z sounding, to which NWS Boulder responded, "That is a classic heavy rain profile...tall skinny cape, saturated all the way up, low LCL, big warm cloud depth. Very efficient rain processes." (NWS Boulder, personal communication).

The 15, 18, and 21Z soundings in the bottom row of Fig. 3.4 show the atmosphere's evolution throughout the day: the level of free convection (LFC) remained low, near 750 mb; surface, mean-layer, and most-unstable CAPE all increased; and precipitable water (PW) gradually increased. As NWS Boulder said in the email to the field team, the atmospheric profile had all of the features of a classic heavy rain profile, especially anomalous PW values. One day before, July 30, 2021, the atmosphere had a similar profile with lower PW values and heavy rainfall occurred in our region of interest. On 31 July 2021, however, little to no rain occurred in our region of interest, as shown in panel c of Fig. 3.2. We aim to investigate the processes taking place near the surface in order to understand what prevented heavy rainfall from taking place on 31 July 2021 when it seemed that all of the necessary and sufficient ingredients were in place. Following (**?**) and the analysis completed in Chapter 2, we will take an ingredients-based approach to our analysis, focusing first on uplift mechanisms and vertical profiles of moisture.

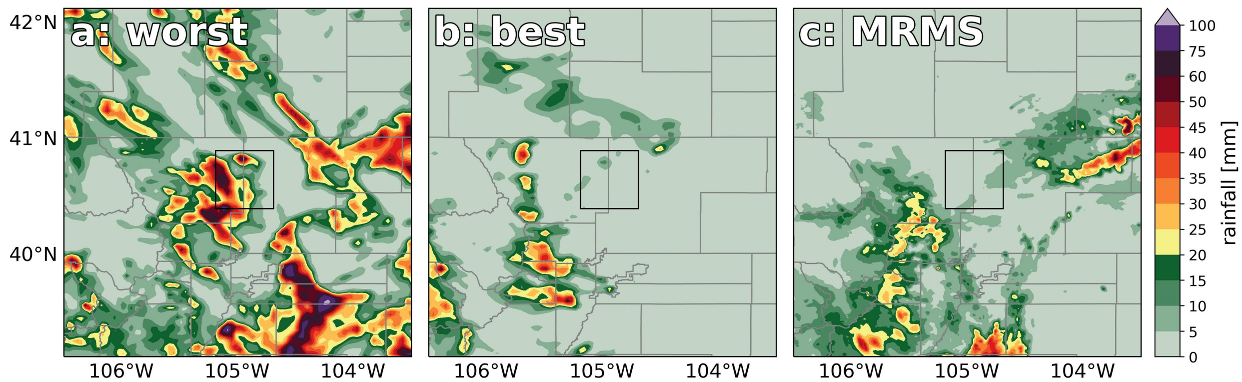


FIG. 3.3. **(a)** Map of rainfall totals for the worst WRF ensemble member for the 24h period beginning at 6Z 31 July 2021. The black box outlines the region of interest around Fort Collins, CO. **(b)** As in panel (a) but for the best WRF ensemble member. **(c)** As in (a) for MRMS.

Rainfall totals in reality peaked around 40 mm in the broader domain and between 5 and 10 mm within the region of interest, despite the forecasts for widespread heavy rainfall and the potential of flooding (Fig 3.3c). While there was lighter widespread rainfall southwest and northeast of Fort Collins, little to no rainfall accumulated within the region of interest. The worst ensemble member produced widespread rainfall with isolated regions of heavy rainfall, with peak totals accumulating at 100 mm in the broader domain and near 85 mm within the region of interest (Fig. 3.3a). The best member was more consistent with reality and produced lighter rainfall west and southwest of Fort Collins, a peak of 10-15 mm in the region of interest, and little to no rainfall to the northeast. While the best member does not perfectly reproduce the rainfall that accumulated in reality, it clearly captures the overall patterns and intensity significantly better than the worst member does. We will investigate near-surface, mesoscale features that may have contributed to the differences in rainfall production in these two members in the following section.

