Radar Wind Retrieval, Ensemble Simulations, and Cluster Analysis of a Heavy Rainfall Event over Northern Taiwan during 1–2 June 2017

Chieh-Ying Ke[a], Pei-Jung Tsai[b], Kao-Shen Chung[a], Yi-Leng Chen[c], Yu-Chieng Liou[a]

[a]*Department of Atmospheric Sciences, National Central University, Taoyuan, Taiwan*

[b]*Central Weather Administration, Taipei, Taiwan*

[c]*Department of Atmospheric Sciences, University of Hawaii at Mānoa, Honolulu, Hawaii, USA*

Corresponding author: Kao-Shen Chung, kaoshen.chung@gmail.com, kschung@atm.ncu.edu.tw

ABSTRACT

This study examines a heavy rainfall event over northern Taiwan from 1–2 June 2017, focusing on the interactions between the marine boundary layer jet (MBLJ), frontal cyclone, and barrier jet (BJ) in influencing rainfall distribution. Three mesoscale convective systems (MCSs) moved eastward, producing intense, prolonged rainfall and transforming the linear convective line from a TS-type (Trailing Stratiform) to a PS-type (Parallel Stratiform) structure. Ensemble simulations, divided into five clusters using k-means clustering with spatial adjustments, reveal distinct rainfall distribution patterns, each linked to specific wind dynamics and mesoscale topographic interactions. Further categorization identified two primary rainfall types: Type A, closely matching the observed event, with substantial inland rainfall driven by a nearby low-pressure center and southwesterly MBLJ; and Type B, with offshore rainfall influenced by a more distant low-pressure center and stronger southerly MBLJ that extends BJs further north. Findings underscore the role of MBLJ, BJ, and cyclone positioning in shaping rainfall distribution, offering insights for more accurate rainfall forecasts in Taiwan's complex terrain by capturing atmospheric variabilities and informing future forecasting models.

Significance Statement

This study explores the factors behind intense rainfall during Taiwan's Mei-Yu season, focusing on how weather systems and the island's mountains shape rainfall patterns. By analyzing different rainfall scenarios, the study identified five main patterns influenced by changing winds and terrain. It also highlighted two key rainfall types: one with more rain inland and another over the ocean, each linked to different weather conditions. These findings help improve our understanding of heavy rainfall in Taiwan and offer valuable insights for better weather prediction.

Key words

Radar retrieval, Ensemble simulation, Clustering, Extremely heavy rainfall

1. Introduction

During the Mei-Yu Season over the Taiwan area from mid-May to mid-June (Chen and Tu 2024), the Mei-Yu front interacts with various mesoscale convective systems (MCSs, Houze (2004)) and the island terrain, causing extreme rainfall events. This significant weather system brings heavy rainfall to Taiwan, serving as a vital water resource but also posing a serious flooding risk. Accurate forecasting of the Mei-Yu front is crucial to mitigate potential damage and effectively manage water resources.

Numerous studies investigated the heavy rainfall events in northern Taiwan. They demonstrate a significant correlation between subsynaptic low-level jets (SLLJs) at the 850-700 hPa layer (G. Chen and Yu 1988; G. Chen et al. 2005), which are part of the secondary circulations associated with the jet/front systems (Chen et al. 1995; Chen et al. (1997), and extreme rainfall events in northern Taiwan. Observations from radiosondes, surface measurements, and aircraft during the 1987 Taiwan Area Mesoscale Experiment (TAMEX) identified a barrier jet around 1 km above ground level, caused by orographic blocking of airflow by Taiwan's terrain(Chen and Li 1995a, 1995b; Li and Chen 1998; Yeh and Chen 2003). This strong, terrain-induced wind significantly impacts heavy rainfall patterns in the region (Chen and Li 1995a; Li and Chen 1998; Li et al. 1997). In the past decades, with advances in computing resources and high-resolution observational data, many studies focused on case studies that employ high-resolution numerical weather prediction (NWP) models (Chen et al. 2013; Lin et al. 2011; Yeh and Chen 2003). During the Mei-Yu season, the marine boundary layer jets (MBLJs) play a crucial role in the horizontal transport of moisture from the northern South China Sea to the Taiwan area, with a wind speed maximum around the 925-hPa level (Tu et al. 2019; Tu et al. 2020; Tu et al. 2022). Chen et al. (2022) found that when an MBLJ or a strong low-level jet (SLLJ) is present, a barrier jet exceeding 14 m s^{-1} forms around the 925-hPa level off northwestern Taiwan, which enhances horizontal moisture transport towards the frontal zone due to island blocking and orographic effects.

Tu et al. (2022) investigated the heavy rainfall event on 2 June 2017 in northern Taiwan by using high-resolution simulations and observations. Their study emphasizes the orographic effects on localized heavy rainfall over northern Taiwan. The unusually shallow east–west-oriented mei-yu front (<850 m) is anchored over the northern side of the Yang-Ming Mountains (with peaks ~1120 m) for 8 h, resulted in torrential rain >600 mm there. Stations within the

Taipei Basin experienced two distinct temperature drops. The first temperature drop is attributable to the shallow postfrontal cold air at the surface is deflected by the Yang-Ming Mountains and moves through the Keelung River and Tamsui River Valley into the Taipei Basin. The second temperature drop is due to the arrival of the deepening postfrontal cold air, which eventually climbs over the Yang-Ming Mountains, and is modified by rainfall evaporation cooling. After the arrival of the shallow cold front, concentrated heavy rainfall occurred over the northwestern coast due to the localized convergence between the barrier jet and the postfrontal cold air whereas heavy rainfall within the Taipei Basin is due to frontal convergence. Cornejo et al. (2024) examined how Taiwan's topography intensifies extreme rainfall associated with Mei-Yu fronts, focusing on a case from 1-3 June 2017. By simulating this event with the WRF Model and reducing terrain height as a sensitivity test, the study found that lower terrain significantly decreased rainfall accumulation and duration in northern Taiwan, primarily due to a weakened orographic barrier jet. Additionally, in the Central Mountain Range, rainfall intensity and duration dropped as orographic effects on frontal convergence diminished. These findings underscore the critical role of Taiwan’s northern topography in enhancing rainfall during Mei-Yu events.

Meteorological radar data, with their high spatiotemporal resolution, provide a robust resource for studying atmospheric phenomena, especially when paired with retrieval techniques such as the WISSDOM wind field inversion method (Liou et al., 2016). In Taiwan, where dense radar coverage intersects complex terrain, these retrievals have been instrumental in examining weather events like fronts, afternoon convection, and typhoons. Studies utilizing Doppler radar, such as Ke et al. (2019) on the 2012 Mei-Yu rainfall event, have shown the role of jet dynamics, particularly the evolution of the barrier jet, in influencing precipitation patterns over Taiwan. Ke et al. (2019) used multiple Doppler radar retrievals to analyze how interactions between the barrier jet and Taiwan’s mountainous topography impacted rainfall distribution and intensity. Their findings demonstrated that the barrier jet strengthened due to orographic effects, which in turn led to enhanced convergence and uplift along the northern coastal region of Taiwan, intensifying rainfall during the Mei-Yu season. These insights significantly deepen our understanding of how mesoscale jet dynamics interact with complex terrain to shape rainfall systems in Taiwan. Such radar-based analyses underscore the essential role of detailed observational data in capturing the complex dynamics of heavy rainfall events during the Mei-Yu season.

Predicting extreme rainfall is inherently challenging due to uncertainties in numerical weather models, which stem from both the complexity of atmospheric processes and limitations in observational data and model resolution (Leutbecher and Palmer 2008). Ensemble forecasting, which involves generating multiple simulations with varied initial conditions, is often employed to address these uncertainties by estimating a range of possible outcomes, thereby improving forecast reliability and capturing a wider scope of model error. This approach helps account for the unpredictable nature of convective systems, essential in extreme precipitation scenarios, and highlights the value of probabilistic methods in meteorology (Buizza et al. 2005; Leutbecher and Palmer 2008). Ensemble modeling, combined with machine learning techniques like clustering, enables the analysis of critical factors such as moisture, wind shear, and mesoscale fronts, ultimately advancing our understanding of rainband evolution. Clustering algorithms classify data points into groups on the basis of similarities and dissimilarities in various parameters, such as sensitive factors, dynamic structure, microphysics contribution, and moisture content. This approach facilitates rapid understanding of the underlying information within data (Wang et al. 2021; Xue et al. 2011; Zheng et al. 2019; Zheng et al. 2017).

This study builds on existing knowledge, applying machine learning and clustering to classify rainband evolution patterns within ensemble simulations, aiming to objectively capture the inherent variability and forecast uncertainty of Mei-Yu season rainfall. The key focus of this study is the interaction between the marine boundary layer jet and the frontal cyclone, which serves as a critical mechanism for determining the location of heavy Mei-Yu rainfall. The research explores various factors influencing rainfall prediction, including marine boundary layer jets, barrier jets, and surface low pressure systems. While the methodology is sound, the study would benefit from a clearer structure, along with a deeper discussion of certain assumptions, physical processes, and broader real-world applications. With further refinements, this framework has the potential to significantly improve the accuracy of extreme rainfall forecasts in Taiwan, providing a valuable tool for future meteorological predictions.

In this study, we employ 3D wind data retrieved by the WISSDOM algorithm as the ground truth to investigate, through ensemble simulations, the mechanisms of the frontal system during the Mei-Yu event on 1–2 June 2017. The radar data, retrieved method, model simulations, and k-means clustering approach are described in section 2. Section 3 provides an overview of the 1–2 June 2017 case. To address the challenges related to the objectivity of the cluster number

and cluster specificity, we adopt a spatially correlated k-means clustering approach. Section 4 presents the results of retrieval using WISSDOM. In Section 5, the application of the modified k-means clustering analysis to ensemble simulations is described. Finally, the conclusions of this study are summarized in Section 6.

2. Data and methodology

a. Observational Data

The observational data used in this study includes weather maps, surface rain gauge data, and composite radar reflectivity images acquired from the CWA. Weather maps at 200 hPa, 500 hPa, 850 hPa, and surface levels provide an overview of the synoptic pattern of the 2017 case. Rain gauge data, satellite imagery, and composite radar reflectivity images are used to capture the evolution of the frontal system.

The rain-gauge observations are from the conventional surface weather stations and the Automatic Rainfall and Meteorological Telemetry System (ARMTS, Kerns et al. (2010)). The enhanced color infrared image (Band 13, 10.4 μm) obtained from the Himawari-8 geostationary satellite depicts a frontal zone located north of Taiwan on June 1, 2017. These images, provided by the Central Weather Administration, also indicates a correlation between this frontal zone and the eastward-moving mesoscale convective systems (MCSs).

This study employs data from two weather radars: RCWF (25.07°N, 121.77°E, 766 m above sea level) and NCU-CPol (24.97°N, 121.18°E, 196 m above sea level; Fig. 6). RCWF is an S-band (10.7-cm) radar located at Wu-Fen-Shan Mountain and operated by the CWA. That employs nine elevation scanning angles (0.5°, 1.5°, 2.4°, 3.4°, 4.3°, 6.0°, 9.9°, 14.6°, and 19.5°) and has a maximum scanning range of 300 km with a Nyquist velocity of 31.0 m s^{-1}. NCU-CPol is a C-band (5.3-cm) dual-polarimetric doppler radar at National Central University, which has nine elevation scanning angles, maximum range of 250 km horizontally, and Nyquist velocity of 31.9 m s^{-1}. To ensure the data's accuracy, a quality control procedure is implemented. Data points were excluded when the cross-correlation coefficient was <0.8 or when high reflectivity (>40 dBZ) coincided with low radial wind speed (<2 m s^{-1}). To process the radial wind data, an unfolding procedure is applied in the spherical coordinate system. All radar observations are then interpolated onto a 3D Cartesian grid using the Constant Altitude Plan

178 Position Indicator (CAPPI) method, with a horizontal resolution of 1 km and a vertical
179 resolution of 0.5 km for use in the retrieval algorithm.

180

181 ***b. The Wind Synthesis System using Doppler Measurements (WISSDOM)***

182 This study employed the WISSDOM algorithm (Liou et al. 2016), a variational method
183 (Liou and Chang 2009; Liou et al. 2012) that merges radar observations with background data
184 from model simulations to reconstruct a 3D wind field with a horizontal resolution of 1 km.
185 The accuracy of this algorithm has been validated in idealized and real-world cases. To guide
186 the wind field toward a realistic atmospheric state, WISSDOM minimizes a cost function with
187 seven key constraints: it aligns the retrieved wind with radar radial wind observations, enforces
188 continuity to maintain an incompressible fluid approximation, incorporates vertical vorticity to
189 capture wind dynamics, and includes a background term to support areas with limited data
190 coverage. Additional terms apply Laplacian smoothing to reduce extreme values, while vertical
191 velocity constraints assume zero vertical velocity at the upper and lower boundaries based on
192 mass conservation principles.

193 This variational approach offers several advantages: (a) it enables reconstruction of the 3D
194 wind field at the radar's native resolution; (b) the immersed boundary method (Tseng and
195 Ferziger 2003) allows for realistic interaction between wind fields and complex topography; (c)
196 diverse data sources, including radiosonde and ground-based observations, can supplement
197 areas without radar coverage; and (d) data from various radar systems can be effectively
198 integrated.

