

Article

An Ensemble-Based Analysis of a Liminal Extreme Rainfall Event near Taiwan

Alexandra S. Cole , Michael M. Bell  and Jennifer C. DeHart * 

Department of Atmospheric Science, Colorado State University, 1371 Campus Delivery, Fort Collins, CO 80523-1371, USA; ascole@colostate.edu (A.S.C.); mmbell@rams.colostate.edu (M.M.B.)

* Correspondence: jcdehart@colostate.edu

Abstract: This study analyzes an ensemble of numerical simulations of a heavy rainfall event east of Taiwan on 9 June 2020. Heavy rainfall was produced by quasi-stationary back-building mesoscale convective systems (MCS) associated with a mei-yu front. Global model forecast skill was poor in 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 moist soundings with low levels of free convection (LFC) and sufficient instability for deep convection. We find that stronger gradients in 100-m θ_e and θ_v in the most accurate members associated with a near-surface frontal boundary focus the lifting mechanism for deep, moist convection and enhanced rainfall. As the simulations progress, stronger southerly winds in the least accurate members advect drier mid-level air into the region of interest and shift the near-surface boundary further north and west. Analysis of the verification ensemble mean analysis reveals a strong near-surface frontal boundary similarly positioned as in the most accurate members and dry air aloft more similar to that in the least accurate members, suggesting that the positioning of the frontal boundary is more critical to accurately reproducing rainfall patterns and intensity in this case. The analyses suggest that subtle details in the simulation of frontal boundaries and mesoscale flow structures can lead to bifurcations in producing extreme or almost no rainfall. Implications for improved probabilistic forecasts of heavy rainfall events will be discussed.

Keywords: mei-yu; extreme rainfall; modeling; mesoscale



Citation: Cole, A.S.; Bell, M.M.; DeHart, J.C. An Ensemble-Based Analysis of a Liminal Extreme Rainfall Event near Taiwan. *Atmosphere* **2022**, *13*, 1011. <https://doi.org/10.3390/atmos13071011>

Academic Editors: Jing-Shan Hong and Ling-Feng Hsiao

Received: 25 May 2022

Accepted: 16 June 2022

Published: 23 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

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 [1]. 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 [2–4]. 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 [4–6].

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) [6,7]. These mesoscale mei-yu fronts and embedded MCSs positively interact and intensify each other [8]. 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 [6,7].

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 [1,3,9,10]. The extreme rainfall that the mei-yu and mei-yu fronts can produce still provides large forecasting challenges.

The mei-yu season in 2020 is recognized as an anomalously rainy season, the reasons for which have been investigated in recent literature. Wang et al. [11] 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. Volonté et al. [12] discovered that the halted northward progression of the Eastern Asian Summer Monsoon Front was associated with anomalously high precipitation in the Yangtze River Basin. The anomalous 2020 season was also found to be impacted by persistent Madden–Julian oscillation phases 1–2 [13]. Mei-yu seasons such as that in 2020 may become more frequent in the future [14], and Liang suggests that there are quantitative challenges in predicting extreme mei-yu seasons such as 2020 [13], 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.

One approach to heavy rainfall forecasting is through the use of Doswell et al. [15]'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:

$$P = \bar{R}D. \quad (1)$$

Equation (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 [15]. Despite the simplicity of Equation (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

$$R = Ewq, \quad (2)$$

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 [15]. We use Doswell et al. [15]'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 [16,17]. 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 [18]. 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 (Holloway and Neelin [19]), and statistically provides a critical threshold for more frequent deep convection [20]. Molinari et al. [21] 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.

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 [22]. 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 [23–25]. 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 [26]. 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 [27].

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 [23,28,29]. 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 [30].

The need for more frequent usage of ensemble forecasts has become more obvious in recent years, as ensembles outperform deterministic forecasts [31,32]. 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.

The heavy rainfall case on 9 June 2020 was selected for this study due to the existence of an apparent bifurcation in the mesoscale ensemble solutions being run in real-time associated with the PRECIP 2020 field campaign dry run. In reality, the event produced over 600 mm east of Taiwan, but some ensemble members produced very little rain while

others produced high totals. Based on the ensemble analysis, we assessed this event as ‘liminal’, defined by the Oxford Dictionary as ‘occupying a position at, or on both sides of, a boundary or threshold’. These ensembles allowed us to study the subtle mesoscale differences associated with the extreme rainfall event to assess the most important factors that determined whether heavy rainfall would occur or not. In this study, we use the WRF ensemble output verifying analysis to examine the ensemble spread in results as well as the meteorological drivers behind the event. We aim to understand why this event was poorly forecast by global models and help forecasters and modelers alike in predicting and modeling events such as this one in the future. The structure of this paper will be as follows. Section 2 will explain the data and model setup and the selection of ensemble members for validation. Section 3 will give a brief summary of the case being studied, analyze the mesoscale features associated with the case in two ensemble members, and provide comparison with more ensemble members. Section 4 will present conclusions.

2. Model Setup and Validation

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 [33]. 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 [34–36]. 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 300×300 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. 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 [37]. 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 Section 2.3, 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.

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:

$$ETS = \frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{misses} + \text{false alarms} - \text{hits}_{\text{random}}}, \quad (3)$$

where hits, misses, and false alarms are defined in the contingency table shown in Table 1 (correct negatives will not be used in this study). Hits due to random chance are calculated via the following equation:

$$\text{hits}_{\text{random}} = \frac{(\text{hits} + \text{misses}) * (\text{hits} + \text{false alarms})}{\text{total}}. \quad (4)$$

Table 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

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.

$$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} \quad (5)$$

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \quad (6)$$

On an 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.

As a first pass, we calculated ETS for all 40 members at a 24-h lead time within the region of interest 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 h) so that we have a different number of chunks for each duration; e.g., we have 12 chunks with 2-h duration and 3 chunks with 8-h 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-h chunk using a 1 mm/h average rain rate, the threshold would be 2 mm, and for a 4-h chunk using a 1 mm/h 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 (3) to achieve an ETS score for each duration. We calculate six ETSs for each member (one for each duration), shown in Figure 1.

From Figure 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/h of rainfall for smoothness), shown in Figure 2. From Figure 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.

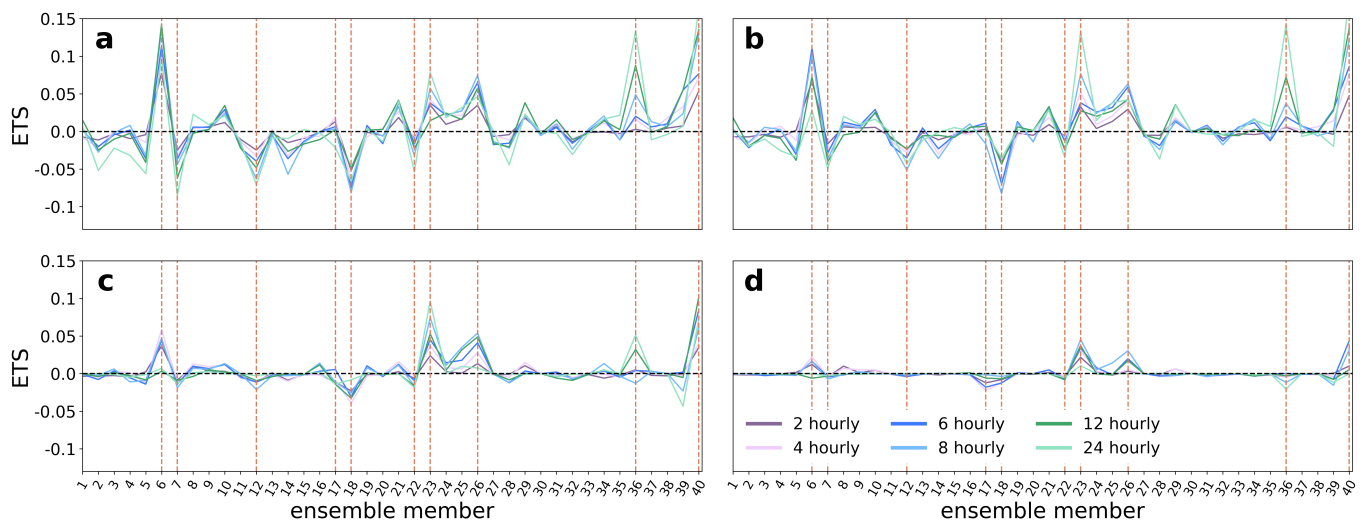


Figure 1. (a) Equitable threat scores for a rainfall rate threshold of 1 mm/h. Hourly periods are differentiated by different colors, which are defined in the legend in panel (d). The five best performing and five worst performing ensemble members are highlighted by vertical dashed orange lines. (b)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