### *3.3.2 Mesoscale analysis*

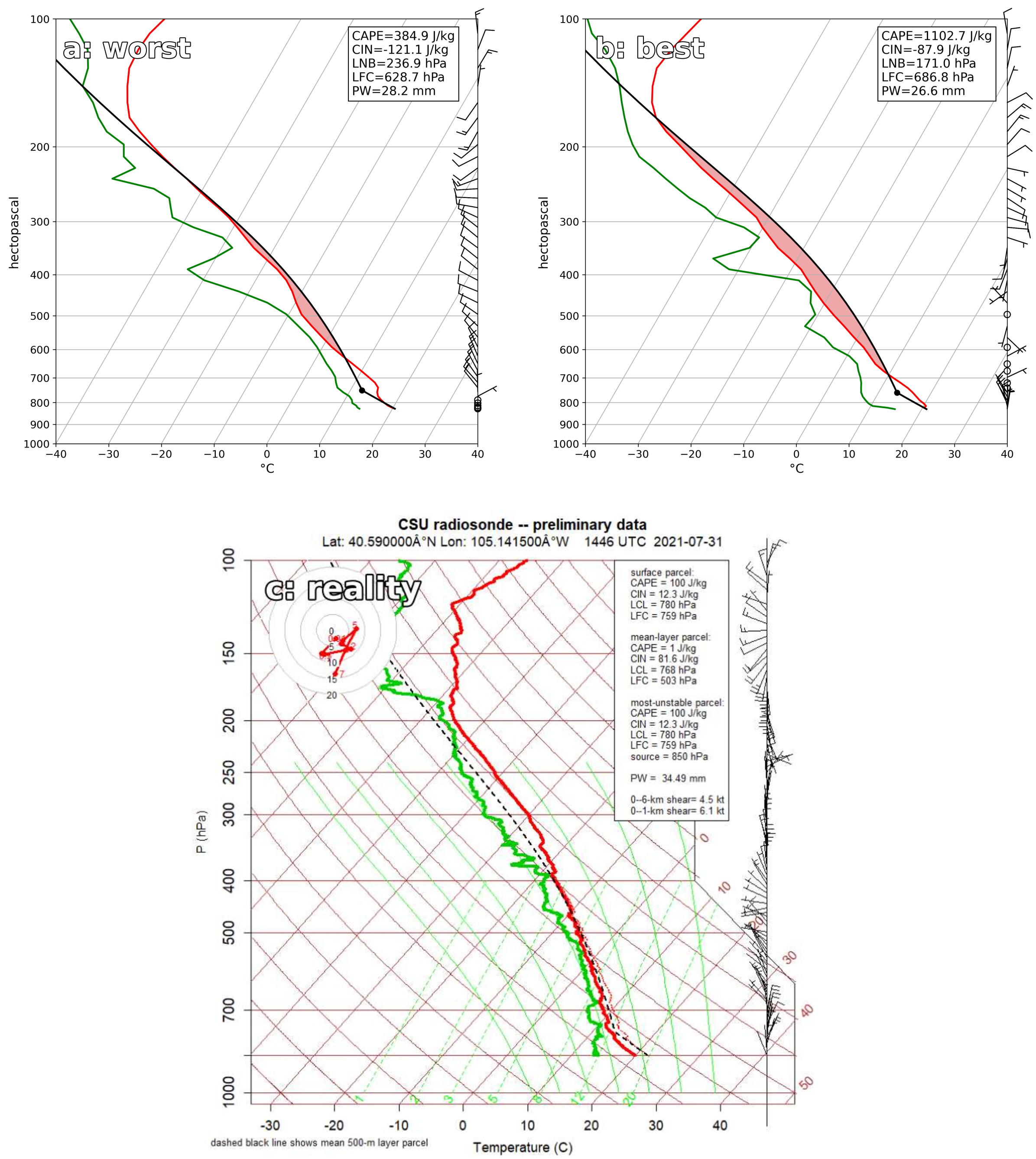


FIG. 3.4. **(a)** Point sounding calculated at approximately 40.5899°N, 105.1416°W, the location of Christman Field, for the best member at 15 Z. **(b)** As in (a) for the worst member. **(c)** As in (a) for the observed sounding from the PRECIP field team at Christman field at 1446 UTC.