199 The WISSDOM wind retrieval algorithm depends on specific assumptions, including
200 constraints on vertical velocity, which may indeed introduce bias in complex atmospheric
201 conditions, such as during strong turbulence or rapid frontal movements. However, previous
202 studies using WISSDOM in Taiwan's complex terrain have demonstrated its reliability. For
203 instance, Ke et al. (2019) applied WISSDOM to retrieve wind fields for a significant rainfall
204 event in northern Taiwan. Their results showed that WISSDOM effectively captured the
205 characteristics of strong convective systems in this region, supporting its validity and
206 robustness in the context of Taiwan's orographically influenced weather patterns.

In this study, WISSDOM was applied over northern Taiwan with a horizontal resolution of 1 km and a vertical resolution of 0.5 km (31 levels, extending up to 15 km). The algorithm utilized hourly radar reflectivity (ZH) and radial velocity (VR) data from NCU C-Pol and RCWF, and the background field from WRF model simulation to retrieve the wind field in this region.

c. Model configuration and ensemble simulation

This study employs the Advanced Research Weather and Research Forecast (WRF) Model version 4.2.1 (Gochis et al. 2017; Skamarock 2019) to supplement missing radar data values in the retrieval algorithm system (referred to as the background term) and perform ensemble analysis. The WRF model configuration incorporated the following physical parameterization schemes: Grell-Devenyi ensemble cumulus scheme only for 27- and 9-km domains (Grell and Dévényi 2002); Goddard 4ICE microphysics scheme (Tao et al. 1989; Tao et al. 2016); Yonsei University planetary boundary layer scheme (Hong and Lim 2006); Dudhia short-wave radiation scheme (Dudhia 1989); and rapid radiative transfer model longwave radiation scheme (Mlawer et al. 1997).

The National Centers for Environmental Prediction (NCEP) final operational global analysis data with a resolution of 0.25° × 0.25° served as the initial and boundary conditions for the WRF simulations. A four-domain nesting strategy was implemented; the domains were designated d01, d02, d03, and d04 and had horizontal grid resolution of 27, 9, 3, and 1 km, respectively (Fig. 1). An ensemble of 128 members was generated by applying WRF-3DVar random perturbations (Barker et al. 2004) in d01. These perturbations were then nested down to the 9-, 3-, and 1-km resolutions. Each ensemble member was integrated over a 30-hour simulation period, starting from 0000 UTC 1 June 2017 to 0600 UTC 2 June 2017, with a spin-up time of 5 hours.

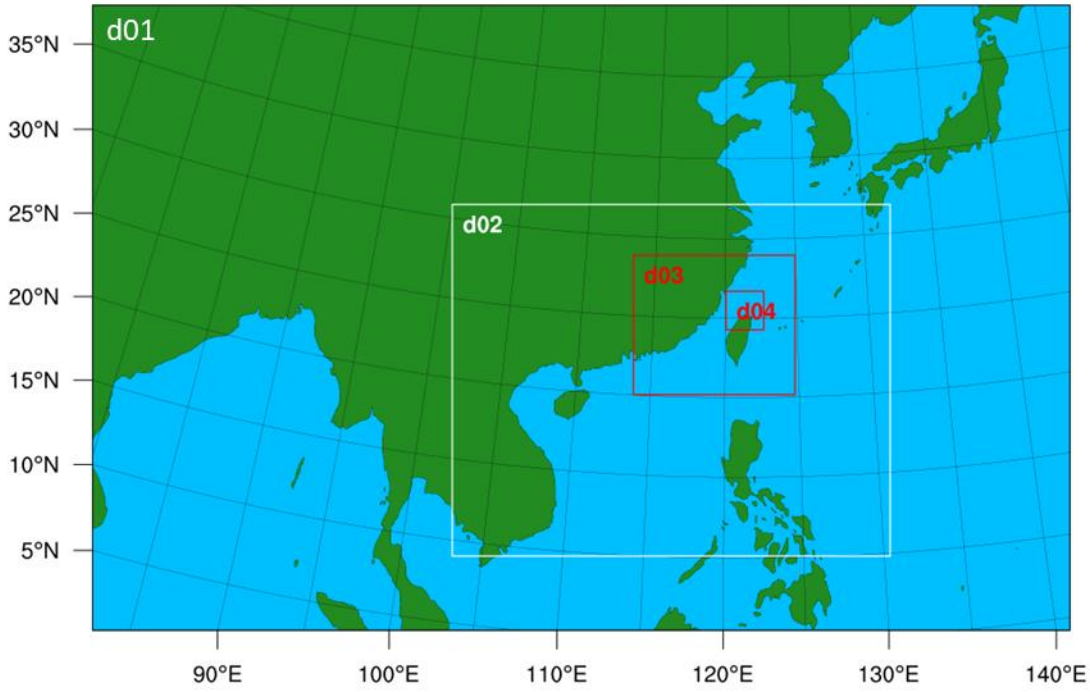


Fig. 1. The WRF domains for the WISSDOM background term and the ensemble simulation. Horizontal grid resolutions of 27, 9, 3, and 1 km were applied for d01, d02, d03, and d04, respectively.

d. Spatially correlated k-means clustering

1) k-means clustering

The k-means clustering method originated in the 1960s and was first employed by MacQueen (1967). In this clustering technique, "n" members are separated into k clusters such that members with similar characteristics are grouped together whereas significant differences in performance exist between the centroids of the groups. The k-means clustering concept, predominantly driven by the squared error (SE), is defined as follows (Teng et al. 2019):

$$D_{in} = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_j - \mu_i\right)^2 \quad (1),$$

$$D_{clu} = \frac{1}{2}\sum_{i=1}^{K}\sum_{m=1}^{K}(\mu_m - \mu_i)^2 \quad (2),$$

where x_j is the jth member of the ith cluster among all K clusters, and the μ_i is the center of the ith cluster. The objective of k-means clustering is to minimize the intracluster distance D_{in} and maximize the mean distance between clustering groups D_{clu}. In the k-means clustering algorithm, the number of clusters k is predetermined before the algorithm is implemented,

which leads to subjectivity. When k is larger, the connections within clusters are stronger but the distinctions between clusters are weaker. Therefore, a larger k may not be the best fit for all datasets. The silhouette coefficient, a metric for assessing clustering quality, is calculated as follows.

$$\text{Silhouette coefficient: } s = \frac{(D_{clu} - D_{in})}{\max(D_{in}, D_{clu})} \quad (3).$$

When s is close to 1, with a larger D_{clu} and smaller D_{in}, the clustering result is considered the best result. Through iterative computations by using k-means clustering and the silhouette coefficient, the optimal value of k from a set of predefined k values can be determined. This process ensures a valid solution across all potential cluster configurations and provides an initial understanding of the ensemble's characteristics.

In this study, the clustering process begins by identifying the extreme rainfall locations for each ensemble member within domain d04, constrained by an 8-hour accumulated rainfall threshold of over 100 mm (Fig. 2a). Next, the desired cluster number k is specified, and the initial centroids are either randomly selected or customized (Fig. 2b). The clustering process then groups all ensemble members into k clusters by minimizing the squared error between each sample (ensemble members) and its nearest centroid (Fig. 2c). For a sensitivity test, Fig. 2d shows the maximum silhouette coefficient values across different cluster numbers (from 3 to 9), calculated by running the clustering process 1000 times with randomly initialized centroids. This approach aims to determine the optimal k -value while considering potential errors from the initial centroid settings in k-means clustering. The silhouette score ranges from -1 to 1, where higher values indicate better-defined clusters. Results from this study show that the highest silhouette score achieved is 0.4067 for five clusters. While the maximum possible silhouette score is 1, scores near zero indicate overlapping clusters, and negative values suggest incorrect clustering. The small range of silhouette scores observed (0.38–0.41) suggests relatively subtle differences between tested k -value, but the 5 clustering provides the most distinct inter-cluster separation and intra-cluster cohesion, making it suitable for further analysis.

2) Spatially correlated by Spatial Correlation Coefficient (SCC)

Building upon the valuable information derived from the distribution and amount of rainfall, the second step was to employ a spatially correlated k-means clustering approach based on mean value of spatial correlation coefficient (SCC). This method is formulated as follows:

$$\text{Mean value of SCC: } c = \frac{1}{K \times n} \sum_{i=1}^{K} \sum_{j=1}^{n} \frac{\sum \left(x_{j,(2D)} - \overline{x_{j,(2D)}}\right)\left(\mu_{i,(2D)} - \overline{\mu_{i,(2D)}}\right)}{\sqrt{\sum \left(x_{j,(2D)} - \overline{x_{j,(2D)}}\right)^2 \sum \left(\mu_{i,(2D)} - \overline{\mu_{i,(2D)}}\right)^2}} \quad (4),$$

where x is the spatial distribution in the jth member of the ith cluster, and μ is the mean pattern of the ith cluster. The objective of spatially correlated k-means clustering is to maximize the mean SCC value for the correlation between each member and its cluster mean pattern, which results in the optimal clustering outcome.

To highlight the impact of heavy rainfall events, only ensemble members experiencing rainfall exceeding 20 mm within 8 hours were retained before conducting the spatially correlated k-means clustering. This study achieved an objective clustering outcome by performing an initial k-means clustering, followed by a spatially correlated k-means clustering approach, as outlined in Fig. 3. The methodology involved three main steps:

1. Initial Clustering: The k-means clustering algorithm was applied with a cluster number of 5, yielding an initial mean spatial correlation coefficient (SCC) of 0.688. Each member was assigned to one of the 5 clusters by maximizing the SCC between each ensemble members and its cluster centroid.
2. Outlier Removal and Cluster Refinement: In the spatially correlated clustering step, any members with an SCC below 0.6, indicating low correlation with the rainfall cluster centroids, were considered outliers and excluded from further analysis. After removing these outliers, the mean SCC improved to 0.736. The centroids of each cluster were then recalculated based on the mean of the assigned ensemble members.
3. Iterative Adjustment and Convergence: Step 2 was repeated iteratively until the centroids and member assignments stabilized, meeting a convergence criterion. The final iteration produced a mean SCC of 0.742.

This integrated approach yielded an objective clustering result of five distinct groups. By prioritizing heavy rainfall information within the spatial clustering framework, this method provided a solid foundation for further analysis.

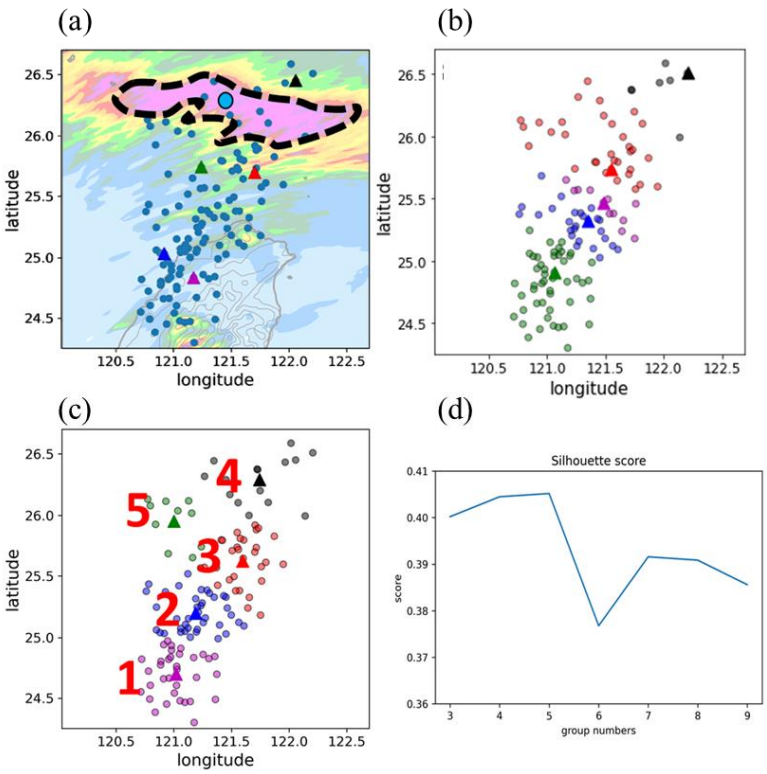


Fig. 2. k-means clustering processes. (a) 100-mm rainfall accumulation area (black dash line) for 8-h and heavy rainfall accumulation center of 128 ensemble members (blue point); (b) Set 5 members randomly from all ensemble members as the initial centroids. (c) Separate all members into 5 clusters through the min squared error between ensemble members and each centroid. (d) Silhouette coefficient of different groups numbers (3 to 9).

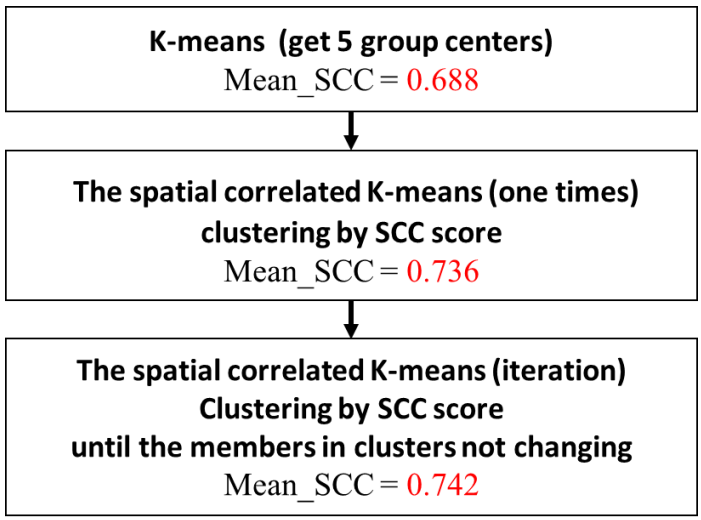


Fig. 3. Spatially correlated k-means clustering processes.