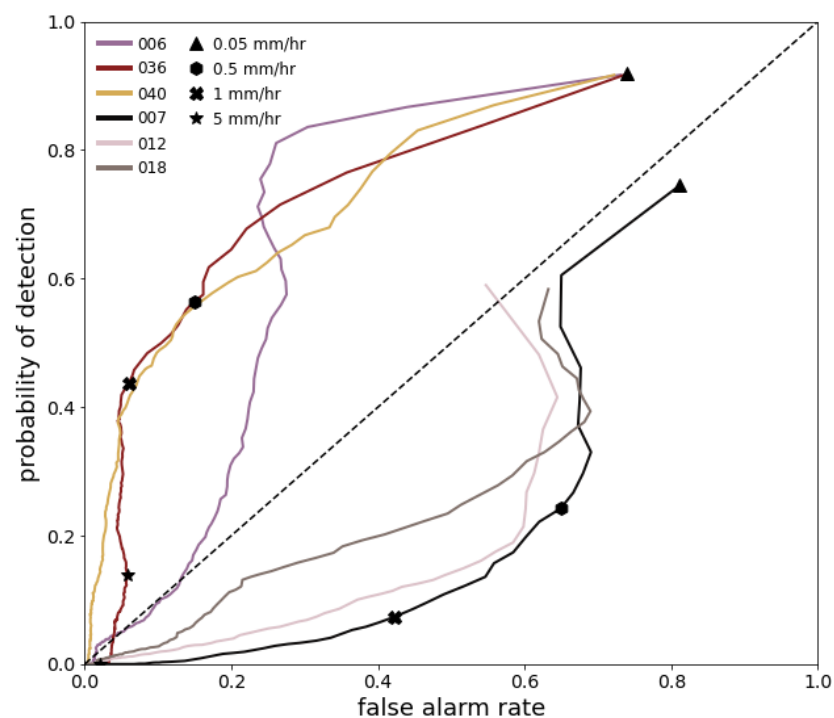


Figure 2. Relative operating characteristic curves for the three best and three worst performing ensemble members. Different rainfall rate thresholds are marked on the single best and worst members only.

3. Analysis

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 3a shows the observed rainfall estimated by the QPESUMS dataset over the 24-h period starting at 00 UTC on 9 June 2020 [37]. QPESUMS is described in Section 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 13 UTC (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.

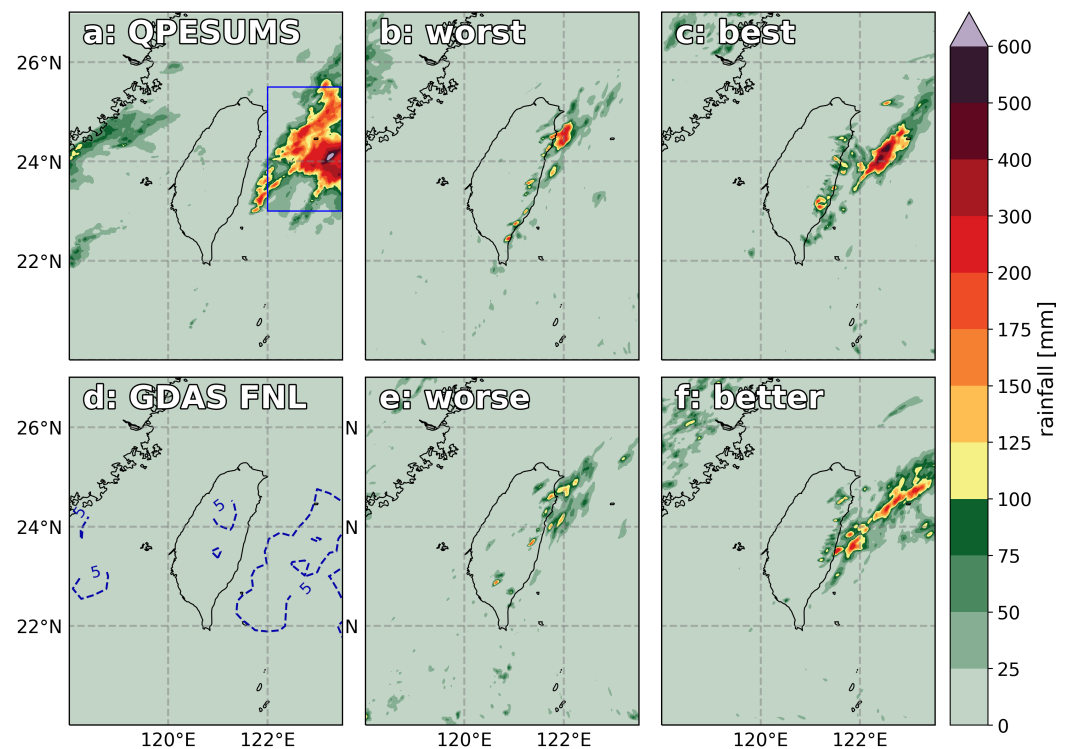


Figure 3. (a) Map of rainfall totals for QPESUMS for the 24-h period beginning at 00 UTC 9 June 2020. (b) As in (a) for the worst WRF ensemble member. (c) As in (a) for the best WRF ensemble member. (d) As in (a) for GDAS FNL. The blue dashed contour lines outline areas that received greater than 5 mm 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 4 shows the brightness temperature from NASA GES DISC satellite imagery. The mei-yu front is along the periphery of the 850 hPa high shown in Figure 5 (described in further detail in Section 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 4, and the wind field is shown in Figure 5.

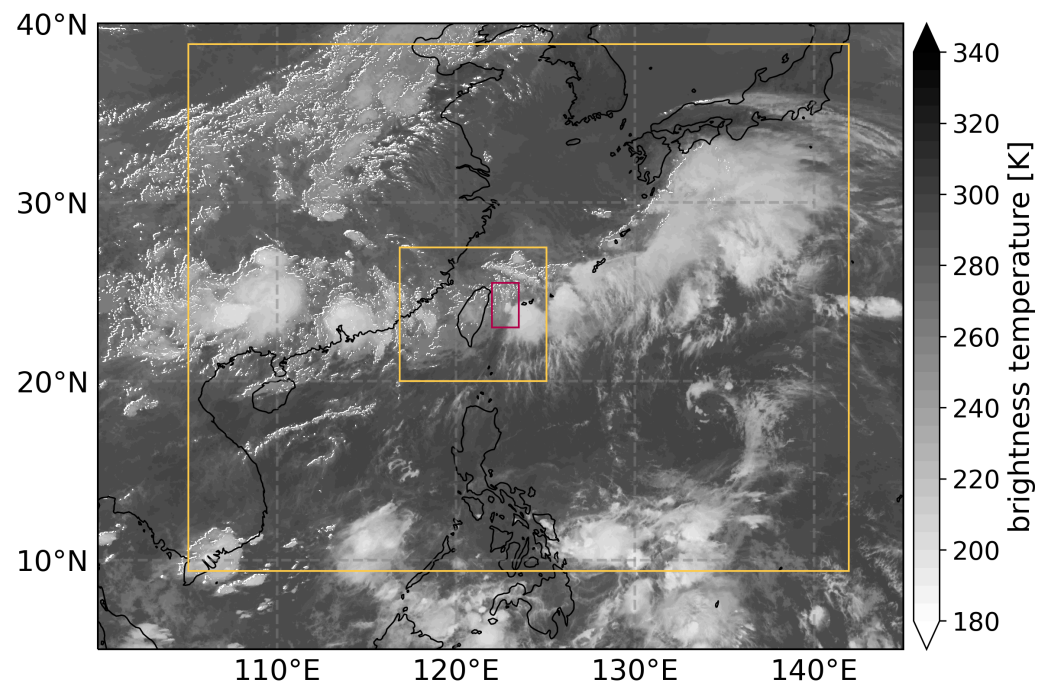


Figure 4. Map of NASA GES DISC brightness temperature at 00 UTC 9 June 2020. The yellow boxes represent the 9-km (**outer**) and 3-km (**inner**) WRF domains used in the PSU WRF-EnKF simulations. The red box shows the region for ensemble statistical calculations and soundings, also called the “region of interest”.