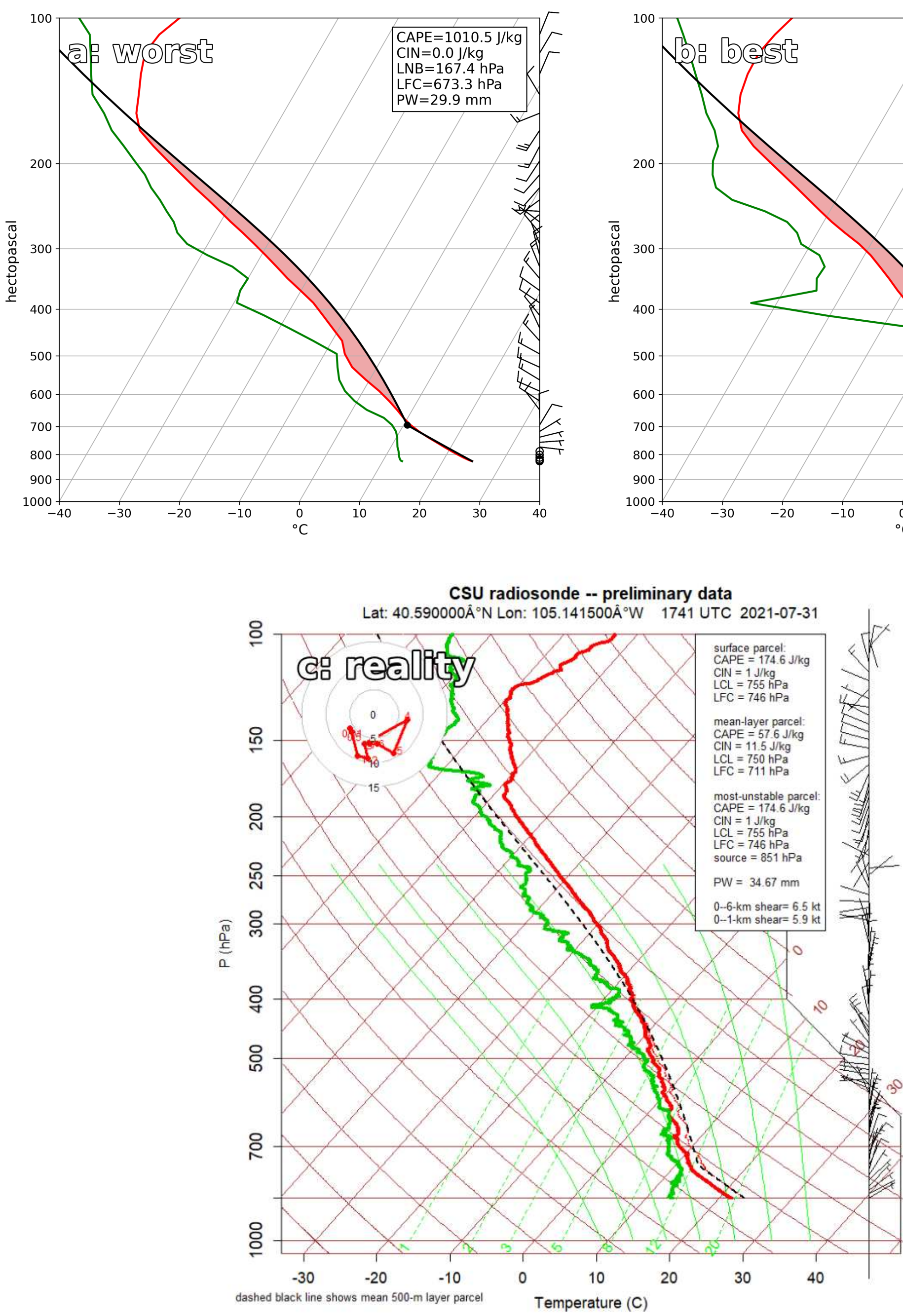


FIG. 3.5. **(a)** Point sounding calculated at approximately 40.5899°N, 105.1416°W, the location of Christman Field, for the best member at 18 Z. **(b)** As in (a) for the worst member. **(c)** As in (a) for the observed sounding from the PRECIP field team at Christman field at 1446 UTC.

