Theoretical and Applied Climatology
https://doi.org/10.1007/s00704-023-04603-0

RESEARCH

Comparative study on the characteristics of rainfall simulation over South Korea by summertime weather patterns according to the use of cumulus parameterization

Ga-Yeong Seo[1] · Joong-Bae Ahn[2]

Received: 28 April 2023 / Accepted: 14 August 2023

Abstract
This study identifies the characteristics of summertime extreme rainfall events (ERE) according to the weather patterns in South Korea and compares the extreme rainfall model simulation skills depending on the use of cumulus parameterization. The Weather Research and Forecasting (WRF) model simulations with and without the Kain–Fritsch cumulus scheme (CU_ON and CU_OFF experiment, respectively) are conducted over a 3 km high-resolution domain. The ERE-occurring days are clustered into four representative weather patterns (northern cyclonic circulation, frontal pattern between low and high, southwestern extratropical cyclone, and dominant positive geopotential height patterns) in terms of the 850 hPa geopotential height anomaly. As the occurrence dates of observed Clusters 1 and 2 overlap with a significant portion of the Changma period, their rainfall is characterized by continuous low-intensity rainfall. In contrast, relatively high-intensity, short-duration rainfall occurs mainly in Clusters 3 and 4. The WRF experiments generally describe the clustered weather patterns well. For CU_ON, the spatial distribution of the daily rainfall anomaly composite in Clusters 1 and 2 is well depicted, but the overall rainfall intensity is underestimated. CU_ON better reproduces the Clusters 1 and 2 type rainfall characterized by long-duration rainfall than CU_OFF. The observed rainfall events exceeding 20 mm h^{-1} intensity with a short-duration are reproduced better in CU_OFF than in CU_ON, showing reasonable performance for sub-daily time-scale rainfall in Clusters 3 and 4. Hence, the CU_ON well depicts the continuous low-intensity ERE type, while CU_OFF captures the ERE type where high-intensity rainfall with a short duration occurs relatively frequently.

1 Introduction

Dynamical downscaling is a method to generate fine-resolution climate data using a regional climate model (RCM), enabling more detailed and reliable climate information (Giorgi and Mearns 1991; Xu et al. 2019). This method can be a useful technique for overcoming the limitations of the general circulation model, which produces relatively coarse resolution output (Gao et al. 2008; Im et al. 2021). Thus, RCMs are reported to have appropriate resolving capability for regional weather and climate, particularly over complex terrain regions (Diro et al. 2012; Torma et al. 2015; Park et al. 2016; Karmacharya et al. 2017; Afrizal and Surussavadee 2018; Qiu et al. 2020). The characteristics and mechanisms of local-scale meteorological phenomena can be explored through dynamical downscaling using RCM.

Recent advances in computational systems have allowed regional climate simulations using a kilometer-scale RCM. The high-resolution model shows a more realistic and detailed structure of the local climate by better resolving the coastlines and topographic features than an RCM with a few tens of kilometers grid scale (Prein et al. 2015; Qiu et al. 2020). In particular, the deep convection process can be resolved explicitly without employing cumulus parameterization using RCM with a horizontal grid resolution of 4 km or less (Prein et al. 2015). The cumulus parameterization, which plays a role in representing the sub-grid scale deep convection process, is concerned with most of the dynamical and physical processes in the model simulation. The RCM parameterization associated with the cloud process has uncertainties because of the incomplete understanding

✉ Joong-Bae Ahn
jbahn@pusan.ac.kr

[1] Department of Atmospheric Sciences, Division of Earth Environmental System, Pusan National University, Busan, South Korea

[2] Department of Atmospheric Sciences, Pusan National University, Geumjeong-Gu, Busan 46241, South Korea

Published online: 24 August 2023

and many assumptions that unexpectedly affect the model performance. The convection-permitting model (CPM), a climate model without using cumulus parameterization schemes, is used for several regional climate studies (Lucas-Picher et al. 2021).