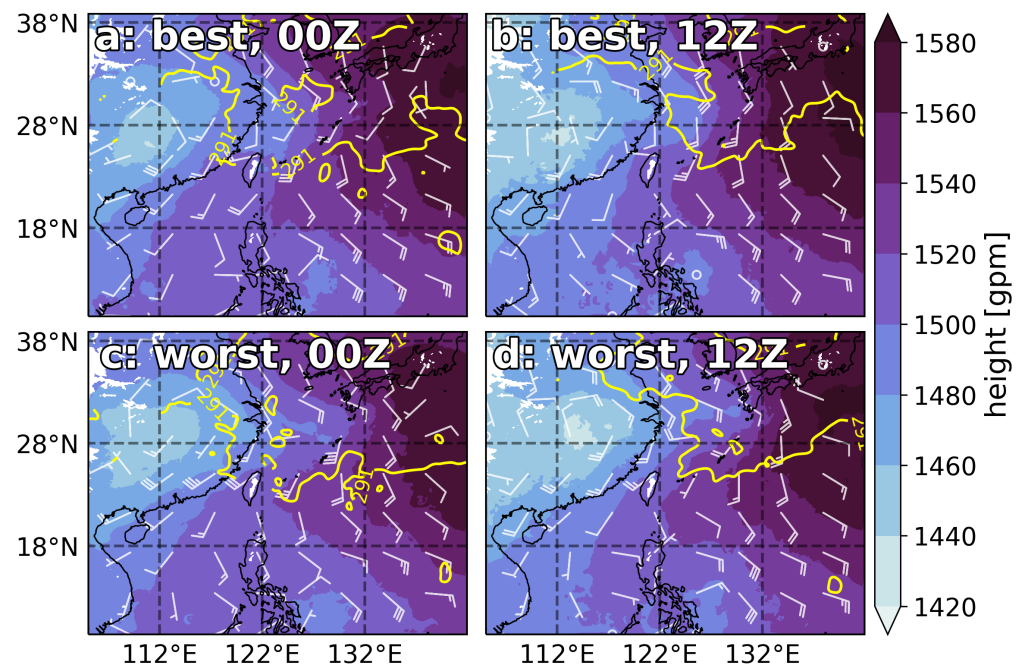


Figure 5. (a) Map of 850 hPa heights (shading), zonal/meridional wind vectors in units of knots (white), and the 291 K temperature contour (yellow) on the 9-km domain for the best member at 00 UTC 9 June 2020. The temperature has been filtered with a 2-dimensional Gaussian filter with a sigma of 3, using the SciPy package [38]. (b) As in (a) at 12 UTC. (c) As in (a) for the worst member at 00 UTC. (d) As in (a) for the worst member at 12 UTC.

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 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 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.

3.2. Synoptic-Scale Analysis

Figure 4 shows the synoptic-scale cloud patterns from satellite brightness temperature at 00 UTC on 9 June 2020. Figure 5 shows the 850 hPa geopotential heights, winds, and 291 K temperature contour for the best (top) and worst (bottom) members at 00 UTC (left) and 12 UTC (right) for the 9-km domain in the WRF ensemble. The synoptic features visible in (Figure 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 00 UTC to 12 UTC, the 850 hPa low over China strengthens slightly. The differences between the best member (top) and worst member (bottom) in Figure 5 are small, but these small differences in the synoptic setup are enough to motivate differences in the mesoscale environment that substantially 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 291 K 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 zonal direction, where the worst member has slightly more westerly winds than the best member.

We calculated a representative sounding using a box average over the area of interest (outlined by the red box in Figure 4) at 12 UTC (Figure 6). We chose 12 UTC 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.

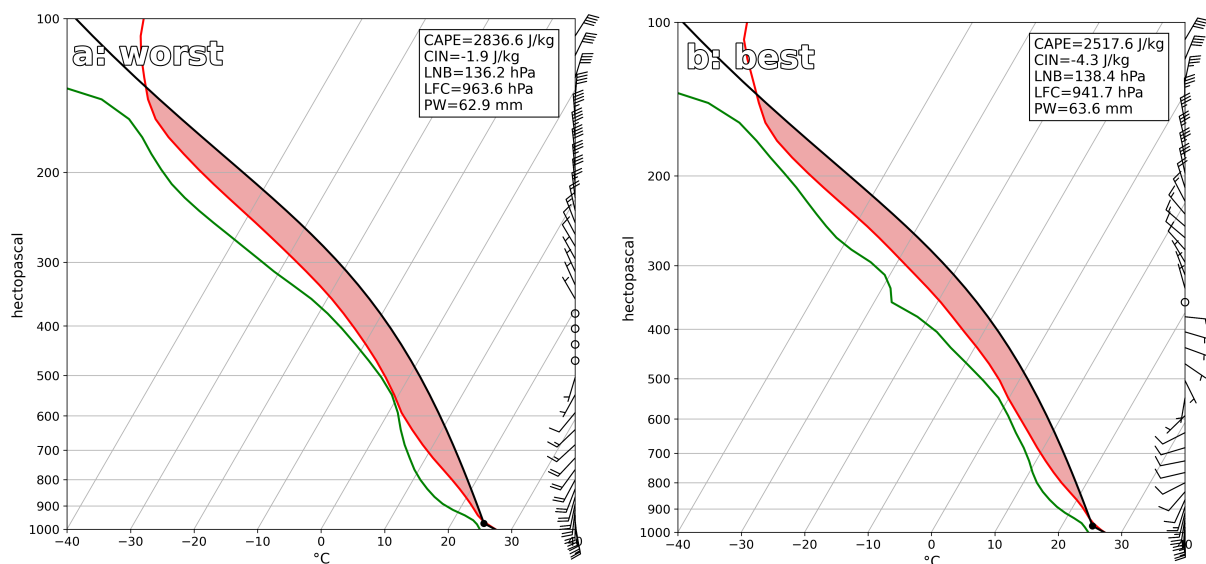


Figure 6. (a) Box-average sounding calculated over the red box in Figure 4 for the worst member at 12 UTC. The green line represents dew point temperature, the red line represents temperature, and the black line traces the sounding. The lifted condensation level (LCL) is marked by a black dot. Convection available potential energy (CAPE, shaded in red), 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.

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, which indicate thermodynamic environments that 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 700 hPa. The worst member's sounding features a drier profile than the best member in the lower atmosphere below 600 hPa. The difference in low-level

moisture also shows up in relative humidity (RH) and column-integrated water vapor (CWV) analysis to be described later. (Figures 7–10).

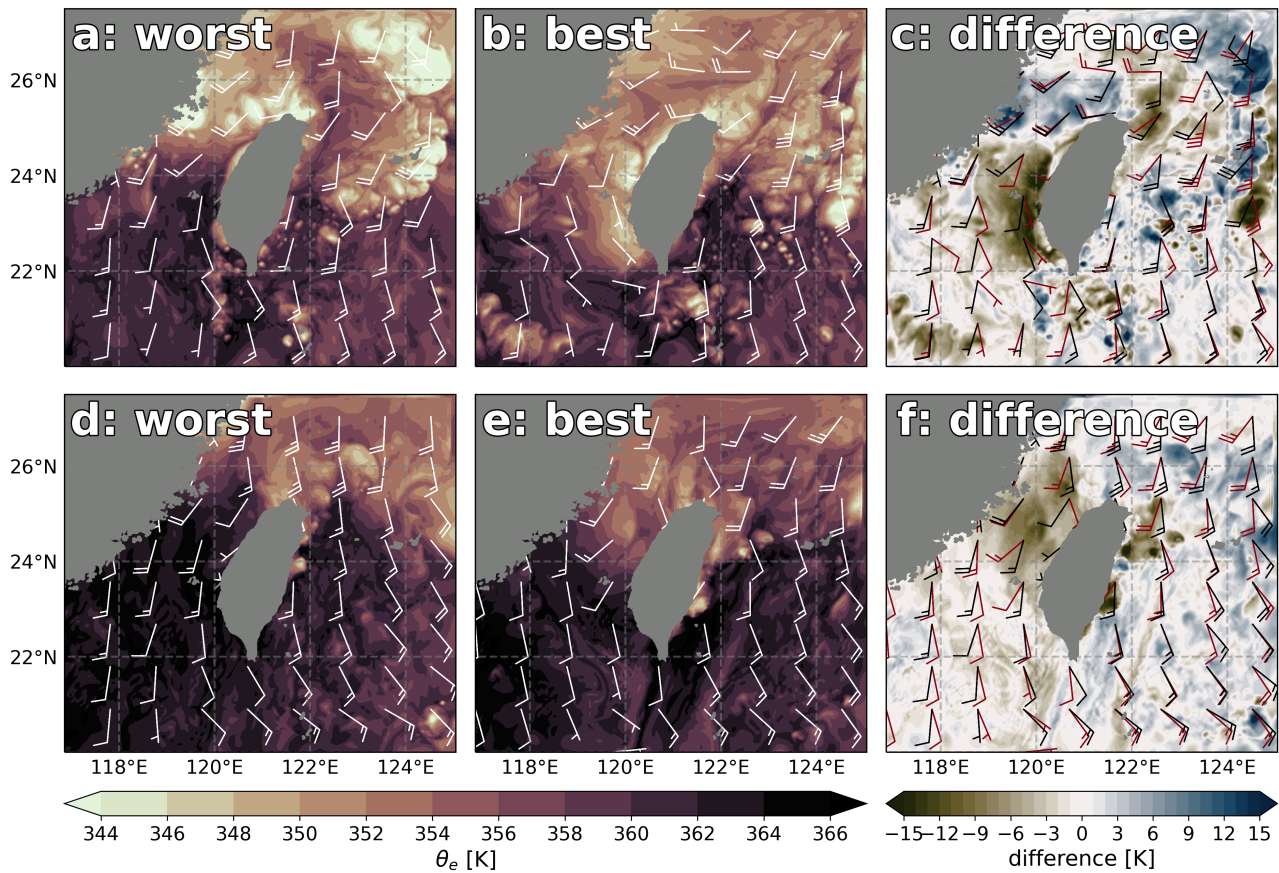


Figure 7. (a) Map of 100-m θ_e (shading) and zonal/meridional winds in units of knots (white) on the 3-km domain for the worst member at 00 UTC. (b) As in (a) for the best member. (c) Map of the difference between panel (a,b) (worst subtracted from best) and zonal/meridional winds (worst in black barbs, best in red barbs). (d) As in (a) for the worst member at 12 UTC. (e) As in (a) for the best member at 12 UTC. (f) As in (c) at 12 UTC.