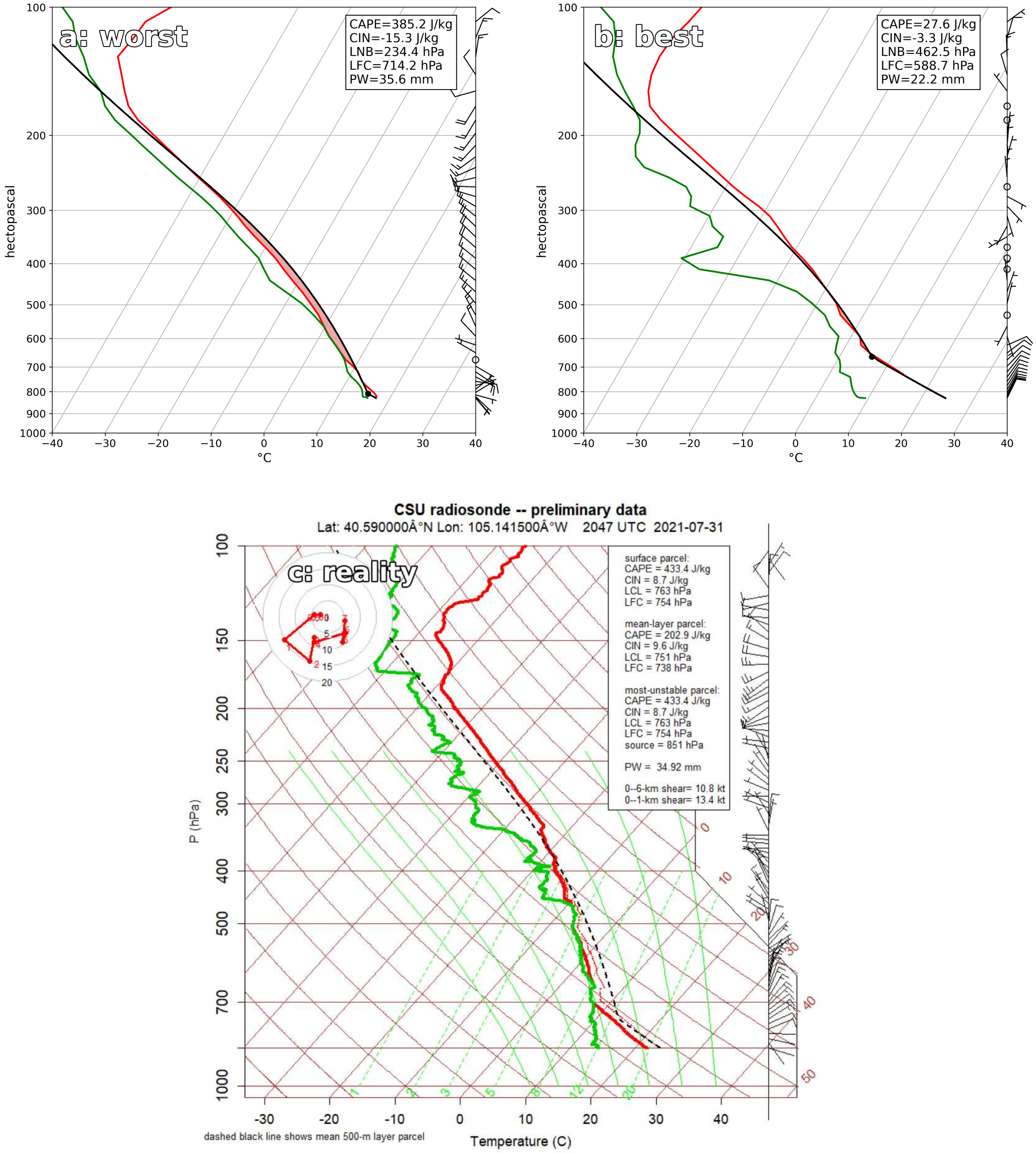


FIG. 3.6. **(a)** Point sounding calculated at approximately 40.5899°N, 105.1416°W, the location of Christman Field, for the best member at 21 Z. **(b)** As in (a) for the worst member. **(c)** As in (a) for the observed sounding from the PRECIP field team at Christman field at 1446 UTC.

Figs. 3.4, 3.5, and 3.6 shows the evolution of the vertical profile at Christman Field over the course of 31 July 2021 at 15, 18, and 21 Z, both in reality (panels a and b in all figures) and in the model (panel c in all figures). Soundings and their respective statistics were calculated using the metpy package (**?**).

Focusing first on the wind profiles, the best member does a remarkable job of accurately reproducing the realistic vertical wind profile at all three times. The wind profile in the worst member, however, is almost a mirror opposite of reality at all three times and all vertical levels, except in the lower atmosphere at 18 and 21 Z. The best member accurately represents the increase in precipitable water over the course of the day, though the increase is more dramatic than in reality. The soundings in reality maintain nearly moist neutral profiles throughout the day, with moderately low levels of CAPE that increase throughout the day (maximizing at 433.4 J/kg of surface-based CAPE). The best members' soundings show moist conditions with a layer of dry air near 400 mb at 15 and 18 Z and a nearly moist neutral profile at 21 Z with no dry air layer. The worst members' soundings show a similar layer of dry air near 400 mb that persists through all times and an overall less moist profile, with precipitable water dipping into the low 20 mm range at 21 Z. The worst member also has larger quantities of CAPE, maximizing near 1600 J/kg at 18 Z, in the first two time steps, then negligible CAPE at the final time step. Though neither member perfectly captures the vertical profile of the atmosphere at all time steps, the soundings produced from the best member's output much more closely match reality than those produced from the worst members' output.

In reality, the sounds are propitious for heavy rainfall but still show conditional instability, which must be released in order for heavy rainfall to occur. There is little convection inhibition (CIN) throughout the day, which removes a barrier for the formation of deep, moist convection. However, there still must be some lifting mechanism in order to release the conditional instability, whether it is the vertical growth of the turbulent boundary layer or mechanical forcing, such as vertical ascent at a front or boundary.