According to Prein et al. (2015), notable improvement has been made in simulating precipitation by employing the CPM instead of using error-prone cumulus parameterization (e.g., Hohenegger et al. 2008; Weusthoff et al. 2010; Prein et al. 2013; Ban et al. 2014; Zhao et al. 2021). Zhao et al. (2021) argued that CPM simulates the water cycle process over the Tibetan Plateau better than the large-scale model, improving the rainfall simulation. Prein et al. (2013) and Ban et al. (2014) argued that CPM could capture the diurnal cycle of summer precipitation in the Alpine region. On the other hand, several studies have also shown improved precipitation reproduction by employing cumulus parameterization in the high-resolution model (Kotroni and Lagouvardos 2004; Deng and Stauffer 2006; Wootten et al. 2016). Ishida et al. (2019) and Lee et al. (2011) reported that the benefits of activating the cumulus scheme in a high-resolution model depend on the dynamic forcing of a precipitation event or domain. Nevertheless, it is unclear whether the simulation ability is improved according to the use of cumulus parameterization in a high-resolution model.

Many studies, including the aforementioned ones, focused on simulating precipitation using a high-resolution climate model over a complex terrain region (Frei et al. 2003; Rauscher et al. 2010). From that perspective, South Korea is an appropriate region for a high-resolution model experiment. Specifically, 70% of the territory consists of mountain areas, and diverse precipitation variations appear depending on the season and region (Bae et al. 2008; Jung et al. 2011). Rainfall from May to September (MJJAS), which accounts for approximately 75% of the total annual precipitation, is a critical meteorological event in the Korean Peninsula (Jung et al. 2001; Kim et al. 2011). The BCM (Baiu in Japan, Changma in Korea, and Meiyu in China) is a complicated phenomenon in which the location and occurrence periods are determined by five different air masses (tropical North Pacific air mass, Okhotsk Sea air mass, tropical continental air mass over North China, tropical monsoonal air mass, and polar continental air mass), and significantly affects South Korea rainfall in MJJAS (Hong and Ahn 2015; Seo et al. 2015; Lee et al. 2017). In addition, rainfall by typhoons and local heavy rainfall occur in the same period, suggesting that various factors can cause MJJAS precipitation in South Korea. Therefore, high-resolution climate modeling is essential for the South Korean region, where precipitation varies from the local to synoptic scales and has complex terrain (Park et al. 2022a). The higher added value of RCM can be obtained in complex topography and coastlines by dynamical downscaling the coarse climate data (Fantini et al. 2018; Ciarlo et al. 2021). Several high-resolution model experiments have been performed to reproduce the MJJAS precipitation phenomenon in South Korea (Lee et al. 2011, 2019; Qiu et al. 2020; Seo and Ahn 2020; Park et al. 2022b); Lee et al. (2019) conducted two downscaling experiments with different horizontal resolutions (5 km and 20 km) and argued that high-resolution RCM is a more suitable tool for extreme hydrological events. Park et al. (2022b) performed high-resolution modeling for rainfall cases in South Korea using the scale awareness cumulus scheme. They showed improvement in simulating convective cells related to heavy rainfall in the gray zone (1–10 km grid spacing). However, most studies conducted short-period simulations or discussed only the added values obtained by increasing the spatial resolution of RCM. Few studies have evaluated rainfall simulations using cumulus parameterization with a high-resolution model.

This study aims to identify the characteristics of extreme rainfall events (ERE) according to the weather patterns in South Korea and compare the model simulation skills depending on the use of cumulus parameterization. For this purpose, EREs are clustered according to the synoptic scale weather pattern, and the ERE simulation performance is evaluated for each cluster. Daily and hourly rainfall data are analyzed for the May to September period (MJJAS), during which the mean monthly rainfall amount of each month exceeds 100 mm. The paper is organized as follows: model configuration and observation data are described in Sect. 2. The definition of ERE and k-means clustering method to divide weather patterns is also included in Sect. 2. Section 3 gives the results of evaluating ERE simulations regarding the daily and hourly time scale and analyzing atmospheric factors related to ERE for each cluster. Section 4 provides a summary and conclusions.