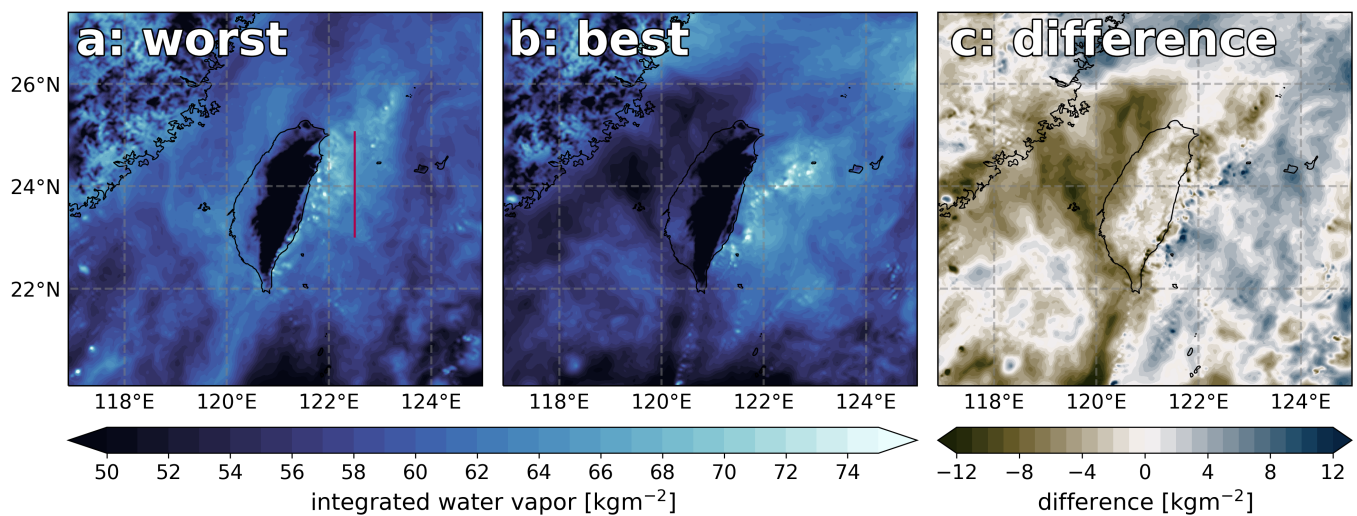


Figure 8. (a) Map of the column-integrated water vapor for the worst member at 12 UTC on the 3 km domain. The red line shows the line across which vertical cross-sections are taken for Figure 12. (b) As in (a) for the best member. (c) Map of the difference between (a,b) (worst subtracted from best).

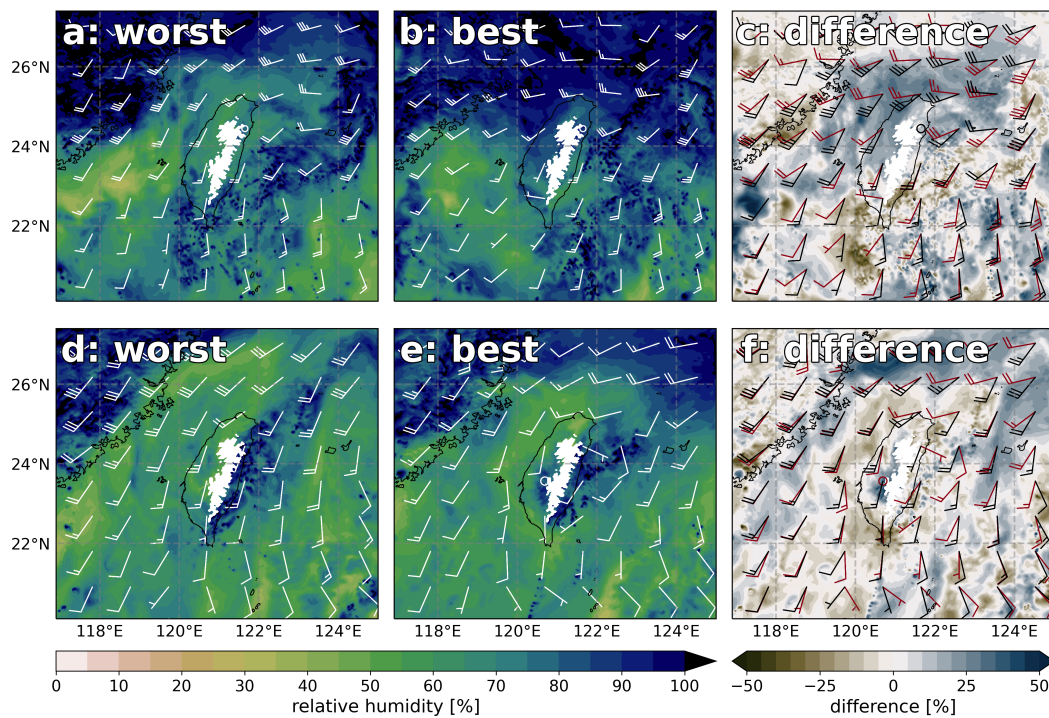


Figure 9. (a) Map of the 850 hPa relative humidity (shading) and zonal/meridional winds in units of knots (white) for the worst member at 00 UTC on the 3-km domain. (b) As in (a) for the best member. (c) Map of the difference between panel (a,b) (worst subtracted from best) and zonal/meridional winds in units of knots (worst in black barbs, best in red barbs). (d) As in (a) at 12 UTC. (e) As in (b) at 12 UTC. (f) As in (c) at 12 UTC.

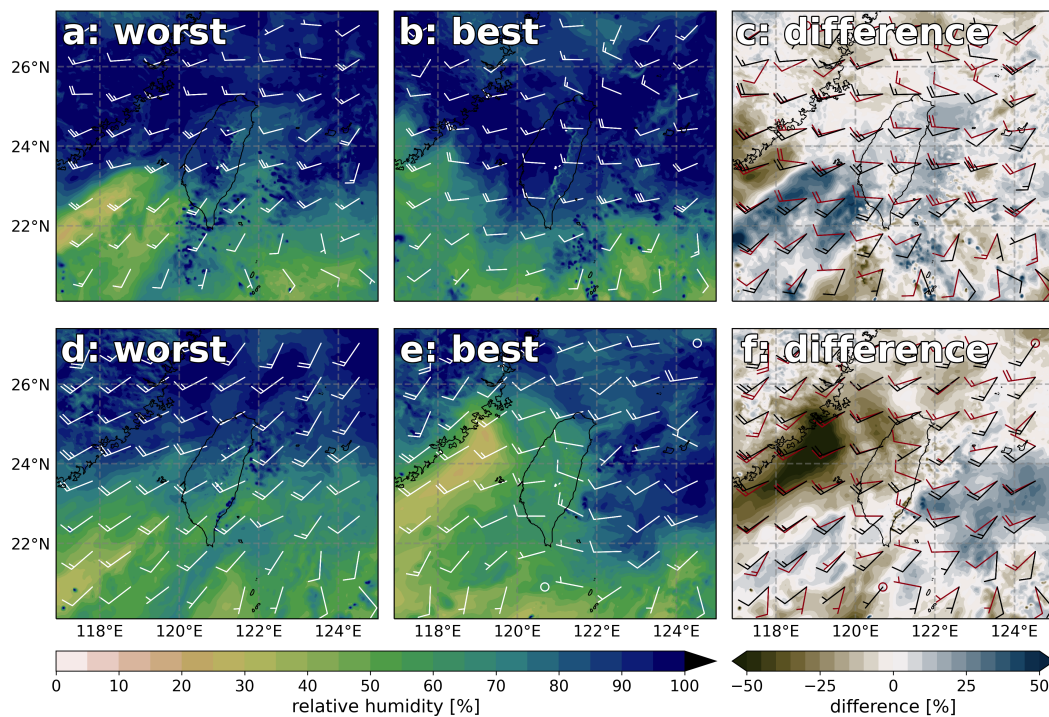


Figure 10. (a) Map of the 700hPa relative humidity (shading) and zonal/meridional winds in units of knots (white) for the worst member at 00 UTC on the 3-km domain. (b) As in (a) for the best member. (c) Map of the difference between panel (a,b) (worst subtracted from best) and zonal/meridional winds in units of knots (worst in black barbs, best in red barbs). (d) As in (a) at 12 UTC. (e) As in (b) at 12 UTC. (f) As in (c) at 12 UTC.