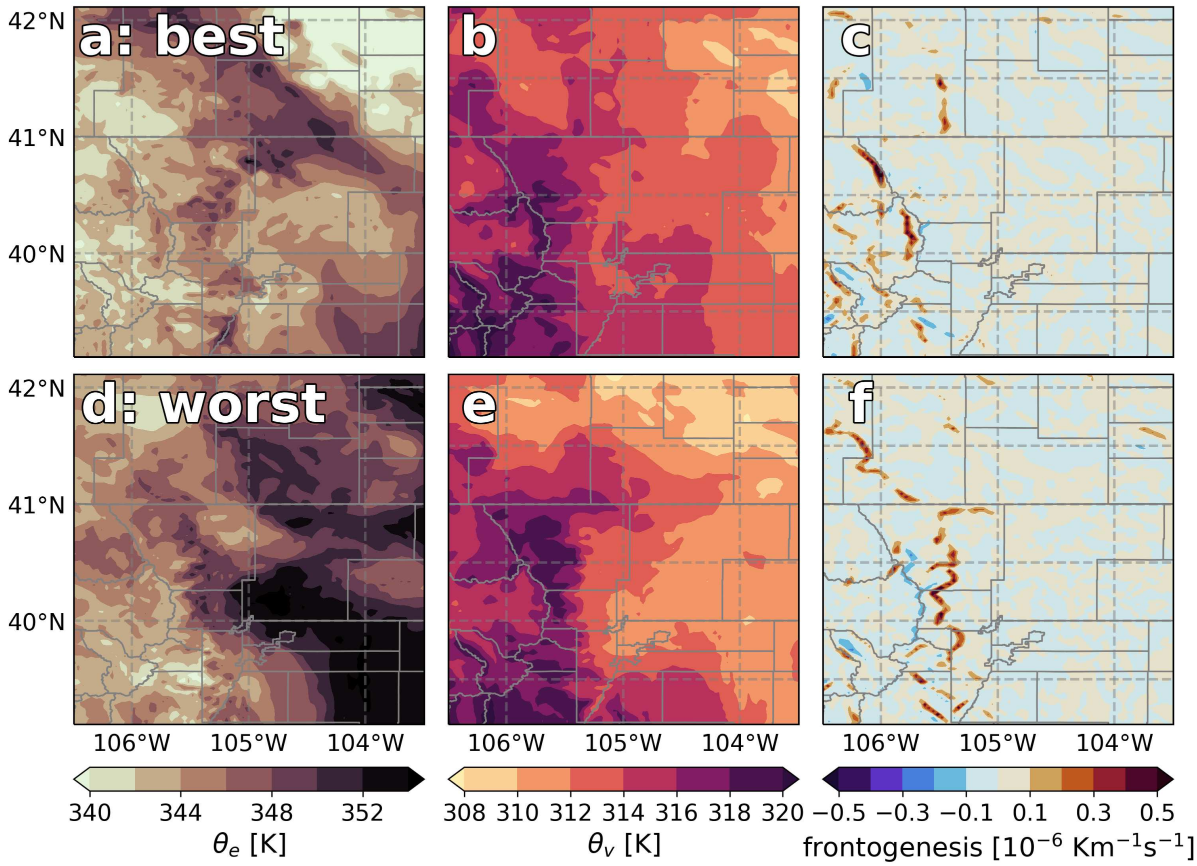


FIG. 3.7. **(a)** Map of surface $\theta_e$ on the 3 km domain for the best member at 15Z. **(b)** Map of surface $\theta_v$ on the 3 km domain for the best member. **(c)** Map of surface frontogenesis on the 3 km domain for the best member. **(d)** As in (a) for the worst member. **(e)** As in (b) for the worst member. **(f)** As in (c) for the worst member.