3. Description of the 1–2 June 2017 Mei-Yu case

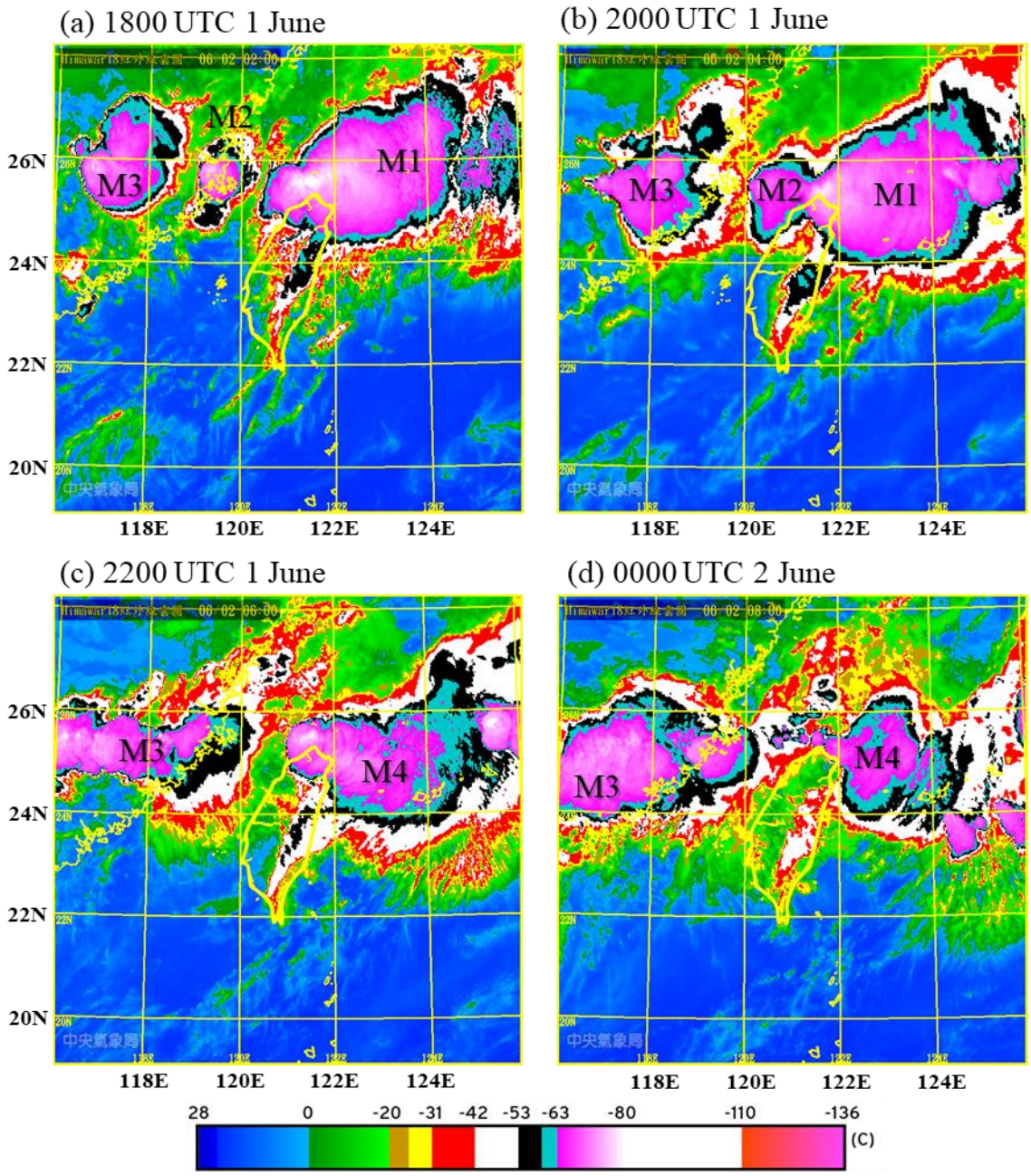
A Mei-Yu front caused very heavy rainfall in the northern Taiwan from 1–2 June 2017. At 1200 UTC 1 June 2017, two Mei-Yu frontal cyclones developed within the 850-hPa pressure trough along the South China coast (Tu et al. 2022). Concurrently, the western Pacific subtropical high exhibited a westward extension into the South China Sea, accompanied by a strong MBLJ on its southeastern side. By 0000 UTC 2 June, the frontal cyclones moved eastward from the South China coast, transiting across northern Taiwan. By 1200 UTC 2 June, the Mei-Yu frontal cyclones and MBLJ moved slowly eastward and then weakened. The simulations presented in Tu et al. (2022) suggest the MBLJ is related to pressure gradients between the deepening of the Mei-Yu frontal cyclones and/or the strengthening of the Northwest Pacific Subtropical High.

Enhanced color infrared satellite imagery from Himawari-8, shown in Fig. 4, captures mesoscale convective systems (MCSs) positioned north of Taiwan on 1 June 2017. At 1800 UTC, three MCSs were aligned in a west-to-east orientation, extending from the South China coast to northern Taiwan. The northernmost MCS, designated M1 and located in the offshore north of Taiwan, was the most developed and highest. M2, a newly forming system, was located in the Taiwan Strait, whereas M3 was in the coastal region of South China. By 2000 UTC, M2 moved eastward and merged with M1. At 2200 UTC, the merged system, M4, was visible, whereas M3 remained in the coastal area of South China. By 0000 UTC 2 June, the areas of lower brightness temperature at higher altitudes of M4 over northern Taiwan decreased considerably. However, the system itself persisted and continued to develop from the South China coast to the Taiwan Strait.

The composite radar observation reflectivity data acquired from the Central Weather Administration (CWA) weather radar network (Fig. 5) provided crucial details regarding the spatial distribution and evolution of the precipitation associated with the Mei-Yu front. At 1800 UTC 1 June, the radar composite reflectivity shows a prominent southwest-to-northeast-oriented squall line along the northern coast of Taiwan, characterized as a TS-type linear convective system within MCSs (Parker and Johnson 2000). Scattered rainfall echoes were also observed over the central and southern mountainous regions. This pattern indicated an environment characterized by abundant moisture, with southwesterly winds on the southern

side of the front transporting moisture to the western mountainous areas of Taiwan. Due to the orographic lifting effect, the precipitation was subsequently triggered at these windward regions. At 2000 UTC, the merger of M2 and M1, identified in the Himawari-8 imagery, was evident in the northernmost part of Taiwan. This was corroborated by the radar reflectivity data displaying the strongest echo in this region. At this time, the original southwest-to-northeast-oriented squall line, initially a trailing stratiform (TS-type) structure, began shifting its stratiform region to the northeast side of the strong convection, evolving into an east-west-oriented parallel stratiform (PS-type) squall line. The progression continued at 2200 UTC, with the heavy rainfall echoes persisting over northern Taiwan. The strong radar reflectivity, initially forming a west-to-east linear pattern, transitioned to a more localized concentration over northern Taiwan. The echoes on the western side weakened, exhibiting a stratiform region, indicative of a squall line adopting a parallel stratiform structure (PS-type). This pattern indicated sustained, intense precipitation in the northern coastal areas. Simultaneously, scattered echoes persisted in the southern mountainous regions, consistent with the continued inflow of abundant moisture by the southwesterly winds. By 0000 UTC 2 June, strong radar echoes persisted along the northern coast of Taiwan, extending eastward and forming a narrow east–west line. The stratiform region to the east shifted further eastward, whereas to the west, the line connected to M3. Overall, this intense linear convection in northern Taiwan transitioned from a TS-type to a PS-type squall line as the Mei-Yu frontal cyclone moved eastward to a position northeast of Taiwan (Tu et al. 2022), with the strong convective system persisting for more than 6 hours. This behavior closely resembles the heavy rainfall event observed in northern Taiwan in 2012 (Ke et al. 2019).

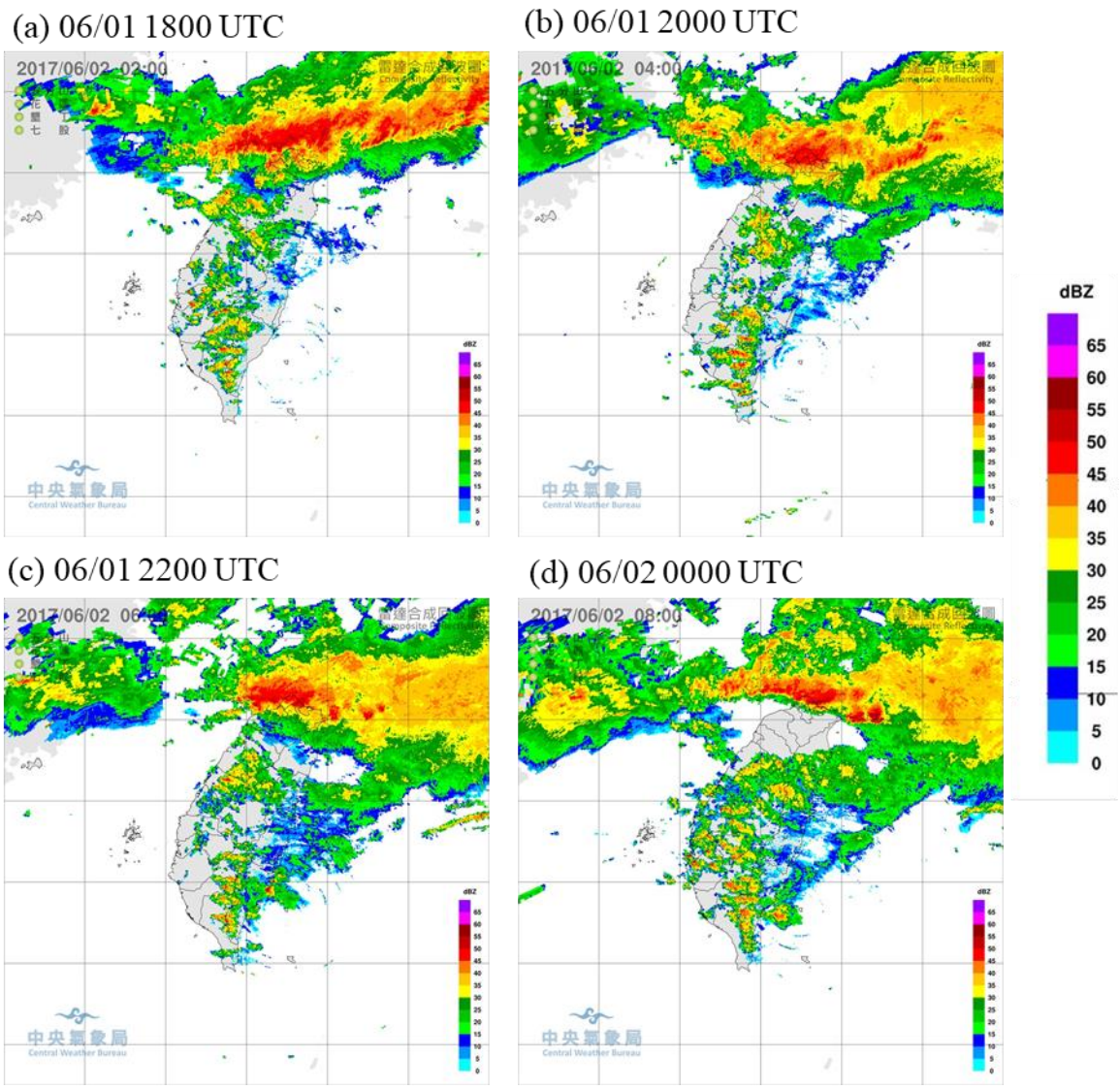
The 8-h accumulated rainfall during this event exceeded 400 mm, and the maximum rainfall concentrated in the northernmost region of Taiwan, as illustrated in Fig. 6. According to Tu et al. (2022, Fig. 7), the rainfall timing at northern stations exhibited two distinct peaks. The first occurred from 1800 UTC 1 June to 0200 UTC 2 June when the shallow cold front was anchored along the windward side of the Yang-Ming Mountains. The second peak was from 0300 UTC to 0700 UTC 2 June after the shallow cold front was able to move over the mountains and into the Taipei Basin. The present study focuses on the first rainfall peak and aims to elucidate its characteristics and underlying processes through ensemble simulations.



377

378 Fig. 4. The color-enhanced infrared images (unit: degree C) from the Himawari-8 geostationary satellite. M1-
379 M4 are mesoscale convective systems. (a) 1800 UTC 1 June; (b) 2000 UTC 1 June; (c) 2200 UTC 1 June;
380 (d) 0000 UTC 2 June.