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.

3.3. Mesoscale Analysis

Equivalent potential temperature (θ_e) acts as a valuable measure of temperature and moisture, as well as unstable, buoyant air. θ_e is calculated using Equation (7) according to Bolton [39], using the metpy package [40].

$$\theta_e = \theta_{DL} \exp \left[\left(\frac{3036}{T_L} - 1.78 \right) * r(1 + 0.448r) \right], \quad (7)$$

where θ_{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 θ_e variability dominates so we will focus on θ_e as a measure for moisture going forward. We choose 100 m to show plots of θ_e (and later θ_v and frontogenesis) as it is a representative layer of air that is typically lifted in buoyant plumes of convection [41]. Gradients in θ_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 (θ_v) for that purpose later on) [42].

There is ample moisture in the area of interest for lifting mechanisms to trigger deep, moist convection, particularly when interacting with a low LFC as shown in the soundings in Figure 6. At 00 UTC, there is an intrusion of warm, moist air near 122 E, 24 N in the best member, which creates a strong θ_e gradient (on the order of 15 K) in the area of interest between the intrusion and the colder, drier air mass just northeast of Taiwan (Figure 7). In the worst member at 00 UTC, the θ_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 θ_e gradient in the area of interest. By 12 UTC, 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 θ_e fields at 00 UTC and 12 UTC. The cooler air in panel f shows that the θ_e gradient in the best member is more prominent than in the worst member at 12 UTC, 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 θ_e gradient or whether the θ_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.

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:

$$IWV = -\frac{1}{g} \int_{p_{bottom}}^{p_{top}} r dp \quad (8)$$

At 12 UTC, the worst member and the best member have similar maximum values of integrated water vapor, but these maximums are displaced (Figure 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 Figure 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 60 s and low 70 s kg m^{-2} indicate large quantities of moisture to begin with but differences on the order of $8\text{--}10 \text{ kg m}^{-2}$ 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.

At 850 hPa, the winds in the worst member remain organized and southerly from 00 UTC (panel a) to 12 UTC (panel d), whereas in the best member, the winds weaken and become less organized over that period (Figure 9). The worst member has greater relative humidity values than the best member in the area of interest at 00 UTC, 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 700 hPa 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 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 (Figure 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 hPa upwards in the atmosphere. This drier air higher in the atmosphere, especially at 00 UTC, 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.

Figure 11 shows the θ_e , θ_v , and frontogenesis parameter for both members at 12 UTC. We calculated θ_e as previously described (Equation (5)). We calculated θ_v using Equation (9) [43] via the metpy package, where θ is the potential temperature, ε is the ratio of water vapor to dry air, and r_v is the water vapor mixing ratio. We calculated frontogenesis (F) with θ_v using Equation 2.3.21 from Bluestein [44] and neglecting the diabatic heating terms (both for simplicity and due to the lack of heating output by the model; Equation (10)), where u , v , and w are the velocities in the x , y , and z directions, respectively.

$$\theta_v = \theta \frac{1 + r_v/\varepsilon}{1 + r_v} \approx \theta(1 + 0.61r_v) \quad (9)$$

$$F = \frac{1}{|\nabla\theta_v|} \left\{ \left(\frac{\partial\theta_v}{\partial x} \right) \left\{ - \left(\frac{\partial u}{\partial x} \frac{\partial\theta_v}{\partial x} \right) - \left(\frac{\partial v}{\partial x} \frac{\partial\theta_v}{\partial y} \right) - \left(\frac{\partial w}{\partial x} \frac{\partial\theta_v}{\partial z} \right) \right\} + \frac{\partial\theta}{\partial y} \left\{ - \left(\frac{\partial u}{\partial y} \frac{\partial\theta_v}{\partial x} \right) - \left(\frac{\partial v}{\partial y} \frac{\partial\theta_v}{\partial y} \right) - \left(\frac{\partial w}{\partial y} \frac{\partial\theta_v}{\partial z} \right) \right\} + \frac{\partial\theta}{\partial z} \left\{ - \left(\frac{\partial u}{\partial z} \frac{\partial\theta_v}{\partial x} \right) - \left(\frac{\partial v}{\partial z} \frac{\partial\theta_v}{\partial y} \right) - \left(\frac{\partial w}{\partial z} \frac{\partial\theta_v}{\partial z} \right) \right\} \right\} \quad (10)$$

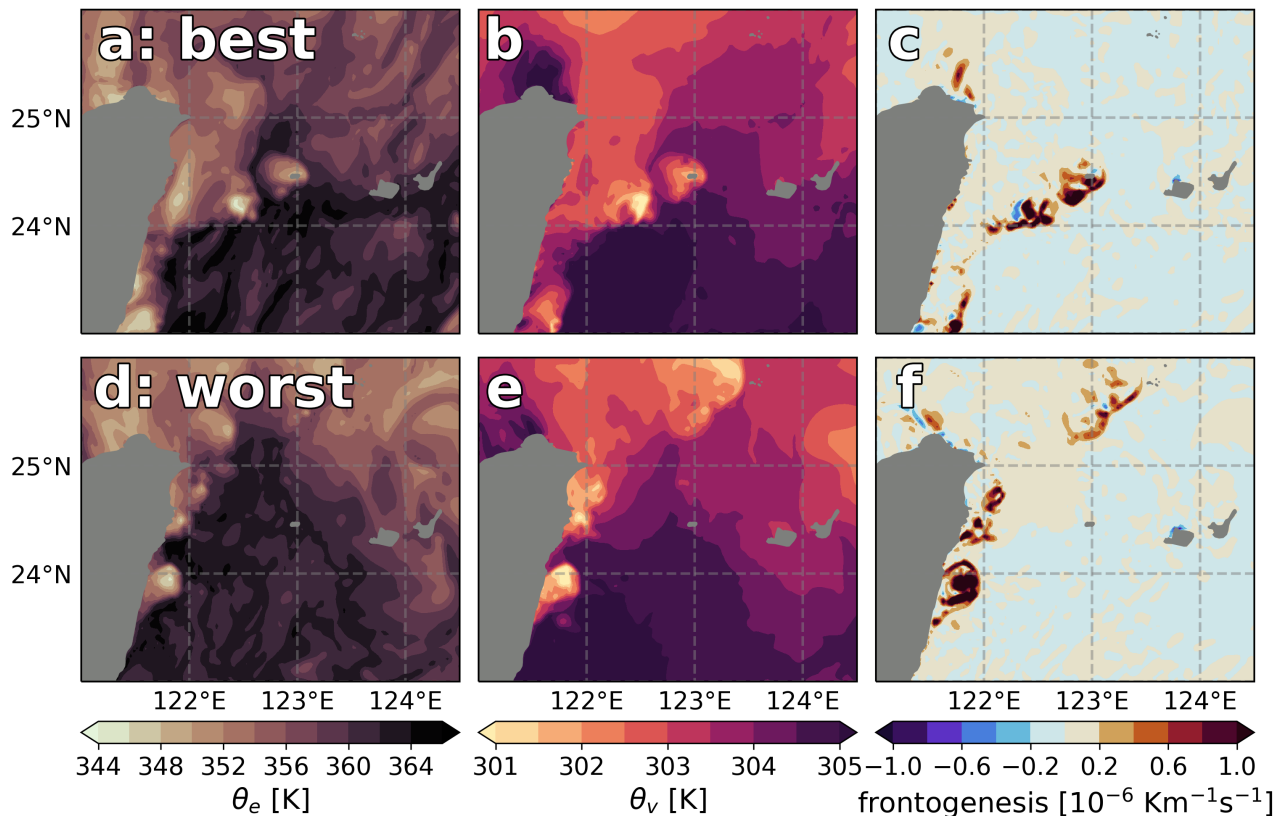


Figure 11. (a) Map of 100-m θ_e for the best member at 12 UTC on the 3-km domain. (b) As in (a) for θ_v . (c) As in (a) for the frontogenesis parameter. (d) As in (a) for the worst member. (e) As in (a) for the frontogenesis parameter in the worst member. (f) (left), θ_v (middle), and the frontogenesis parameter (right) for the best member (top row) and the worst member (bottom row) on the 3-km domain at 12 UTC.

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 (θ_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 θ_e and θ_v , the gradients are important components in Figure 11. Panels (a) and (b) and (d) and (e) highlight the similarities between the θ_e and θ_v gradients, with the primary difference being that the differences across the θ_e gradients are on the order of 15–20 K while the differences across the gradients associated with θ_v fronts are closer to 4–5 K, indicating that the moisture aspect of the gradients is stronger than the temperature gradient. Together, these two variables 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.

In both θ_e and θ_v , we see the differences in both location and strength of the frontal boundary. The frontogenesis parameter shows that frontogenesis is indeed taking place along these gradients. In the best member, both variables highlight the southeastward extension of the boundary, as far east as Yonaguni at 123 E and as far south as 24 N.