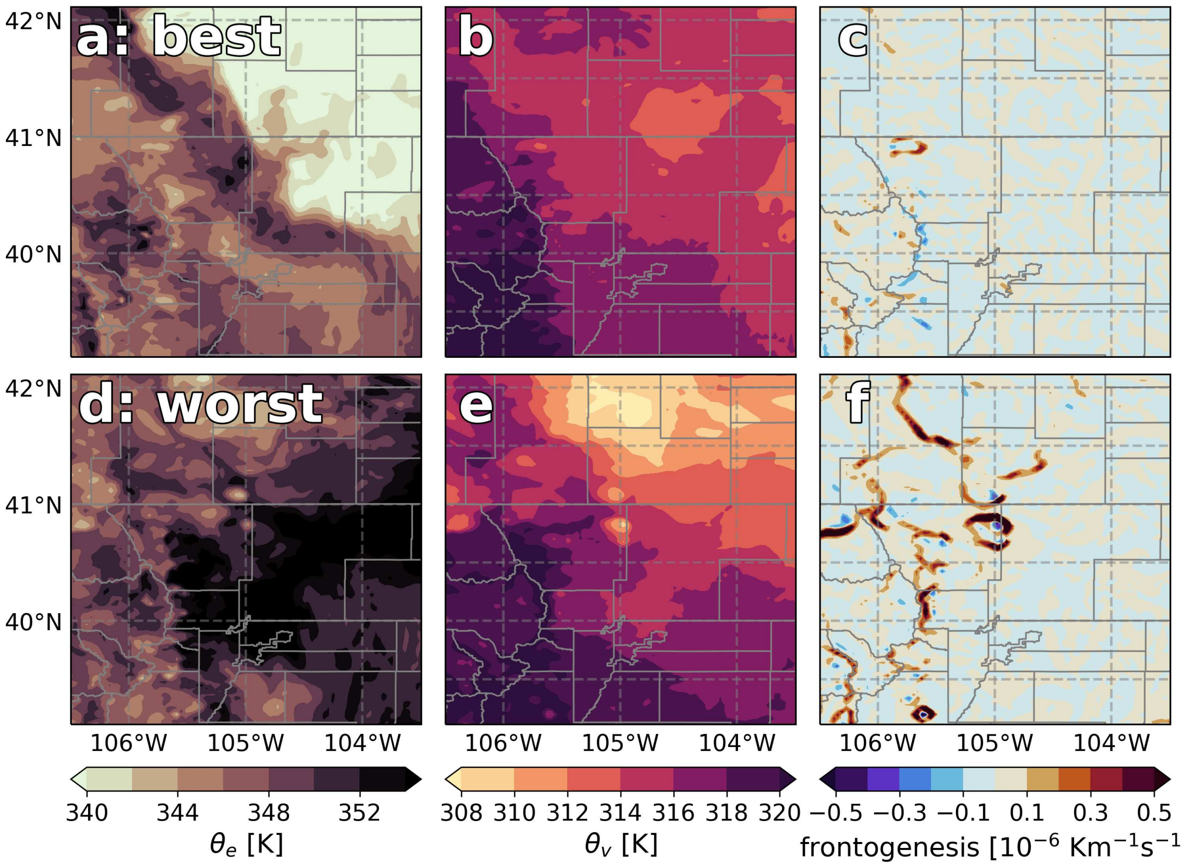


FIG. 3.8. **(a)** Map of surface $\theta_e$ on the 3 km domain for the best member at 18 Z. **(b)** Map of surface $\theta_v$ on the 3 km domain for the best member. **(c)** Map of surface frontogenesis on the 3 km domain for the best member. **(d)** As in (a) for the worst member. **(e)** As in (b) for the worst member. **(f)** As in (c) for the worst member.

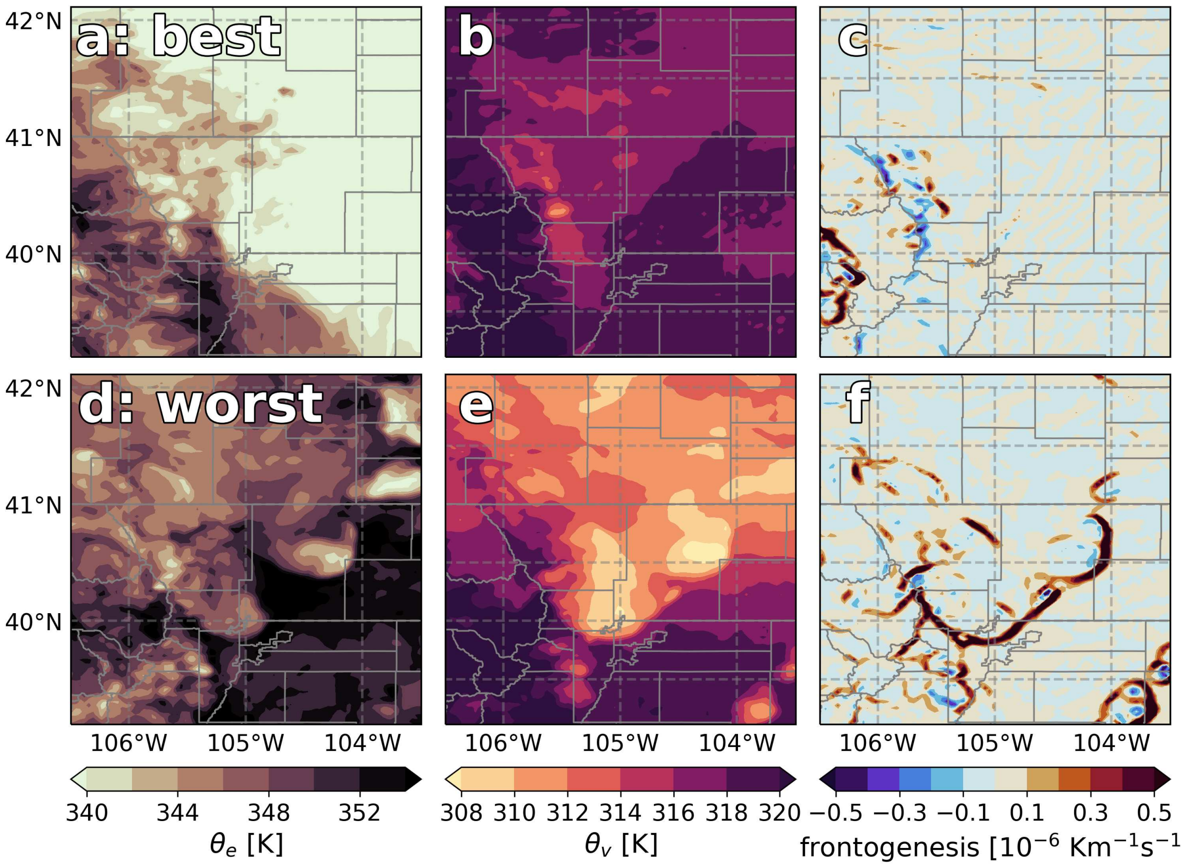


FIG. 3.9. **(a)** Map of surface $\theta_e$ on the 3 km domain for the best member at 21 Z. **(b)** Map of surface $\theta_v$ on the 3 km domain for the best member. **(c)** Map of surface frontogenesis on the 3 km domain for the best member. **(d)** As in (a) for the worst member. **(e)** As in (b) for the worst member. **(f)** As in (c) for the worst member.