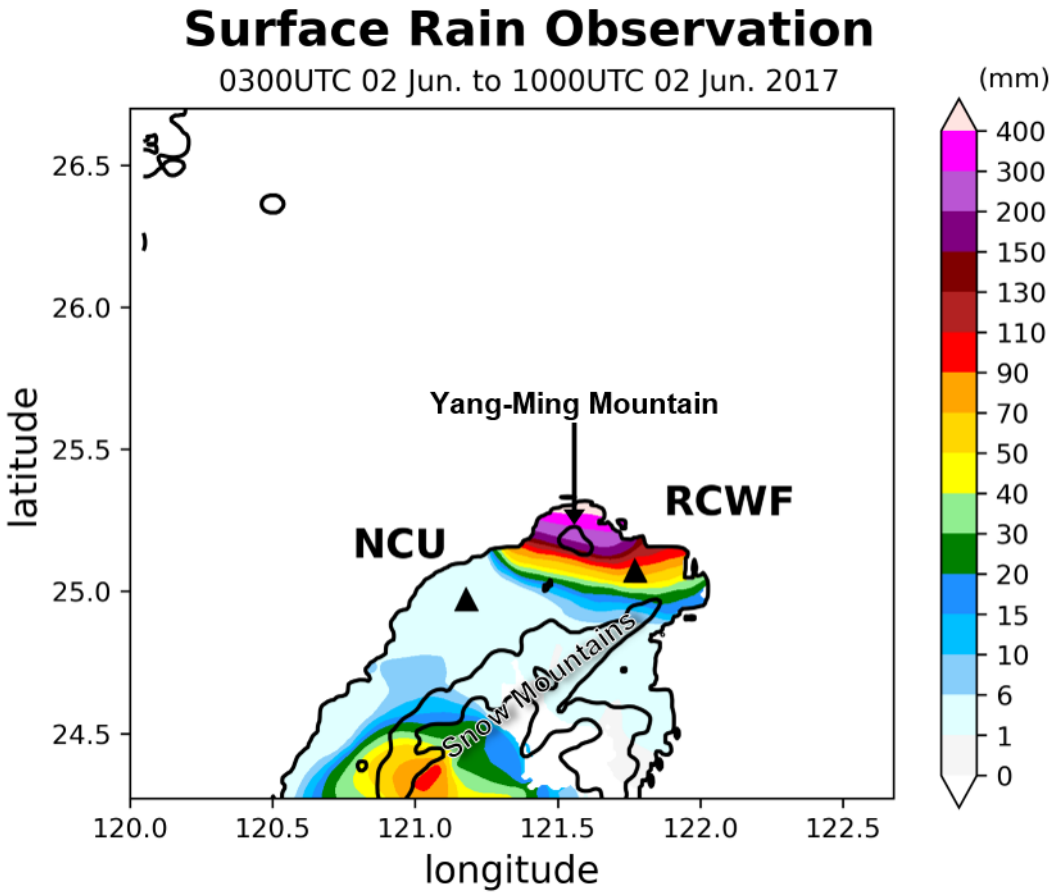
381



382

383 Fig. 5. Composite reflectivity from 1600 UTC 1 June to 0200 UTC 2 June.

384



385

386 Fig. 6. Observed 8-h accumulated rainfall data from ARMTS for 1800 UTC 1 June 2017 to 0200 UTC 2 June
387 2017. The black triangles denote the locations of two radar sites, RCWF and NCU C-Pol.

388

389 ## 4. WISSDOM retrieval of the Mei-Yu case of 1–2 June 2017

390 At 1800 UTC 1 June, three eastward-moving MCSs were aligned in a west–east
391 orientation north of Taiwan (Fig. 4a). Notably, M1 was located northwest of Taiwan, as
392 evidenced by the vertical motion and the wind shear line. At the 1-km level, the prevailing
393 southwesterly wind was over the Taiwan Strait and extended north of Taiwan (Fig. 7a, d, g, j).
394 Over time, M1 moved eastward to the northern edge of Taiwan (Fig. 7a, d). The MBLJ
395 impinges on the Central Mountain Range with orographically induced high pressure at the
396 foothills of the mountains on the windward side (Li and Chen 1998; Yeh and Chen 2003; Tu et
397 al. 2020). This, in turn, drove a downstream wind maximum over northwestern Taiwan (Fig.
398 7d). Furthermore, the Yang-Ming Mountains acted as a barrier, hindering southward movement
399 of the convective line when the shallow cold front was anchored along the northern coast (Tu

et al., 2020). The rapid development of prefrontal convections occurred on the western windward slopes north of the Snow Mountains due to orographic lifting (Fig. 7g). The robust frontal convergence with vertical motion is particularly evident at the frontal position by 1800 UTC (Fig. 7j). In this updraft zone, a downward motion forms between two updrafts, blocks the BJ, and weakens the coupling between the convective system and low-level wind. The downward motion may be due to the cold pool generated by the strong convective activity (Tu et al. 2022).

At 2100 UTC, M2 moves further eastward and merged with M1, which led to pronounced intensification of the vertical velocity over northern Taiwan (Fig. 7b, e, h, k). The BJ in the prefrontal region, coupled with the westerly wind in the north, provides enhanced convergence and rain to the frontal zone over the northwestern coast. The weakened convection, influenced by evaporative cooling effect (Tu et al., 2022), coincides with reorganization of the two MCSs, which led to widespread high wind speeds in northwestern Taiwan around 2100 UTC (Fig. 4b). This phenomenon, in turn, increases the strength and northward tilt of the front in the subsequent hour (not shown). The absence of a sufficiently thick layer of cold air near the surface (Tu et al. 2022) hinders the frontal rainband's progression beyond the northern edge of Taiwan, resulting in persistent heavy rainfall in the region. Furthermore, as highlighted by Ke et al. (2019), prefrontal line convection also strongly contributes to the extreme rainfall. At 0000 UTC 2 June, M3 reaches its mature stage and subsequently dissipates northeastward (Fig. 4d), resulting in a reduction in the strong prefrontal winds inland and a corresponding decrease in vertical motion.

After 0000 UTC 2 June, a new development of a new convective system is observed west of the weakening M4 (Fig. 4d). Fresh convective cells form westward of the frontal zone and propagate eastward, leading to a process of repeated system regeneration. The combined effects of the complex terrain and prefrontal low-level winds reduce the tilt of the frontal convection, thereby enhancing its dynamic force in the vertical direction and leading to an intensification of vertical velocity at 0200 UTC 2 June (Fig. 7c, f, i, l). This observation suggests that as the Mei-Yu front proceeded inland, the associated convection became narrower and stronger. The observational analyses and high-resolution numerical simulations from Tu et al. (2022) and WISSDOM 3-dimensional retrieval winds provide valuable insights into the near-surface and regional evolution of the front and physical processes associated with heavy rainfall, however, a comprehensive understanding of uncertainties on the simulated multiscale interactions for the

Mei-Yu system remains elusive. We use the numerical ensemble simulations to identify the key factors that may contribute to the uncertainties and spread in simulated model results.

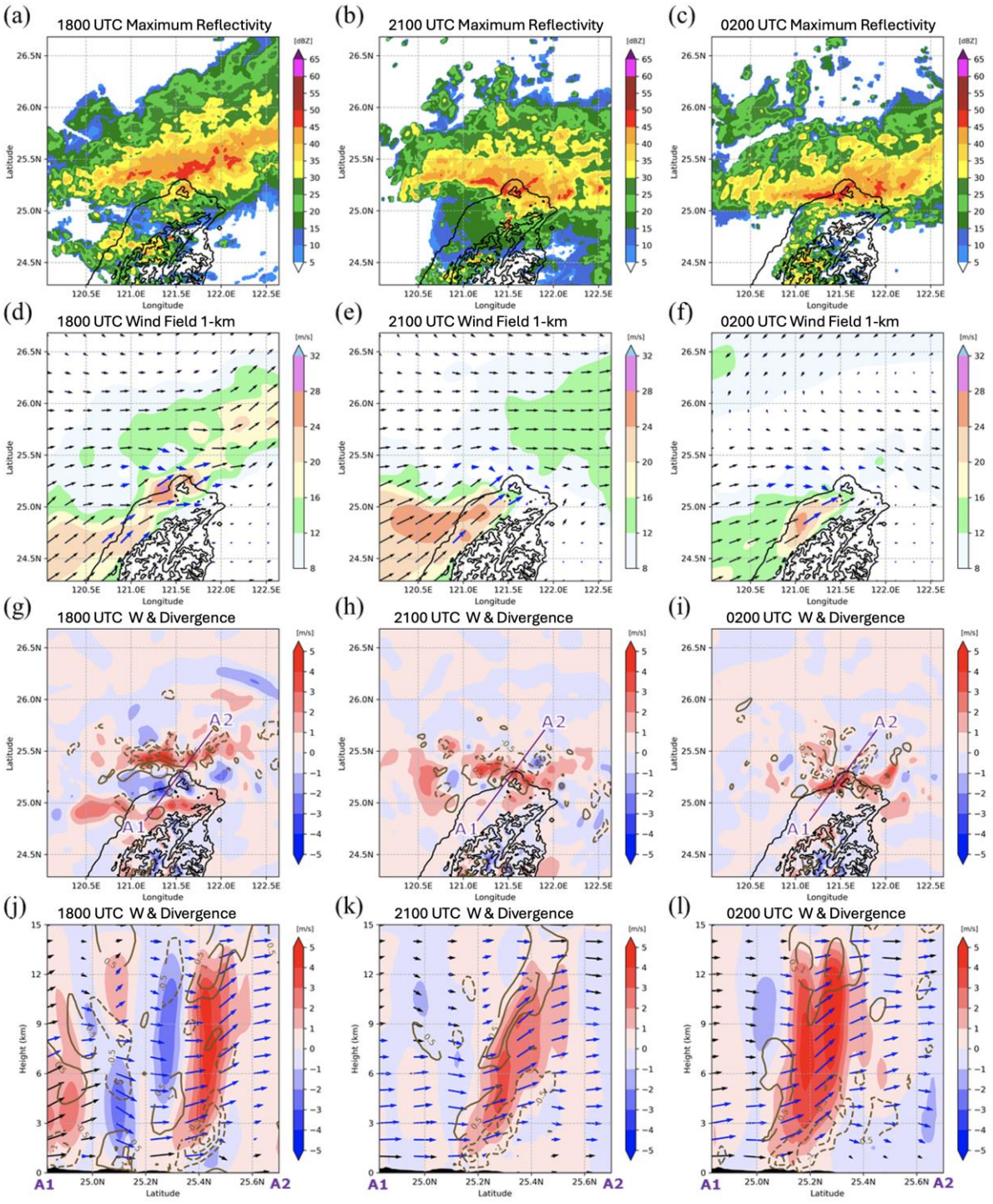


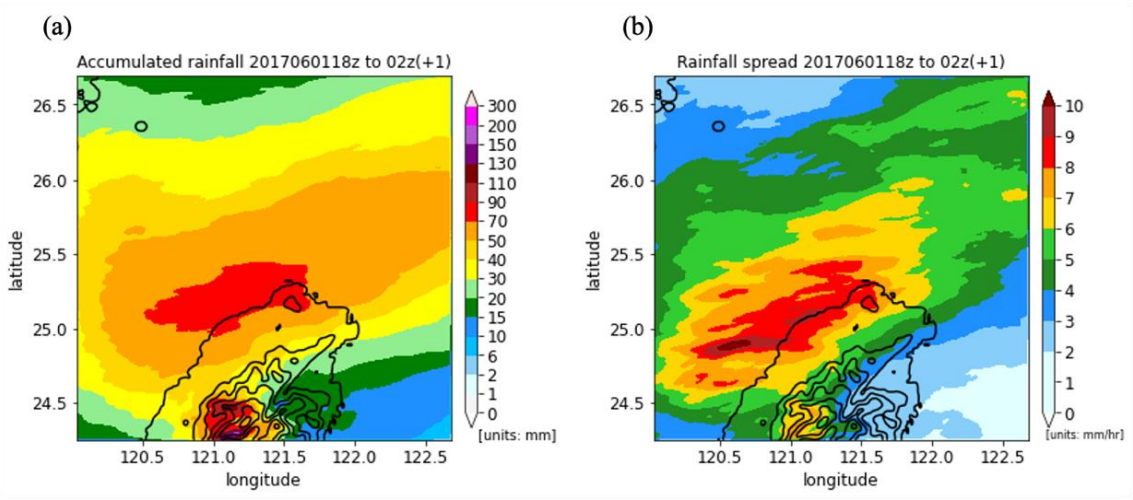
Fig. 7. Retrieved results at 1800 UTC (a, d, g, j) and 2100 UTC (b, e, h, k) 1 June 2017, and at 0200 UTC 2 June 2017 (c, f, i, l). (a)–(c) Composite radar reflectivity from the RCWF and the NCU C-Pol radar. (d)–(f) Horizontal wind direction (vector) and speed (shaded, unit: m s^{-1}) at 1 km. The blue vectors indicate regions

where wind data were derived from radar observations. (g)–(i) Vertical velocity (shaded, unit: m s^{-1}) at 5 km and the divergence region (brown contour, unit: 10^{-3} s^{-1}) at 4.5 km. (j)–(l) Vertical cross section (purple line A1-A2 in g) of vertical velocity (shaded, unit: m s^{-1}), wind direction (vector, blue vector for radar coverage region), and divergence region (brown contour, unit: 10^{-3} s^{-1}).