The frontal boundary in the worst member is more confined to the coast of Taiwan, not extending much past 122°E and only dropping south of 24°N along the coast. This difference in frontal boundary location, combined with the aforementioned moisture differences, strongly contributes to the difference in intensity and location of deep, moist convection between the two members. The frontal forcing evident by the frontal boundaries and frontogenesis parameter is necessary to release the conditional instability shown in the soundings. 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 Figure 11 may be highlighting the existence of convection where stronger positive values are seen. However, whether the frontogenesis is occurring primarily as a result of dynamic processes or as a result of the convection does not negate the conclusion that the positioning of the θ_e gradients and θ_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 affirms that the θ_e gradients and θ_v fronts are providing uplift mechanisms contributing to that convection due to the co-location between the variables.

Figure 12 shows a cross-section taken along the red line in Figure 8a with the relative humidity shown in shading, θ_v isentropes shown in white dashed contours, and wind shown in the red wind barbs in the lower 5 km of the atmosphere. θ_v is filtered using a 2-dimensional Gaussian filter with $\sigma = 1$, using the SciPy package [38]. 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 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 θ_e gradient in Figure 8, in Figure 12 we see a θ_v front in a similar location. The shift of the θ_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 θ_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 θ_v isentropes acts as the primary lifting mechanism. However, in the worst member the tightening and shifting of the θ_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 θ_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 12 suggests that it is a combination of the θ_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.

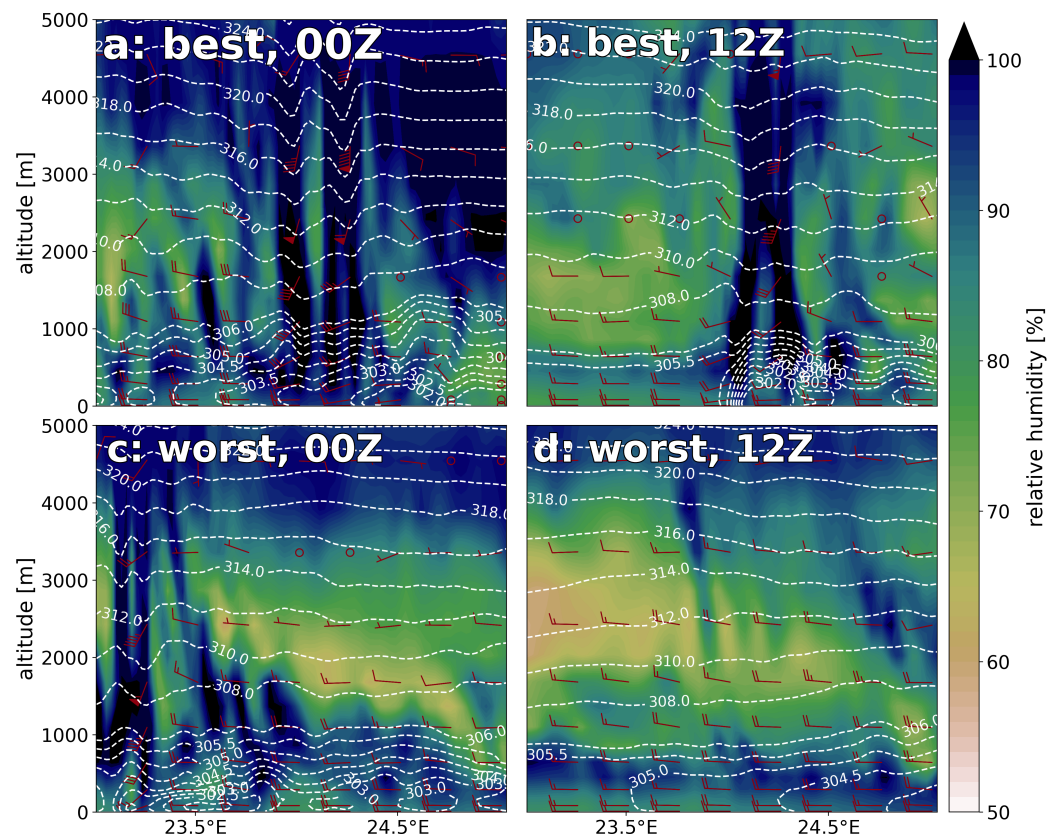


Figure 12. (a) Vertical cross-section of relative humidity (shading) and θ_v (white dashed contours) taken along the red line in Figure 8 for the best member at 00 UTC. Meridional/vertical wind vectors in units of knots are plotted in red wind barbs, with the vertical component multiplied by a factor of 10 to highlight vertical motion. (b) As in (a) at 12 UTC. (c) As in (a) for the worst member at 00 UTC. (d) As in (a) for the worst member at 12 UTC.

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 Figure 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 θ_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 Figure 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 Figures 9 and 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 aloft 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 Figures 9 and 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, give us confidence in our conclusions surrounding what lead to the bifurcation in rainfall totals among ensemble members.

In order to further generalize our conclusions and gain 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 2.3) for different meteorological fields and the statistical significance of those differences. We calculated statistical significance using a paired *t*-test as shown in Equation (11),

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \delta_{1,2}}{\hat{\sigma} \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}, \hat{\sigma} = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}} \quad (11)$$

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 Figure 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 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 Figure 13 also confirm what we have concluded in previous analysis: for both θ_e and θ_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 00 and 12 UTC than those that least accurately reproduce the rainfall pattern and totals.

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 θ_e and θ_v , as in Figure 11, and the vertical cross-section of

relative humidity, θ_v , and winds, as in Figure 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.

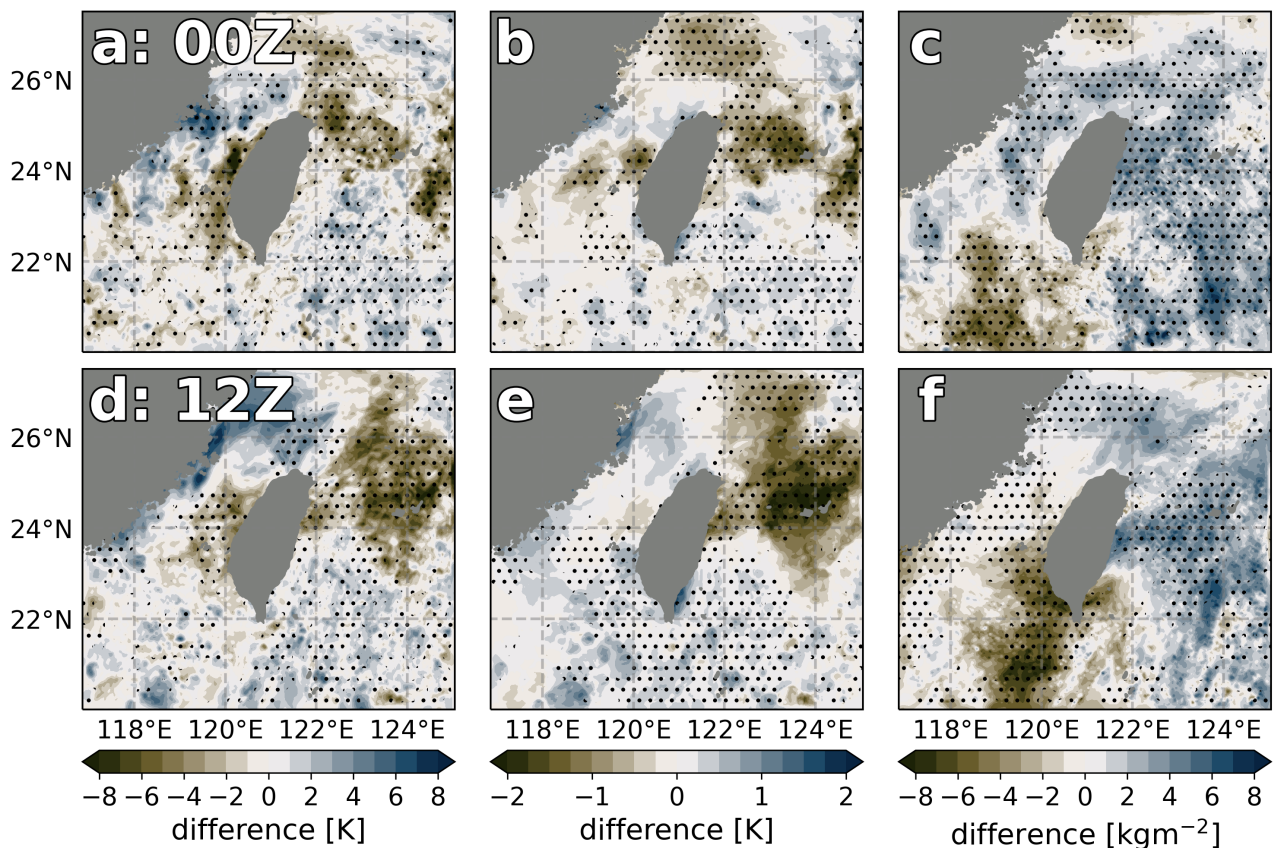


Figure 13. (a) Map of the difference in 100-m θ_e between the mean of the top five and mean of the bottom five ensemble members at 00 UTC. (b) As in (a) for the difference in 100-m θ_v at 00 UTC. (c) As in (a) for the difference in integrated water vapor at 00 UTC. (d) As in (a) at 12 UTC. (e) As in (b) at 12 UTC. (f) As in (c) at 12 UTC.