Figs 3.7, 3.8, and 3.9 show the evolution of the surface $\theta_e$, $\theta_v$, and frontogenesis at 15, 18, and 21 Z, respectively. Note the difference in colorbar scales between these figures and their counterparts in Chapter 2 (Fig. 2.12) (in Chapter 2, we see higher $\theta_e$ values, lower $\theta_v$ values, and higher positive and negative frontogenesis values). Frontogenesis was calculated with $\theta_v$ using Equation 2.6 from **?** and neglecting the diabatic heating terms (both for simplicity and due to the lack of heating output by the model). Both $\theta_v$ and $\theta_e$ were calculated using the metpy package; $\theta_v$ is calculated according to **?** and $\theta_e$ is calculated according to **?** (Eqn. 2.3 and 2.5).

At 15 Z, both members feature similar profiles in the three variables, particularly $\theta_v$ (b and e) (Fig 3.7). In panels c and f, the worst member (bottom row) features frontogenesis occurring further east than that in the best member, though the north-south extent is similar between the two. The strongest differences are visible in panels a and d, which show an intrusion of cold air in the upper northeast corner of the best member, and a strong surface $\theta_e$ gradient near 40°, 105°W in the worst member. The members diverge rapidly as the simulation progresses to 18 and 21 Z (Fig 3.8 and 3.9). The cold air intrusion seen in $\theta_e$ in the best member at 15 Z moves further southeast at both 18 and 21 Z. The $\theta_v$ and frontogenesis fields at 18 Z reflect this intrusion in the best member, with a largely uniform $\theta_v$ field and little to no frontogenesis or frontolysis (negative frontogenesis). It is likely that this cold air intrusion is not associated with a frontal boundary, as there is not a strong gradient in $\theta_v$ which would represent a front. At 21 Z, the intrusion of cold air breaks up slightly near 105.5°W, 40.5°N, west of the region of interest, which is reflected by a small, isolated region of higher $\theta_v$ and stronger frontogenesis and frontolysis southwest of the region of interest. As the simulation progresses in the worst member, a more established front emerges in surface $\theta_v$ and moves southward toward 40°N by 21 Z. The $\theta_e$ field is less uniform, featuring isolated regions of low $\theta_e$ at 21 Z, which likely represent cold pools from existing convection. Frontogenesis strengthens from 15 to 18 Z, and again to 21 Z, illustrating convection moving further northeast at 18 Z and further southeast at 21 Z. In conjunction, these plots illustrate the intrusion of cold, dry air from the northeast over time and the subsequent lack of frontogenesis and deep, moist convection in the best member and the increase in widespread, strengthening convection associated with a surface $\theta_v$ front in the worst member, both of which are expected to have played large roles in the lack (or presence) of widespread heavy rainfall totals.

### 3.4 Conclusions and future work

This preliminary analysis began to investigate a mesoscale WRF ensemble simulation of an event in which heavy rainfall and potential flooding was forecast but little to no rainfall accumulated. We selected the best- and worst-performing ensemble members by comparing maximum rainfall totals in our region of interest against observational rainfall data and qualitatively selecting those with the most and least similar spatial patterns to reality. Soundings captured in-situ showed a moist environment with a low LFC and moderate values of CAPE sufficient to support deep, moist convection if air was lifted above the LFC. Our results suggest that a near-surface intrusion of cooler, drier air from the northeast prevented deep, moist convection from forming in the region of interest in the member which most accurately reproduced rainfall totals and patterns according to reality. This intrusion was not present in the least accurate member, which allowed a surface $\theta_v$ front to form, providing an uplift mechanism in an environment with high values of CAPE and a low LFC, triggering deep, moist convection in the region of interest and across the broader domain.

Based on observational data and the experience of the field team launching balloons on 31 July 2021, there was a strong cloud deck and little solar insolation over the course of the day. As a result, there was less surface heating, which likely prevented strong surface uplift, preventing air from rising above the low LFC to form deep, moist convection. There were anomalously high precipitable water values in the region, as shown by the soundings which originated at Christman Field, which would have provided sufficient moisture for heavy rainfall, as forecast. Further analysis of this case is needed to confirm these observations and hypothesis, such as evaluation of temperature and solar insolation over the course of the day. Future work should continue to examine the mesoscale meteorological variables that contributed to the differences in rainfall totals and patterns between members. As previously discussed, regridding the MRMS data to match the native WRF grid in order to perform point calculations would provide valuable statistical insight. In this way, a probabilistic approach could be taken to assess ensemble performance and find areas in which to propose changes for model improvement. Future work would also benefit from including a comprehensive examination of the radar data gathered by the PRE-CIP 2021 team in order to track storm systems and cells as the day unfolded. Further investigation of this case could elucidate mesoscale features that can promote or inhibit deep, moist convection from forming in a conducive environment with anomalously high moisture. This case provides strong potential to improve model accuracy and short-range forecasts of heavy rainfall events.