5. Cluster analysis of ensemble simulations

a. Results of ensemble simulations

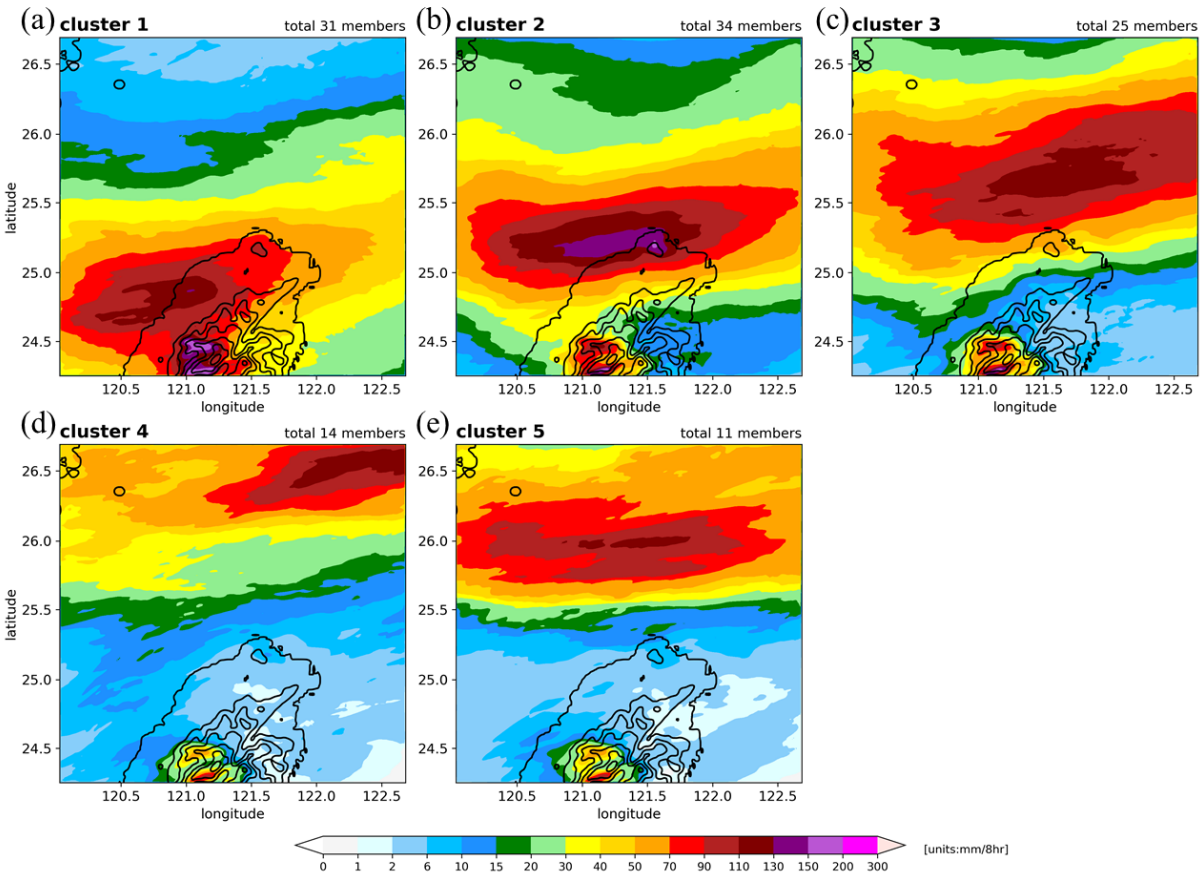
The 1-km-resolution ensemble mean rainfall in domain d04 exhibits a similar pattern to the observations, with the extreme value occurring north of the Yang-Ming Mountains and over the northwestern coast (Fig. 6). The simulated rainfall amount (Fig. 8a), however, was less than observed. This discrepancy is attributed to the ensemble averaging process, which tends to smooth out maximum values of rainfall, leading to lower peak rainfall intensity. Tu et al. (2002) attributes the enhanced rainfall over the northwestern coast to the localized convergence between the BJ and the postfrontal northwesterly flow associated with a frontal cycle (their Figs. 13a, b). The simulated rainfall pattern exhibits considerable spread along the northwest coast of Taiwan, oriented in a north-northeast to south-southwest direction. This spread arises primarily from the large uncertainties in the models associated with the frontal position and low-level winds across the various ensemble members (Fig. 8b). A combined k-means clustering approach is employed that considered both linear and spatial features. This clustering identifies five distinct groups with clusters located in the: (1) inland region; (2) north edge of Taiwan; and (3)–(5) the offshore region north of Taiwan (Fig. 9). The number of members in clusters 1–5 was 31, 34, 25, 14, and 11, respectively. Notably, cluster 2 (Fig. 9) exhibits the closest agreement with the observed rainfall distribution (Fig. 6).



463

464 Fig. 8. The (a) 8-h ensemble mean rainfall and (b) 8-h ensemble spread of rainfall (per hour) during the
465 rainfall period from 1800 UTC 1 June to 0200 UTC 2 June 2017.

466



467

468 Fig. 9. The 8-h ensemble mean rainfall of five clusters during the rainfall period from 1800 UTC 1 June to
469 0200 UTC 2 June 2017. (a) cluster 1; (b) cluster 2; (c) cluster 3; (d) cluster 4; (e) cluster 5.

b. Dynamic characteristic of clustering result

To investigate the pre-frontal rainfall and heavy rainfall characteristics for the different clusters, the 850-hPa cluster mean wind anomalies and cluster mean relative vorticity fields derived from d03 are used (Fig. 10). During the pre-frontal rainfall period, a prevailing southwesterly wind dominates all clusters, with a vortex associated with the low-level frontal mesocyclone positioned north of Taiwan. On the lee side of Taiwan terrain, a lee trough is observed in each cluster (Li and Chen 1998; Tu et al. 2014; Yeh and Chen 2003). In western Taiwan, southwesterly winds ahead of the front lead to the formation of a BJ due to blocking and mass accumulation on the southwestern windward side of Taiwan's terrain. The orographic blocking results in a windward pressure ridge at the foothills with an orographically induced high pressure over the southwestern windward coast. The Lagrangian flow acceleration of the northern branch of the deflected airflow along the northwest coast, known as BJ, is mainly related to the pressure gradient force as the airflow moves down the orographically induced pressure ridge (Yeh and Chen 2003). Using model sensitivity tests, Yeh and Chen (2003) show that the location and strength of the orographically induced low-level strong winds along the northwest coast of Taiwan (e.g., BJ) also depend on the impinging angle and the wind speed of the prevailing winds. Through cluster analysis, we further show that the variations in the MBLJ strength are a primary factor influencing the location of the frontal rainband. During the heavy rainfall phase (Fig. 10b), the distinctions between clusters are evident. Cluster 1 exhibits minimal evidence of a BJ. By contrast, clusters 2 and 3 exhibit clear traces of a BJ along the northwestern coast of Taiwan, which closely resemble the wind fields retrieved by WISSDOM (Fig. 7e). Both clusters 4 and 5 also exhibit a notable BJ presence, although it extended further north, which is consistent with the respective rainfall locations.

Although the prevailing southwesterly flows play a crucial role in moisture transport, the availability of moisture and warm temperatures contribute to "heavy" rainfall events. In this case, under the prevailing southwesterly monsoon flow, clusters numbers 1, 2, and 3 exhibit high equivalent potential temperature exceeding 350 K over Taiwan, indicating a highly moist atmospheric environment (Fig. 11a, b, c). Additionally, the equivalent potential temperature areas of clusters 4 and 5 only affect the offshore region northwest of Taiwan (Fig. 11d, e). This

result suggests that, along with the moisture tongue, thermodynamic conditions may also influence the location of rainfall during this event.

An analysis of geopotential height evolution at 850 and 925 hPa alongside the corresponding rainfall patterns (Fig. 12) reveals that the trough exhibited a northward tilt. This suggests a potential deceleration of the low-level trough due to the influence of the northern Taiwan terrain, leading to gradual alignment with the phase of the 850-hPa trough. The fast-moving trough is associated with peak rainfall accumulations at northern Taiwan for clusters 1, 2, and 3 at approximately 1800 UTC 1 June. These peaks coincide with the trough's proximity to the mountainous region and propagate southward. By contrast, in cluster 4, a stronger Northwest Pacific Subtropical High over southern Taiwan combined with the slowest-moving frontal cyclone results in a larger north-south pressure gradient than for other clustering groups. This configuration resulted in the most northerly positioned rainband of all clusters, with the rainband primarily confined to the north of Taiwan. The eastward movement of the Mei-Yu frontal cyclone corresponds to the large increments in rain rates from 1600 UTC to 2000 UTC, which suggests that the rainband is influenced by the movement of the frontal cyclone.

To gain a better understanding of the interplay among MBLJ, BJ, frontal cyclone evolution and localized convergence, we conducted hodograph analyses for the BJ domain (red box, Fig. 13a) and southwestern domain (SW domain, black box, Fig. 13b). We selected the mean wind values in the NE and SW domains during the 12 hours period surrounding the event, with an average centered around 0600 UTC from 0500 UTC 1 June to 0700 UTC 1 June, as well as the mean wind values in the BJ domain during the rainfall period from 1800 UTC 1 June to 0200 UTC 2 June, for hodograph analysis. These values were derived from the five cluster-averaged wind patterns.

An analysis of the MBLJ region (SW domain) reveals a strengthening trend in all five clusters, with wind speeds exceeding 15 m s^{-1} before 0800 UTC. Two groups emerge within these clusters: group 1, which comprised clusters 1, 2, and 3; and group 2, which comprised clusters 4 and 5. Around 1800 UTC, all clusters reach their maximum meridional wind when the frontal cyclone was positioned offshore from South China and subsequently reorganized north of Taiwan. As the frontal cyclone progresses eastward, the pressure gradient weakens, leading to a decrease in meridional wind compared with the prefrontal period. However, the zonal component remains strong. In the SW domain, particularly near the core region of the

MBLJ, cluster 5 exhibits a distinct wind pattern. After 0800 UTC 1 June, winds in this cluster strengthen to over 15 m s^{-1} and turn counterclockwise from west-southwesterlies to southwesterlies. Consequently, the increase in northeasterly flow in the north and decrease in meridional wind in the south result in the faster southward propagation of the frontal system after the analysis period.

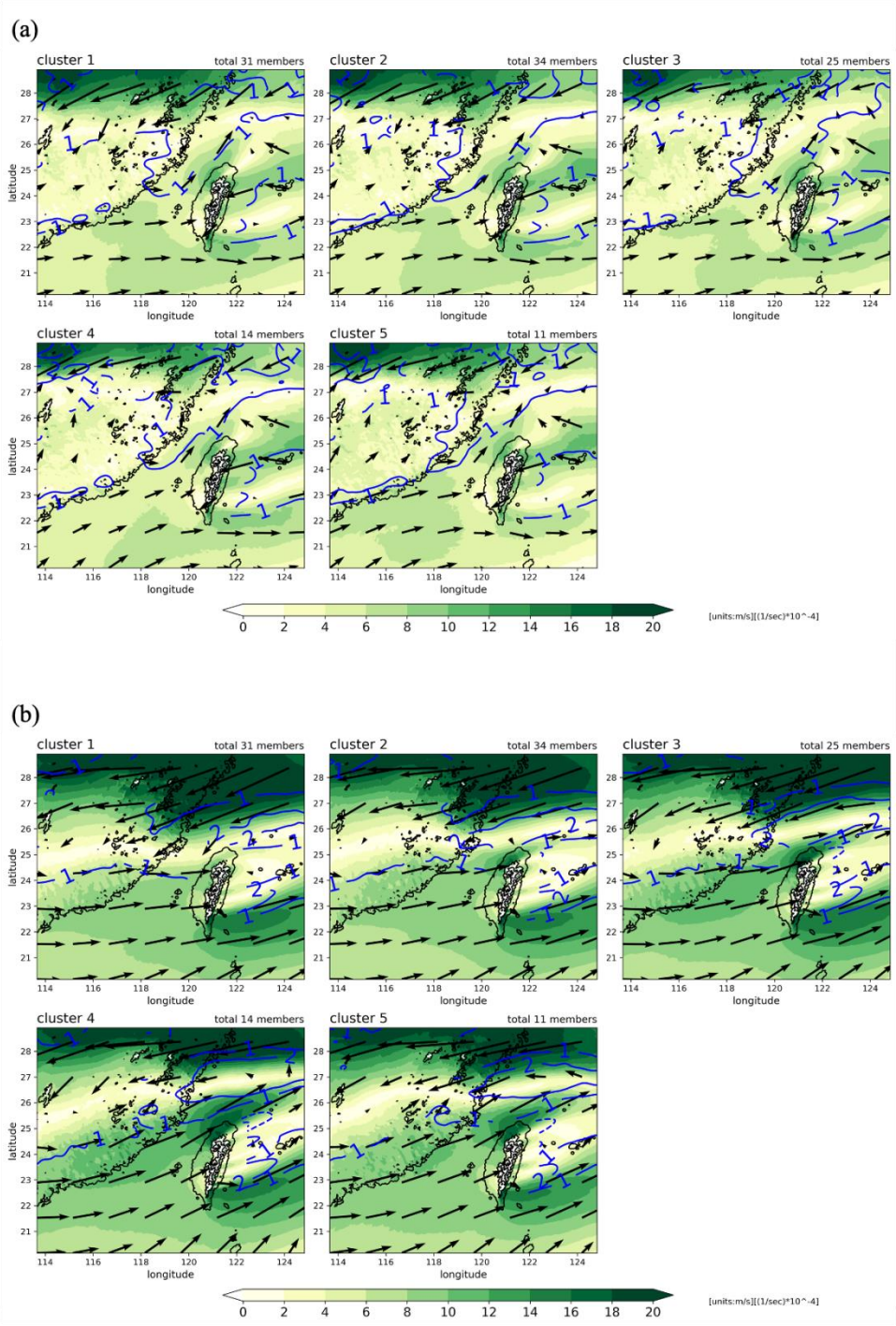
In the BJ domain (Fig. 13b), the initial wind pattern is dominated by a prevailing southwesterly flow across the Taiwan Strait. However, as wind speeds strengthen in southwestern Taiwan (Fig. 13c), a corresponding increase in wind speed with a slight directional change is observed along the northern edge of the Snow Mountains (not shown). After 1000 UTC, the passage of a frontal cyclone and the advancing wind shear line affect the five wind patterns, leading to an increase in the zonal wind component. As the wind shear line reaches northern Taiwan, a growing number of ensemble members are affected by the postfrontal wind, resulting in a reduction of the meridional wind in each cluster. After 1800 UTC, cluster 1 exhibits a significant change in wind direction, with a marked decrease in the southerly wind component (Fig. 13d). This shift is consistent with the fact that cluster 1 has the southernmost ensemble-averaged rainfall location among all clusters. Compared with cluster 1, clusters 2 and 3 experience a smaller reduction in the southerly wind component and maintain stronger westerly winds. Notably, cluster 4, which has the northernmost rainfall location, exhibits the least decrease in southerly wind after 1400 UTC. This finding indicates a strong correlation between wind direction and wind speed in the BJ domain and the rainfall location.