Panels a and b in Figure 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 θ_e and θ_v with similar positioning, and the front also shows up clearly in panel d in the θ_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 7, 11 and 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 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 9b).

Panel c in Figure 14 shows the 850 hPa relative humidity in the verification ensemble mean at 12 UTC. Note the difference in colorbar scale between panels c and d. The best and worst member (Figure 9d,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.

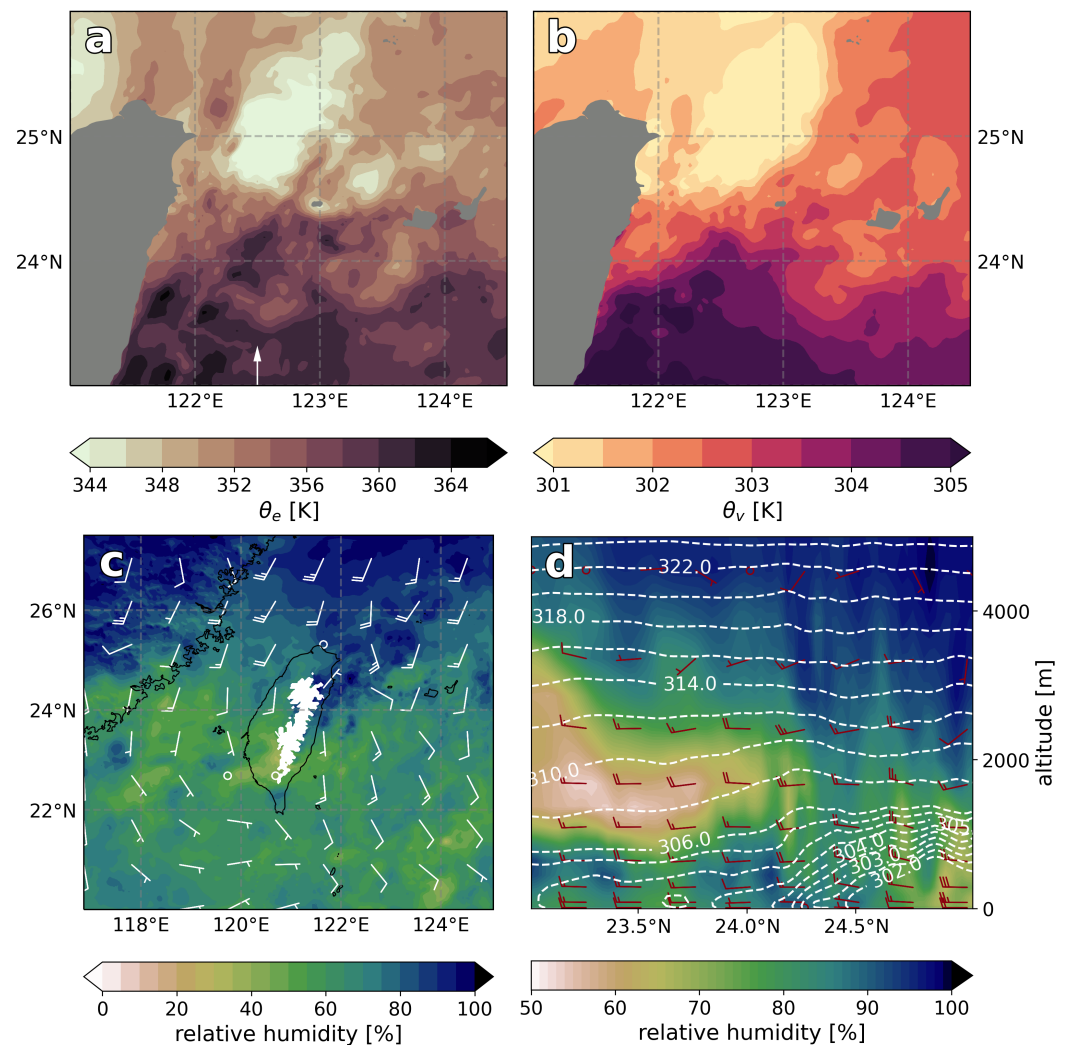


Figure 14. (a) Map of 100-m θ_e for the verification ensemble mean at 12 UTC 9 June 2020. (b) As in (a) for 100-m θ_v . (c) Map of 850 hPa relative humidity for the verification ensemble mean at 12 UTC 9 June 2020. (d) Vertical cross-section of relative humidity (shading) and θ_v (white dashed contours) taken along the red line in Figure 8 for the verification ensemble mean at 12 UTC 9 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.

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 in this case as our analysis of RH differences in Figures 9 and 10 suggested. Rather, the strength and location of the near-surface front, shown in θ_v , is the most 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.

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 9 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 ensemble 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, lower tropospheric 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 θ_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 θ_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 θ_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 Examination. 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 (Clark et al. [14]), 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 Liang et al. [13], 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.

Author Contributions: Conceptualization, A.S.C. and M.M.B.; methodology, A.S.C., M.M.B. and J.C.D.; validation, A.S.C. and M.M.B.; formal analysis A.S.C. and M.M.B.; writing—original draft preparation, A.S.C.; writing—review and editing, A.S.C., M.M.B. and J.C.D.; funding acquisition M.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation award number AGS-1854559.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to embargo under project data policy through August 2022.

Acknowledgments: We would like to thank Yunji Zhang for conducting the numerical simulations and providing the ensemble output. We would also like to thank the anonymous reviewers for their helpful comments which improved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CAPE	Convective Available Potential Energy
CIN	Convective Inhibition
CWV	Column-integrated Water Vapor
EnKF	Ensemble-Kalman Filter
ETS	Equitable Threat Score
LCL	Lifted Condensation Level
LFC	Level of Free Convection
LNB	Level of Neutral Buoyancy
MCS	Mesoscale Convective System
MDPI	Multidisciplinary Digital Publishing Institute
PRECIP	Prediction of Rainfall Extremes Campaign in the Pacific
PW	Precipitable Water
QPESUMS	Quantitative Precipitation Estimation and Segregation Using Multiple Sensors
QPF	Quantitative Precipitation Forecast
RH	Relative Humidity
ROC	Relative Operating Characteristic
WRF	Weather Research and Forecasting