CHAPTER 4

# CONCLUSIONS AND FUTURE WORK

As extreme precipitation events grow in frequency and magnitude as environments get warmer and moister, predicting heavy rainfall will become increasingly more important (**??**). Forecasting these extreme events will also become increasingly important for mountainous regions with high potential for flash flooding, such as in the Big Johnson River Flood of 1973. As computational power has grown and modeling accuracy has increased, due to advancements in parameterizations and other efforts, numerical weather prediction has become an increasingly powerful tool in researching and forecasting weather phenomena. One particularly valuable advancement that has become more widely used in recent years has been ensemble forecasts, which provide a broader set of potential realities than a deterministic forecast. The increased usage of ensembles in research can improve parameterizations and make them more widely available to forecasters and reasonable for forecasters to use.

In order to further understanding of the meteorological processes behind extreme rainfall events via model ensembles, we analyzed two events which took place during separate components of the PRECIP field campaign: a heavy rainfall case on 09 June 2020 near Taiwan and Yonaguni, Japan, in which global models forecast little rain but widespread heavy rainfall occurred, with peak accumulation greater than 600 mm and widespread accumulation greater than 100 mm; and a case on 31 July 2021 in Northern Colorado in which models forecast heavy rainfall in our region of interest, near and surrounding Fort Collins, CO, but little to no rainfall actually occurred. These cases provide a valuable contrast and allow us to explore the mechanisms that contribute to heavy rainfall, or the lack thereof, in two largely different environments–one warm, tropical environment with consistently high moisture, and one land-locked, continental environment with significantly lower moisture.

To investigate the 09 June 2020 case, we analyzed the PSU WRF-EnKF ensemble made available via the PRECIP 2020 dry run portion of the PRECIP 2020 campaign. We selected the most and least accurate members by statistical comparison against observational rainfall data for a full meteorological analysis. We found that the location and intensity of the shallow near-surface frontal boundary were key to ensemble member accuracy. In the two best members, south-southwesterly flow ascended over the sloped isentropes, rose above the LFC, releasing conditional instability, and encountered consistently high moisture throughout the column to produce back-building deep convection. In the two worst members, the near-surface gradients were weaker and more confined along Taiwan's coast, which led to less rainfall displaced west of the observed rainfall. We analyzed the ensemble mean at the first time step of the following model run (12Z 09 June 2020), which occurred after data assimilation had completed and which we used as our verification time step. Analysis of the mean revealed the greater importance of the location and sharpness of the near-surface front as compared to the low-tropospheric moisture above the boundary layer. The member which most accurately reproduced the rainfall more accurately captured the location of the front as compared to the mean, while the member which least accurately reproduced the rainfall more accurately captured the location of the low-tropospheric, near 850 to 700 mb, dry air as compared to the mean.

In our preliminary analysis of the 31 July 2021 case, we again analyzed the PSU WRF-EnKF ensemble, as well as a number of observational data sources, including in-situ soundings captured by the PRECIP field team and MRMS observational rainfall estimates. We selected the most and least accurate ensemble members according to their maximum rainfall accumulation within the region of interest and qualitative rainfall patterns in the broader domain as compared to the MRMS rainfall totals. In the best member, a near-surface intrusion of drier air prevented deep, moist convection from forming in the region of interest. A lack of an intrusion in the least accurate member allowed the environment to stay moist enough and, when coupled with a low LFC and high values of CAPE, allowed deep, moist convection to develop in the region of interest and across the broader domain. The deep, moist convection generated gust fronts, which contributed to the $\theta_v$ boundaries seen in Figs. 3.7, 3.8, and 3.9. We hypothesize, based on observations from the PRECIP field team, of which the author was a part that day, that a strong cloud deck and little solar insolation over the course of the day restricted surface heating and, thus, buoyancy buoyancy, preventing air from rising above the LFC and forming deep, moist convection. Further analysis is required to confirm the observations and should include, but not be limited to analysis of temperature and solar insolation, as well as other meteorological variables; further statistical analysis and probabilistic analysis of the ensemble; and comprehensive analysis of the radar data gathered during the case by the PRECIP radar team.

In conjunction, our studies suggest that increased accuracy in simulating the location of frontal boundaries, specifically those which do not appear to have strong synoptic forcing, and meso- to synoptic scale winds may lead to improved accuracy in models and emphasize the value of ensemble forecasts and efforts that lead to their improvement. Ensembles provide the unique ability to investigate both meteorological features which were influential in accurate members and detrimental in inaccurate members. Future research should continue to evaluate ensembles in extreme rainfall cases–those in which heavy rainfall is not forecast but occurs, and vice versa–and further the probabilistic approach to understand how ensembles can be useful for forecasters and continue to understand the drivers behind heavy rainfall events, especially those on smaller scales.