Figure 14 shows the latitude–time section of 8-h accumulated rainfall for each cluster, spanning latitudes from 24.38°N to 27.93°N. The rainfall patterns are primarily influenced by the evolution of the frontal cyclone, modulated by both wind patterns and pressure gradients. The rainfall peaks for clusters 1, 2, and 3 move southward over time, with slight variation in timing that consistently correspond to the position of the frontal cyclone. Conversely, for clusters 4 and 5, the rainfall peaks remained primarily north of Taiwan due to the slower and more northern eastward-moving Mei-Yu frontal cyclone, eventually culminating in maximum rainfall around 26.5°N. Examination of the model probability in ensembles reveals that the simulated members in each cluster initially had similar hourly rainfall features, however, these ensembles diverged as the model integration progressed (Fig. 14b). Previous studies established a positive relationship between the strength of the BJ and rainfall amount, attributable to localized convergence induced by the jet (Chen et al. 2005; Chen et al. 2018; Ke et al. 2019; Li

563 and Chen 1998; Li et al. 1997). The prefrontal jet intensity is known to be closely related to
564 convergence at the interface between air masses. In the temporal diagrams of the heavy rainfall
565 event, the probability distribution of rainfall captured by the ensemble clusters indicates high
566 probability of heavy rainfall across all groups (Fig. 14b). This suggests a considerable degree
567 of consistency in rainfall characteristics between the cluster ensembles. Notably, the cluster 2
568 ensembles demonstrate remarkable ability to effectively capture the probability of heavy
569 rainfall.

570

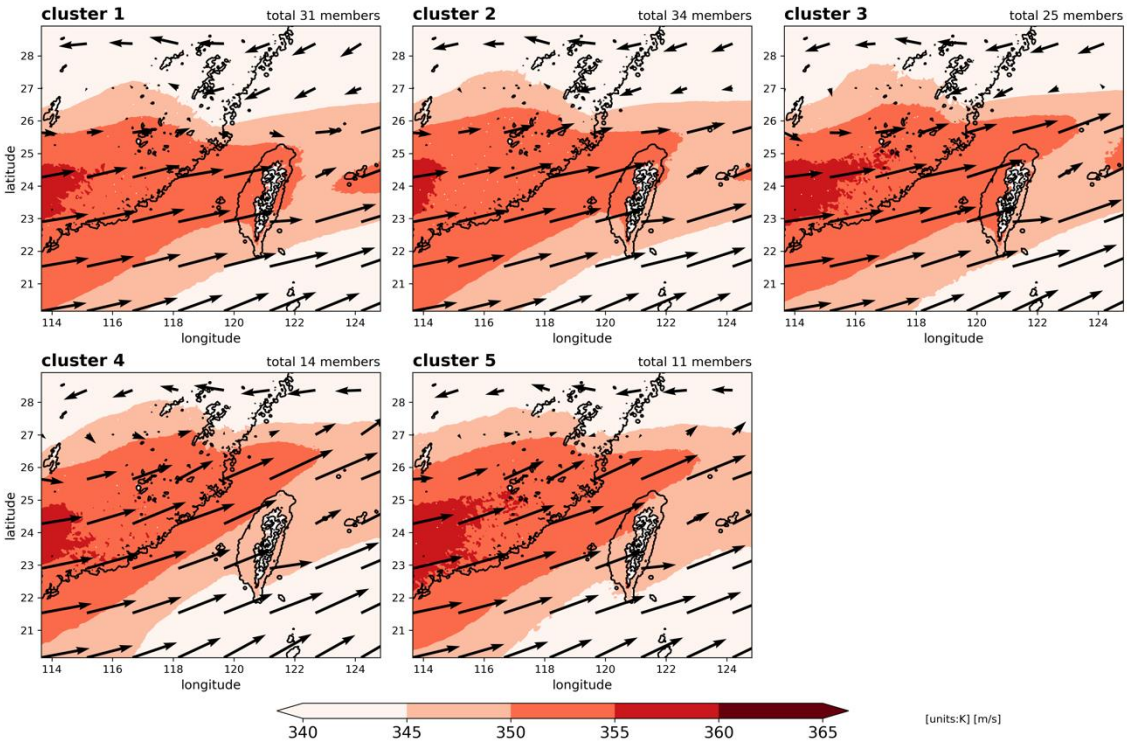
571



572

573 Fig. 10. The 850-hPa mean anomaly winds (shaded, unit: m s^{-1}; vector) and mean relative vorticity (blue
574 contour, unit: s^{-1}) of five identified clusters during the (a) prefrontal rainfall period (from 0500 UTC 1 June
575 to 0700 UTC 1 June) and (b) heavy rainfall period (from 1800 UTC 1 June to 0200 UTC 2 June).

576

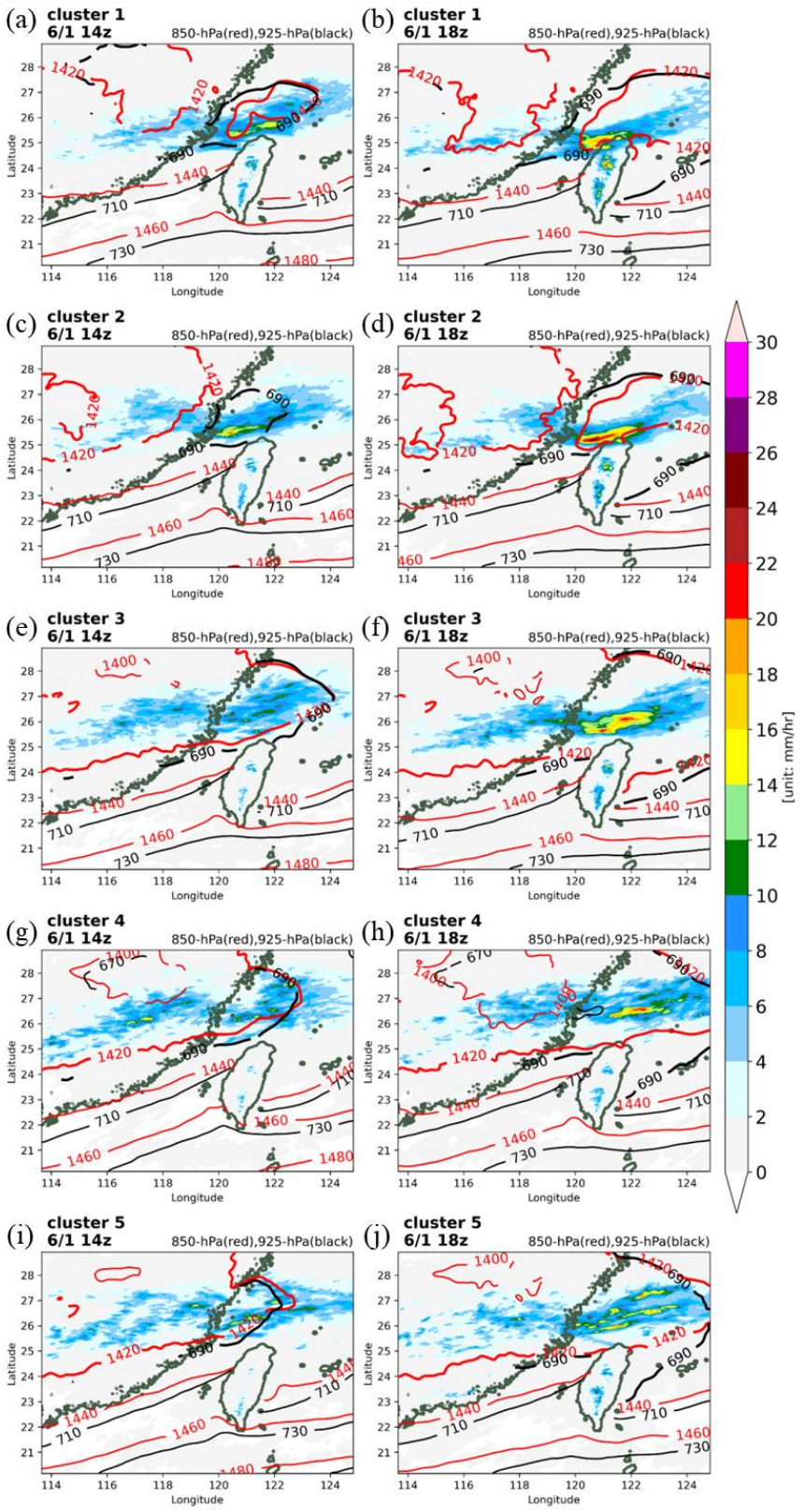


577

578 Fig. 11. Mean equivalent potential temperature (shaded) and 850-hPa mean wind (vector) during the rainfall
579 period for each k-means cluster, from 1800 UTC 1 June to 0200 UTC 2 June 2017.

580

581



582

583 Fig. 12. Simulated clustered hourly rainfall pattern (shaded, unit: mm h^{-1}) and geopotential height (contoured,
584 unit: gpm, interval: 10 gpm) at 850 hPa (red contour line) and 925 hPa (black contour line) from 1400 UTC
585 (left column) to 1800 UTC (right column) 1 June 2017.

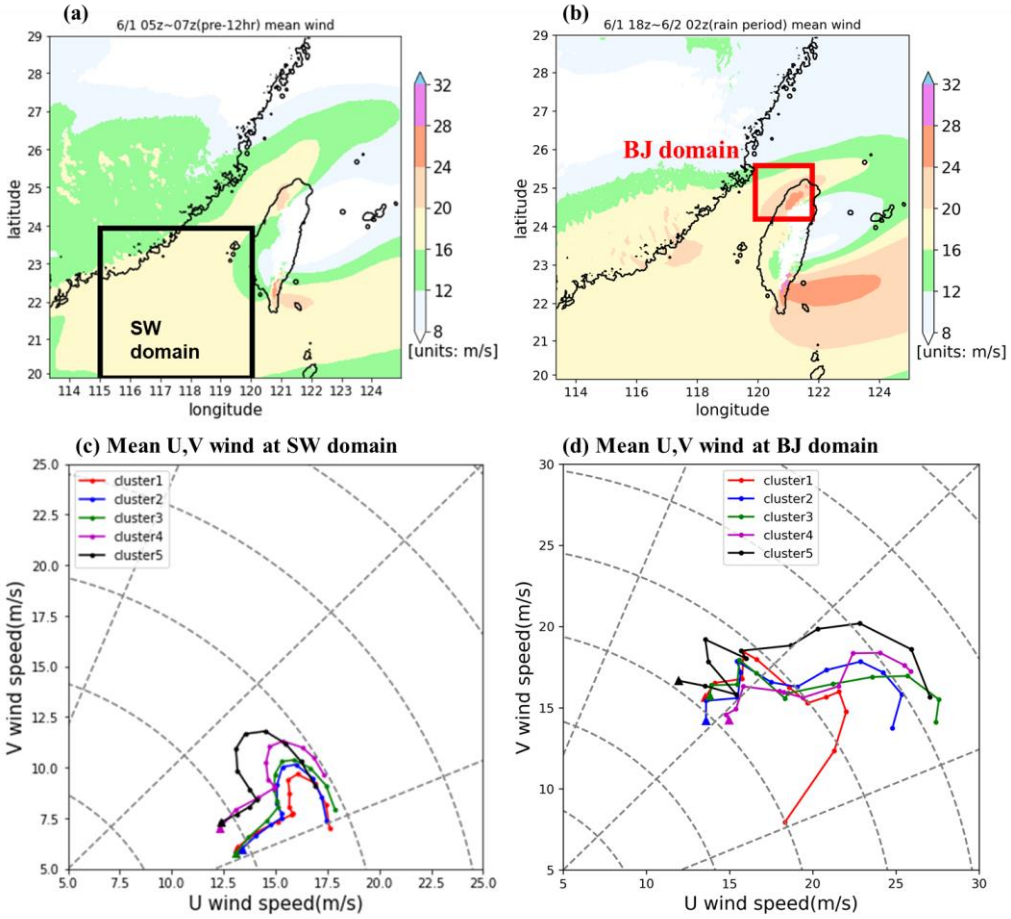
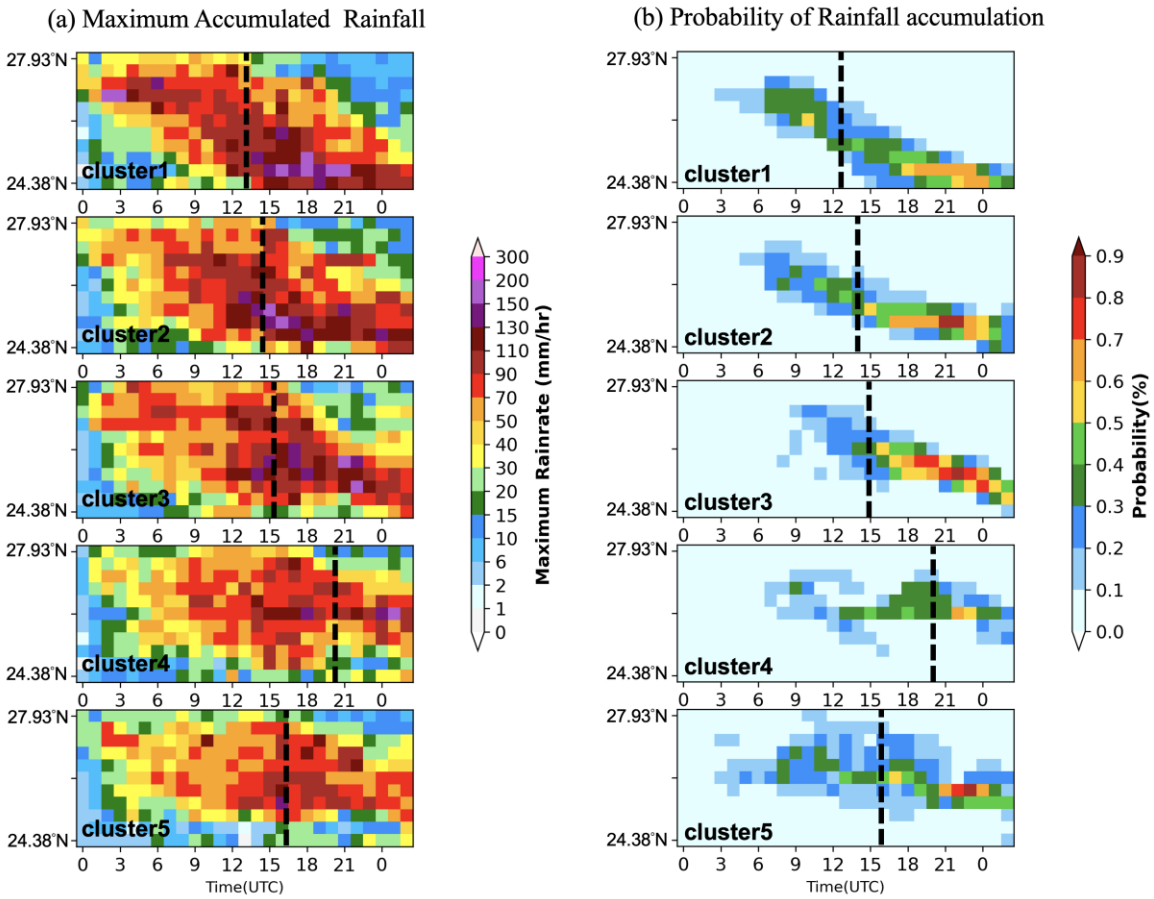