References

1. Kerns, B.W.J.; Chen, Y.L.; Chang, M.Y. The diurnal cycle of winds, rain, and clouds over Taiwan during the Mei-Yu, summer, and autumn rainfall regimes. *Mon. Wea. Rev.* **2010**, *138*, 497–516. [\[CrossRef\]](#)
2. Sampe, T.; Xie, S.P. Large-scale dynamics of the meiyu-baiu rainband: Environmental forcing by the westerly jet. *J. Clim.* **2010**, *23*, 113–134. [\[CrossRef\]](#)
3. Chen, G.T.J. Mesoscale Features Observed in the Taiwan Mei-Yu Season. *J. Meteor. Soc. Jpn.* **1992**, *70*, 497–516. [\[CrossRef\]](#)
4. Chen, T.J.G.; Chang, C.P. The Structure and Vorticity Budget of an Early Summer Monsoon Trough (Mei-Yu) over Southeastern China and Japan. *Mon. Wea. Rev.* **1980**, *108*, 942–953. [\[CrossRef\]](#)
5. Akiyama, T. The large-scale aspects of the characteristic features of the Baiu front. *Pap. Meteor. Geophys.* **1973**, *24*, 157–188. [\[CrossRef\]](#)
6. Chen, G.T.J.; Yu, C.C. Study of Low-Level Jet and Extremely Heavy Rainfall over Northern Taiwan in the Mei-Yu Season. *Mon. Wea. Rev.* **1988**, *116*, 884–891. [\[CrossRef\]](#)
7. Xu, W.; Zipser, E.J.; Liu, C. Rainfall characteristics and convective properties of Mei-Yu precipitation systems over South China, Taiwan, and the South China sea. Part I: TRMM observations. *Mon. Wea. Rev.* **2009**, *137*, 4261–4275. [\[CrossRef\]](#)
8. Chen, S.J.; Kuo, Y.H.; Wang, W.; Tao, Z.Y.; Cui, B.O. A Modeling Case Study of Heavy Rainstorms along the Mei-Yu Front. *Mon. Wea. Rev.* **1998**, *126*, 2330–2351. [\[CrossRef\]](#)
9. Chen, C.S.; Chen, Y.L.; Liu, C.L.; Lin, P.L.; Chen, W.C. Statistics of heavy rainfall occurrences in Taiwan. *Wea. Forecast.* **2007**, *22*, 981–1002. [\[CrossRef\]](#)
10. Chen, C.S.; Chen, Y.L. The Rainfall Characteristics of Taiwan. *Mon. Wea. Rev.* **2003**, *131*, 132–1341. [\[CrossRef\]](#)
11. Wang, T.; Wei, K.; Ma, J. Atmospheric rivers and mei-yu rainfall in China: A case study of summer 2020. *Adv. Atmos. Sci.* **2021**, *38*, 2137–2152. [\[CrossRef\]](#)
12. Volonté, A.; Muetzelfeldt, M.; Schiemann, R.; Turner, A.G.; Klingaman, N. Magnitude, scale, and dynamics of the 2020 mei-yu rains and floods over China. *Adv. Atmos. Sci.* **2021**, *38*, 2082–2096. [\[CrossRef\]](#)
13. Liang, P.; Hu, Z.Z.; Ding, Y.; Qian, Q. The extreme mei-yu season in 2020: Role of the madden-Julian oscillation and the cooperative influence of the pacific and Indian Oceans. *Adv. Atmos. Sci.* **2021**, *38*, 2040–2054. [\[CrossRef\]](#)
14. Clark, R.T.; Wu, P.; Zhang, L.; Li, C. The anomalous Mei-yu rainfall of summer 2020 from a circulation clustering perspective: Current and possible future prevalence. *Adv. Atmos. Sci.* **2021**, *38*, 2010–2022. [\[CrossRef\]](#)
15. Doswell, C.A.; Brooks, H.E.; Maddox, R.A. Flash Flood Forecasting: An Ingredients-Based Methodology. *Wea. Forecast.* **1996**, *11*, 560–581. [\[CrossRef\]](#)
16. Lemone, M.A.; Zipser, E.J.; Trier, S.B. The Role of Environmental Shear and Thermodynamic Conditions in Determining the Structure and Evolution of Mesoscale Convective Systems during TOGA COARE. *J. Atmos. Sci.* **1998**, *55*, 3493–3518. [\[CrossRef\]](#)
17. Rotunno, R.; Klemp, J.B.; Weisman, M.L. A Theory for Strong, Long-Lived Squall Lines. *J. Atmos. Sci.* **1988**, *45*, 463–485. [\[CrossRef\]](#)
18. James, R.P.; Markowski, P.M. A numerical investigation of the effects of dry air aloft on deep convection. *Mon. Wea. Rev.* **2010**, *138*, 140–161. [\[CrossRef\]](#)
19. Holloway, C.E.; Neelin, D.J. Moisture vertical structure, column water vapor, and tropical deep convection. *J. Atmos. Sci.* **2009**, *66*, 1665–1683. [\[CrossRef\]](#)
20. Neelin, J.D.; Peters, O.; Hales, K. The transition to strong convection. *J. Atmos. Sci.* **2009**, *66*, 2367–2384. [\[CrossRef\]](#)
21. Molinari, J.; Romps, D.M.; Vollaro, D.; Nguyen, L. CAPE in tropical cyclones. *J. Atmos. Sci.* **2012**, *69*, 2452–2463. [\[CrossRef\]](#)
22. Sukovich, E.M.; Ralph, F.M.; Barthold, F.E.; Reynolds, D.W.; Novak, D.R. Extreme quantitative precipitation forecast performance at the weather prediction center from 2001 to 2011. *Wea. Forecast.* **2014**, *29*, 894–911. [\[CrossRef\]](#)
23. Bryan, H. G.; Wyngaard, Fritsch, J.C.A.; Michael, J. Resolution requirements for the simulation of deep moist convection. *mon* **2003**, *131*, 2394–2416. [\[CrossRef\]](#)
24. Smith, M.B.; Torn, R.D.; Corbosiero, K.L.; Pegion, P. Ensemble variability in rainfall forecasts of Hurricane Irene (2011). *Weather Forecast.* **2020**, *35*, 1761–1781. [\[CrossRef\]](#)
25. Jee, J.B.; Kim, S. Sensitivity study on high-resolution WRF precipitation forecast for a heavy rainfall event. *Atmosphere* **2017**, *8*, 96. [\[CrossRef\]](#)
26. Wang, C.C. The more rain, the better the model performs—The dependency of quantitative precipitation forecast skill on rainfall amount for typhoons in Taiwan. *Mon. Weather Rev.* **2015**, *143*, 1723–1748. [\[CrossRef\]](#)
27. Nielsen, E.R.; Schumacher, R.S. Using convection-allowing ensembles to understand the predictability of an extreme rainfall event. *Mon. Wea. Rev.* **2016**, *144*, 3651–3676. [\[CrossRef\]](#)
28. Lorenz, E.N. The predictability of a flow which possesses many scales of motion. *Tellus* **1969**, *21*, 289–307. [\[CrossRef\]](#)
29. Durran, D.R.; Weyn, J.A. Thunderstorms do not get butterflies. *Bull. Am. Meteor. Soc.* **2016**, *97*, 237–243. [\[CrossRef\]](#)
30. Fang, X.; Kuo, Y.H. Improving ensemble-based quantitative precipitation forecasts for topography-enhanced typhoon heavy rainfall over Taiwan with a modified probability-matching technique. *Mon. Wea. Rev.* **2013**, *141*, 3908–3932. [\[CrossRef\]](#)
31. Boucher, M.A.; Anctil, F.; Perreault, L.; Tremblay, D. A comparison between ensemble and deterministic hydrological forecasts in an operational context. *Adv. Geosci.* **2011**, *29*, 85–94. [\[CrossRef\]](#)
32. Schumacher, R.S.; Davis, C.A. Ensemble-based forecast uncertainty analysis of diverse heavy rainfall events. *Weather Forecast.* **2010**, *25*, 1103–1122. [\[CrossRef\]](#)

33. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Liu, Z.; Berner, J.; Wang, W.; Powers, J.G.; Duda, M.G.; Barker, D.; et al. *A Description of the Advanced Research WRF Model Version 4*; Technical Report, NCAR/TN-475+ STR; National Center for Atmospheric Research: Boulder, CO, USA, 2019. [\[CrossRef\]](#)
34. Zhang, F.; Weng, Y.; Sippel, J.A.; Meng, Z.; Bishop, C.H. Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Weather Rev.* **2009**, *137*, 2105–2125. [\[CrossRef\]](#)
35. Weng, Y.; Zhang, F. Assimilating airborne Doppler radar observations with an ensemble Kalman filter for convection-permitting hurricane initialization and prediction: Katrina (2005). *Mon. Weather Rev.* **2012**, *140*, 841–859. [\[CrossRef\]](#)
36. Zhang, F.; Minamide, M.; Nystrom, R.G.; Chen, X.; Lin, S.J.; Harris, L.M. Improving Harvey forecasts with next-generation weather satellites: Advanced hurricane analysis and prediction with assimilation of GOES-R all-sky radiances. *Bull. Am. Meteorol. Soc.* **2019**, *100*, 1217–1222. [\[CrossRef\]](#)
37. Chang, P.L.; Zhang, J.; Tang, Y.S.; Tang, L.; Lin, P.F.; Langston, C.; Kaney, B.; Chen, C.R.; Howard, K. An operational multi-radar multi-sensor QPE system in Taiwan. *Bull. Am. Meteorol. Soc.* **2021**, *102*, E555–E577. [\[CrossRef\]](#)
38. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [\[CrossRef\]](#)
39. Bolton, D. The Computation of Equivalent Potential Temperature. *Mon. Wea. Rev.* **1980**, *108*, 1046–1053. [\[CrossRef\]](#)
40. May, R.M.; Arms, S.C.; Marsh, P.; Bruning, E.; Leeman, J.R.; Goebbert, K.; Thielen, J.E.; Bruick, Z.S.; Camron, M.D. MetPy: A Python Package for Meteorological Data. 2022. Available online: <https://github.com/Unidata/MetPy> (accessed on 15 August 2019).
41. Romps, D.M.; Kuang, Z. Do undiluted convective plumes exist in the upper tropical troposphere? *J. Atmos. Sci.* **2010**, *67*, 468–484. [\[CrossRef\]](#)
42. Thomas, C.M.; Schultz, D.M. What are the best thermodynamic quantity and function to define a front in gridded model output? *Bull. Am. Meteorol. Soc.* **2019**, *100*, 873–895. [\[CrossRef\]](#)
43. Markowski, P.; Richardson, Y. *Mesoscale Meteorology in Midlatitudes*; Advancing Weather and Climate Science; Wiley: Hoboken, NJ, USA, 2011.
44. Bluestein, H. *Synoptic-Dynamic Meteorology in Midlatitudes: Principles of Kinematics and Dynamics*; Oxford University Press: Oxford, UK, 1992.