Fig. 13. (a) Mean wind field at 925 hPa during the 12 h period surrounding the event, with an average centered around 0600 UTC (from 0500 UTC 1 June to 0700 UTC 1 June). The black square denotes the SW domain. (b) Mean wind field at 925 hPa during the rainfall period (from 1800 UTC 1 June to 0200 UTC 2 June). The red square denotes the BJ domain. (c) Cluster-averaged hodographs of the simulated 850-hPa mean wind in the SW domain from 0200 UTC 1 June to 0000 UTC 2 June (dots; interval: 2 h), with the first time (0200 UTC 1 June) marked by a triangle. (d) Same as panel (c) but for the BJ domain.



596

597 Fig. 14. Time series of (a) maximum accumulated rainfall over 8 h and (b) the probability of hourly rainfall
598 accumulation exceeding 10 mm for 24.38°N to 27.93°N. Figures in each panel correspond to clusters 1 (top)
599 to 5 (bottom). The dashed line indicates the period when the frontal cyclone affected the region.

600

601 ***c. Schematic of two types of rainfall center***

602 The trajectory of the low-level low-pressure system is a critical factor of the frontal system,
603 influencing not only the spatial distribution but also the temporal evolution of heavy rainfall in
604 the investigated event. The pressure gradient force induces meridional acceleration downstream
605 of the pressure ridge along the western coast. As the frontal cyclone progresses eastward
606 downstream of the jet, the southerly wind intensifies on the eastern flank of the frontal cyclone.
607 The preceding analysis of mid-to-low-level frontal cyclone, wind patterns, and their
608 associations with rainfall characteristics reveal two distinct types of frontal cyclone movements
609 and their corresponding rainfall patterns in the five clusters.

To further illustrate these distinct categories, we examine a representative member from each type (Member 106 from cluster 2 and Member 29 from cluster 4; Fig. 15). The Hovmöller diagram (Fig. 15) depicting wind and rainfall patterns at 121°E reveals a clear correlation between rainfall distribution and frontal cyclone movement. The initial rainfall pattern aligns with the frontal cyclone and gradually shifts southward until 1500 UTC. Subsequently, it accumulates north of the topography before 0000 UTC 2 June. This behavior is reminiscent of observations classified as Type A. This scenario is characterized by a deceleration of the frontal cyclone and enhanced convergence, attributed to terrain blockage in the north. Consequently, a notable increase in accumulated rainfall and stronger prefrontal winds are noted compared with the preceding period. With the arrival of postfrontal northerly winds, the rainband strengthens and moves southward. In contrast, the Type-B scenario involves rainfall primarily located south of the frontal cyclone, albeit with less precipitation than that observed in the Type-A scenario.

The preceding analysis allows for the classification of the heavy rainfall event in northern Taiwan into two distinct types. These mechanisms are illustrated in the schematic presented in Fig. 16. In Type-A scenarios (Fig. 16a), the MBLJ exhibits a weaker southerly component in the prevailing southwestern airflow. This airflow approaches the Central Mountain Range in southwestern Taiwan at a greater angle to the north-south orientation of the terrain. The low-level frontal cyclone passes the northern edge of Taiwan at approximately 26°N, leading to the rainband largely stalling at the convergence region between the southern edge of the low and the core region of the BJ. Due to the eastward propagation of the low-pressure system, the jet intensifies to the east of the low. As the frontal cyclone propagates faster, the rainband moves southward earlier, as observed in cluster 2. In Type-B scenarios (Fig. 16b), a stronger meridional component prevails from the west to north of Taiwan. The connection with the southerly from the east of the low-level trough would lead to an eastward-moving rather than southward-moving trough. As the low-pressure system moves away from the terrain, the Marine Boundary Layer Jet (MBLJ+) exhibits a pronounced southerly wind particularly as the rainband's location shifts further away from northern Taiwan.

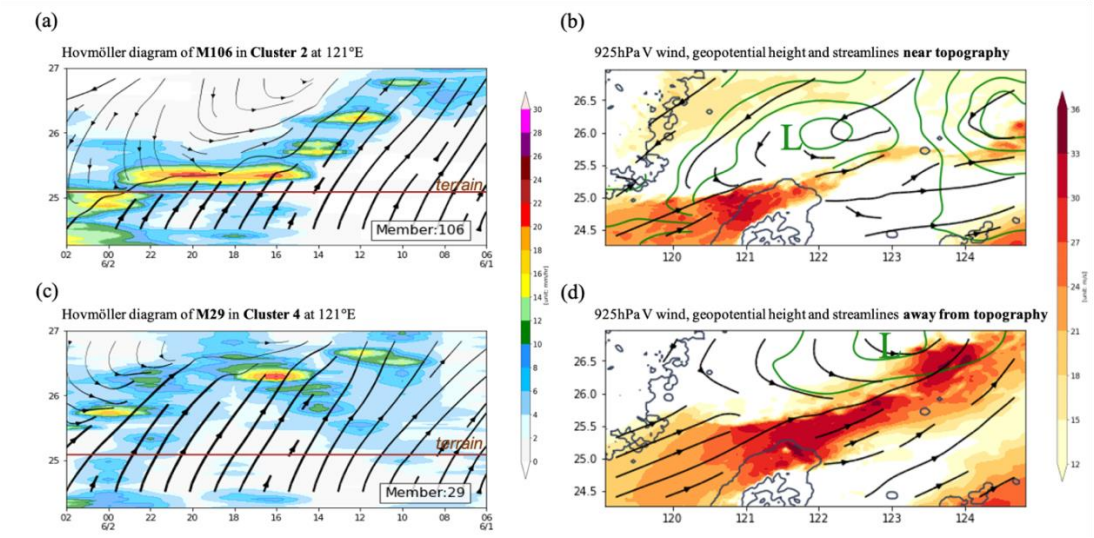


Fig. 15. (a) and (c) Hovmöller diagrams for Member 106 and Member 29 at 121°E depicting simulated streamline patterns and rainfall distribution from 0600 UTC 1 June. (b) and (d) Schematic of the correlation between the low-level low, meridional wind, and frontal system. (b) For the Type-A scenario, the low-level low is located at approximately 26°N with the barrier jet located at the prefrontal region. (d) For Type-B, the low-level low is located at approximately 26.5°N, far from the topography, with the barrier jet prevailing from south to east of the low.

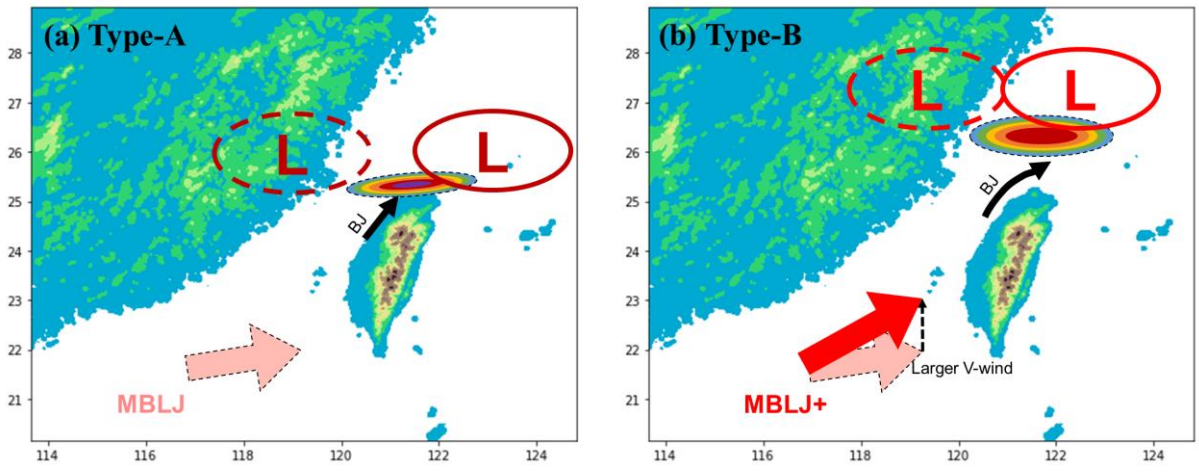


Fig. 16. Schematic of the correlation between the frontal cyclone (red L), marine boundary layer jet (MBLJ, pink vector), barrier jet (BJ, black vector) and frontal heavy rain area (blue-to-red shaded). (a) Type-A: The frontal heavy rainfall located northern Taiwan, depicting the interaction between the MBLJ, BJ, and frontal cyclone near northern Taiwan. (b) Type-B: The frontal heavy rainfall located in the offshore area north of Taiwan. The frontal cyclone forms away from northern Taiwan. The red arrow represents MBLJ+, which has a more pronounced southerly component compared to MBLJ. The barrier jet extends further north to the southern edge of the heavy rainfall area.

6. Summary

During 1–2 June 2017, a shallow east–west orientated front stagnated over northern Taiwan, resulting in extreme rainfall—more than 400 mm within 8 h along the northern coast—under abundant moisture and strong winds from both the MBLJ and BJ. This study explores the uncertainties from an ensemble perspective, analyzing the possible outcomes and category differences, rather than simply a deterministic simulation. We made a detailed comparison between the simulated results from ensemble members and observed data, including satellite and radar reflectivity data and three-dimensional wind fields derived from the Wind Synthesis System using Doppler Measurements (WISSDOM).

By employing spatially correlated k-means clustering on the rainfall distributions from 128 WRF model ensemble members, this study effectively categorized the simulated heavy rainfall event into five groups. This approach provides a comprehensive understanding of the factors influencing the model uncertainties for this heavy rainfall event, especially the relationship among jets, cyclones, and heavy rainfall.

1) Successive MCSs and Squall Line Transition**:** Three eastward-moving mesoscale convective systems (M1, M2, M3) contributed to heavy rainfall over northern Taiwan. M1 initiated strong convergence near the coast at 1800 UTC 1 June, creating a downstream wind maximum and weakening the barrier jet. By 2100 UTC, M2 merged with M1, further enhancing vertical velocity and intensifying rainfall. During this time, radar showed a squall line transitioning from a trailing stratiform (TS-type) to a parallel stratiform (PS-type) pattern, reflecting prefrontal line convection effects. After M3 dissipated northeastward by 0000 UTC 2 June, rainfall and prefrontal winds decreased inland.
2) Distinct Rainfall Patterns by Clustering: Spatial k-means clustering divided 128 ensemble members into five rainfall pattern clusters, each associated with unique wind dynamics and rainfall locations. Clusters showed rainfall concentrated over the ocean, inland, or along specific coastal regions, with variations tied to meridional wind strength and interaction with topography.
3) Wind-Driven Dynamics: Analysis showed a relationship between frontal cyclone proximity and rainfall distribution, with close frontal cyclone positioning leading to

stronger southwest meridional winds and sustained frontal convection. Post-cyclone passage, the southwest wind's meridional component weakened and shifted direction, aligning with shifts in rainfall intensity.

4) Two Rainfall Distribution Types: Two primary rainfall distributions emerged. Type A, similar to the observed event, produced heavy inland rainfall in northern Taiwan, driven by a close low-pressure center and southwesterly MBLJ winds. Type B showed offshore rainfall, influenced by a more distant low-pressure center and stronger southerly MBLJ winds, extending barrier jets further north.

This study underscores the importance of using ensemble simulations to understand the variability in extreme weather patterns, especially for mesoscale events like heavy rainfall. Unlike traditional deterministic simulations, which often overlook the range of possible outcomes and the inherent uncertainties in weather modeling, ensemble simulations reveal how small perturbations in initial conditions can lead to significant differences in rainfall distribution and intensity. By applying spatial clustering to group ensemble members based on rainfall patterns, this study highlights how variations in atmospheric dynamics—such as the marine boundary layer jet (MBLJ), frontal cyclone positioning, and barrier jet (BJ) intensity—impact rainfall outcomes across different scenarios.

The ensemble-based framework, integrated with clustering techniques, enhances the understanding of mesoscale weather phenomena by highlighting the probabilistic nature of rainfall distributions and intensities. This method effectively captures environmental differences and provides insights into the likelihood of various convective rainfall patterns, aiding in forecast accuracy and disaster preparedness. It offers a robust approach for improving extreme precipitation forecasts through four key applications: 1) Improving nowcasting by quickly identifying the most likely cluster for ongoing events, enabling timely and precise warnings; 2) Enhancing probabilistic forecasts by leveraging cluster-based rainfall distributions to deliver more reliable outcomes for decision-making; 3) Supporting targeted mitigation strategies by identifying regions at risk of heavy rainfall, facilitating early warning systems and infrastructure enhancements; and 4) Evaluating and improving numerical weather prediction models by identifying ensemble performance biases and uncertainties, guiding model refinement. This adaptable methodology provides a comprehensive foundation for advancing atmospheric research and improving weather forecasting across diverse meteorological scenarios and regions.

Acknowledgments.

This work was supported by the National Science and Technology Council (NSTC) of Taiwan through Research Grant NSTC 113-2111-M-008-030. We thank to National Center for High-performance Computing (NCHC) of National Applied Research Laboratories (NARLabs) in Taiwan for providing computational and storage resources.

Data Availability Statement.

The reanalysis data used in this study were obtained from NCEP-FNL (https://doi.org/10.5065/D6M043C6). The WRF model employed in this study can be accessed at the Github repository (GitHub - wrf-model/WRF at release-v4.2.1). The observational datasets used in this research are obtained from the Atmospheric Science Research and Application Databank (ASRAD; https://asrad.pccu.edu.tw) funded and maintained by the National Science and Technology Council and Chinese Culture University of Taiwan, respectively. The codes of WISSDOM for retrieving three-dimensional wind can be obtained by request to the co-author Dr. Yu-Chieng Liou (tyliou@atm.ncu.edu.tw). The numerical model setups and experimental results are available from the corresponding author Dr. Kao-Shen Chung (kschung@atm.ncu.edu.tw). The clustering results presented in this study were obtained using Python's (version 3.8.6) k-means clustering algorithm, available through the scikit-learn library (version 0.23.2) with detailed documentation available at https://scikit-learn.org.

REFERENCE

Barker, D. M., W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results. *Monthly Weather Review*, **132,** 897-914.

Buizza, R., P. L. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, 2005: A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. *Mon. Wea. Rev.*, **133,** 1076–1097.

Chen, C.-S., W.-C. Chen, Y.-L. Chen, P.-L. Lin, and H.-C. Lai, 2005: Investigation of orographic effects on two heavy rainfall events over southwestern Taiwan during the Mei-yu season. *Atmospheric Research*, **73,** 101-130.

Chen, C.-Y., Y.-L. Chen, C.-S. Chen, P.-L. Lin, and C.-L. Liu, 2013: Revisiting the Heavy Rainfall Event over Northern Taiwan on 3 June 1984. *Terrestrial, Atmospheric and Oceanic Sciences*, **24**.

Chen, Y.-L., and J. Li, 1995a: Characteristics of Surface Airflow and Pressure Patterns over the Island of Taiwan during TAMEX. *Monthly Weather Review*, **123,** 695-716.

——, 1995b: Large-Scale Conditions Favorable for the Development of Heavy Rainfall during TAMEX IOP 3. *Monthly Weather Review*, **123,** 2978–3002.

Chen, Y.-L., and C.-C. Tu, 2024: The Mei-Yu Season Over the Taiwan Area. *Systems and Environmental Sciences*.

Chen, Y.-L., X. A. Chen, S. Chen, and Y.-H. Kuo, 1997: A Numerical Study of the Low-Level Jet during TAMEX IOP 5. *Monthly Weather Review*, **125,** 2583-2604.

Chen, Y.-L., Y.-J. Chu, C.-S. Chen, C.-C. Tu, J.-H. Teng, and P.-L. Lin, 2018: Analysis and Simulations of a Heavy Rainfall Event over Northern Taiwan during 11–12 June 2012. *Monthly Weather Review*, **146,** 2697-2715.

Chen, Y.-L., C.-C. Tu, F. Hsiao, C.-S. Chen, P.-L. Lin, and P.-H. Lin, 2022: An Overview of Low-Level Jets (LLJs) and Their Roles in Heavy Rainfall over the Taiwan Area during the Early Summer Rainy Season. *Meteorology*, **1,** 64-112.

764 Cornejo, I. C., A. K. Rowe, K. L. Rasmussen, and J. C. DeHart, 2024: Orographic Controls on
765 Extreme Precipitation Associated with a Mei-Yu Front. *Monthly Weather Review*, **152,**
766 531-551.

767 Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon
768 Experiment using a mesoscale two–dimensional model. *Journal of the atmospheric*
769 *sciences*, **46,** 3077–3107.

770 Gochis, D. J., and Coauthors, 2017: The Weather Research and Forecasting Model: Overview,
771 System Efforts, and Future Directions. *B Am Meteorol Soc*, **98,** 1717-1737.

772 Grell, G. A., and D. Dévényi, 2002: A generalized approach to parameterizing convection
773 combining ensemble and data assimilation techniques. *Geophysical Research Letters*, **29,**
774 38-31-38-34.

775 Hong, S., and J. J. Lim, 2006: The WRF Single-Moment 6-Class Microphysics Scheme
776 (WSM6). *Asia-pacific Journal of Atmospheric Sciences*, **42,** 129-151.

777 Houze, R. A., 2004: Mesoscale convective systems. *Reviews of Geophysics*, **42**.

778 Ke, C.-Y., K.-S. Chung, T.-C. Chen Wang, and Y.-C. Liou, 2019: Analysis of heavy rainfall and
779 barrier-jet evolution during Mei-Yu season using multiple Doppler radar retrievals: a case
780 study on 11 June 2012. *Tellus A: Dynamic Meteorology and Oceanography*, **71,** 1-21.

781 Kerns, B. W. J., Y.-L. Chen, and M.-Y. Chang, 2010: The Diurnal Cycle of Winds, Rain, and
782 Clouds over Taiwan during the Mei-Yu, Summer, and Autumn Rainfall Regimes.
783 *Monthly Weather Review*, **138,** 497-516.

784 Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *Journal of Computational*
785 *Physics*, **227,** 3515-3539.

786 Li, J., and Y. L. Chen, 1998: Barrier jets during TAMEX. *Monthly Weather Review*, **126,** 959-
787 971.

788 Li, J., Y. L. Chen, and W. C. Lee, 1997: Analysis of a heavy rainfall event during TAMEX.
789 *Monthly Weather Review*, **125,** 1060-1082.

Lin, P.-L., Y.-L. Chen, C.-S. Chen, C.-L. Liu, and C.-Y. Chen, 2011: Numerical experiments investigating the orographic effects on a heavy rainfall event over the northwestern coast of Taiwan during TAMEX IOP 13. *Meteorology and Atmospheric Physics*, **114,** 35-50.

Liou, Y.-C., and Y.-J. Chang, 2009: A Variational Multiple–Doppler Radar Three-Dimensional Wind Synthesis Method and Its Impacts on Thermodynamic Retrieval. *Monthly Weather Review*, **137,** 3992-4010.

Liou, Y.-C., S.-F. Chang, and J. Sun, 2012: An Application of the Immersed Boundary Method for Recovering the Three-Dimensional Wind Fields over Complex Terrain Using Multiple-Doppler Radar Data. *Monthly Weather Review*, **140,** 1603-1619.

Liou, Y.-C., T.-C. Chen Wang, and P.-Y. Huang, 2016: The Inland Eyewall Reintensification of Typhoon Fanapi (2010) Documented from an Observational Perspective Using Multiple-Doppler Radar and Surface Measurements. *Monthly Weather Review*, **144,** 241-261.

MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, **1,** 281-297.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, **102,** 16663-16682.

Parker, M. D., and R. H. Johnson, 2000: Organizational Modes of Midlatitude Mesoscale Convective Systems. *Monthly Weather Review*, **128,** 3413-3436.

Skamarock, W. C., and Coauthors, 2019: A Description of the Advanced Research WRF Model Version 4**,** NCAR Tech. Note NCAR/TN-556+STR 145 pp.

Tao, W.-K., J. Simpson, and M. McCumber, 1989: An Ice-Water Saturation Adjustment. *Mon. Wea. Rev.*, **117,** 231–235.

Tao, W. K., D. Wu, S. Lang, J. D. Chern, C. Peters-Lidard, A. Fridlind, and T. Matsui, 2016: High-resolution NU-WRF simulations of a deep convective-precipitation system during MC3E: Further improvements and comparisons between Goddard microphysics schemes and observations. *Journal of Geophysical Research: Atmospheres*, **121,** 1278-1305.

819 Teng, H.-F., C.-S. Lee, H.-H. Hsu, J. M. Done, and G. J. Holland, 2019: Tropical Cloud Cluster
820 Environments and Their Importance for Tropical Cyclone Formation. *Journal of Climate*,
821 **32,** 4069-4088.

822 Tseng, Y.-H., and J. H. Ferziger, 2003: A ghost-cell immersed boundary method for flow in
823 complex geometry. *Journal of Computational Physics*, **192,** 593-623.

824 Tu, C.-C., Y.-L. Chen, P.-L. Lin, and Y. Du, 2019: Characteristics of the Marine Boundary Layer
825 Jet over the South China Sea during the Early Summer Rainy Season of Taiwan. *Monthly*
826 *Weather Review*, **147,** 457-475.

827 Tu, C.-C., Y.-L. Chen, P.-L. Lin, and P.-H. Lin, 2020: The relationship between the boundary
828 layer moisture transport from the South China Sea and heavy rainfall over Taiwan.
829 *Terrestrial, Atmospheric and Oceanic Sciences*, **31,** 159-176.

830 Tu, C.-C., Y.-L. Chen, P.-L. Lin, and M.-Q. Huang, 2022: Analysis and Simulations of a Heavy
831 Rainfall Event Associated with the Passage of a Shallow Front over Northern Taiwan on
832 2 June 2017. *Monthly Weather Review*, **150,** 505-528.

833 Tu, C.-C., Y.-L. Chen, C.-S. Chen, P.-L. Lin, and P.-H. Lin, 2014: A Comparison of Two Heavy
834 Rainfall Events during the Terrain-Influenced Monsoon Rainfall Experiment (TiMREX)
835 2008. *Monthly Weather Review*, **142,** 2436-2463.

836 Wang, C.-C., M.-S. Li, C.-S. Chang, P.-Y. Chuang, S.-H. Chen, and K. Tsuboki, 2021:
837 Ensemble-based sensitivity analysis and predictability of an extreme rainfall event over
838 northern Taiwan in the Mei-yu season: The 2 June 2017 case. *Atmospheric Research*,
839 **259**.

840 Xue, M., X. Wang, A. Johnson, and F. Kong, 2011: Hierarchical Cluster Analysis of a
841 Convection-Allowing Ensemble during the Hazardous Weather Testbed 2009 Spring
842 Experiment. Part II: Ensemble Clustering over the Whole Experiment Period. *Monthly*
843 *Weather Review*, **139,** 3694-3710.

844 Yeh, H.-C., and Y.-L. Chen, 2003: Numerical Simulations of the Barrier Jet over Northwestern
845 Taiwan during the Mei-Yu Season. *Monthly Weather Review*, **131,** 1396-1407.

Zheng, M., E. K. M. Chang, and B. A. Colle, 2019: Evaluating U.S. East Coast Winter Storms in a Multimodel Ensemble Using EOF and Clustering Approaches. *Monthly Weather Review*, **147,** 1967-1987.

Zheng, M., E. K. M. Chang, B. A. Colle, Y. Luo, and Y. Zhu, 2017: Applying Fuzzy Clustering to a Multimodel Ensemble for U.S. East Coast Winter Storms: Scenario Identification and Forecast Verification. *Weather and Forecasting*, **32,** 881-903.