2 Data and methods

2.1 Model configuration

The Weather Research and Forecasting (WRF) version 4.0 (Skamarock et al. 2019) is used as the regional climate model in this study for conducting dynamical downscaling over South Korea. The European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA 5) reanalysis data provided by ECMWF is used as a lateral boundary condition with a spatial resolution of 31 km and a temporal resolution of 6 h (Hersbach et al. 2020). A two-way nesting technique is applied to downscale the 31 km data to a few kilometers. Domain 1 covers the Northeast Asia region centered in South Korea (36.5°N, 127.8°E) with a 9 km horizontal resolution. (Fig. 1a). Domain 2 focuses on South Korea with a 3 km horizontal resolution (Fig. 1b).

The physical schemes used in both Domains 1 and 2 are the Goddard microphysics scheme (Tao et al. 1989, 2016),

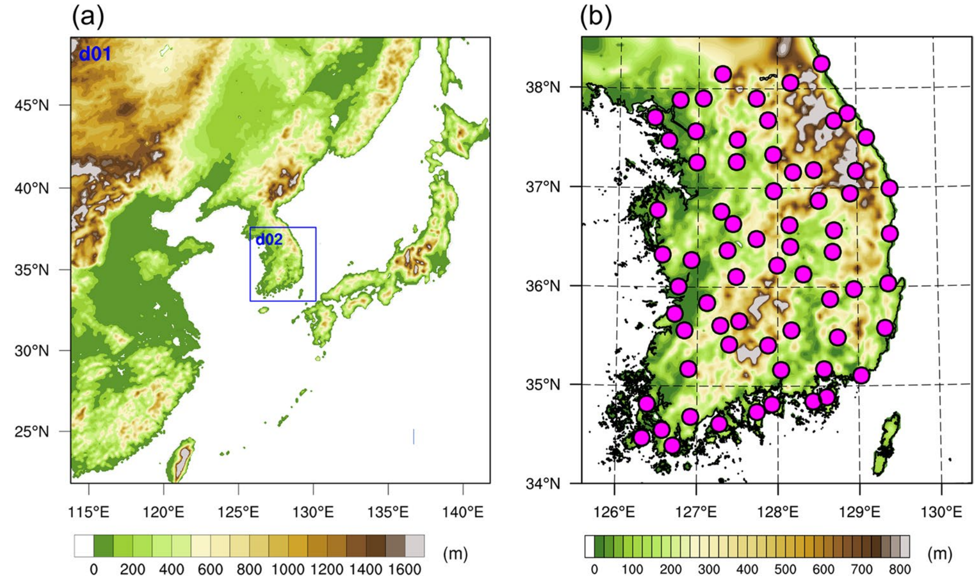

Fig. 1 **a** Topography (unit: m) of WRF configuration domain. d01: Domain 1 with 9 km grid spacing and d02: Domain 2 with 3 km grid spacing. **b** Topography (unit: m) of domain 2. The pink dots denote the location of the ASOS stations

the Community Atmosphere Model (CAM) longwave and shortwave radiation scheme (Collins et al. 2002), the Noah land surface model scheme (Chen and Dudhia 2001), and the Yonsei University Scheme (YSU) planetary boundary layer scheme (Hong et al. 2006) (Table 1). Two experiments with different cumulus parametrization settings are conducted to evaluate the high-resolution WRF simulations according to the cumulus parametrization usage. The experiment using the Kain–Fritsch cumulus scheme (Kain 2004; KF) in Domains 1 and 2 is denoted as CU_ON, and the experiment using the KF scheme only in Domain 1 and turning off the cumulus scheme in Domain 2 is denoted as CU_OFF. The KF scheme is the mass flux type scheme that uses a trigger function considering the thermodynamic status and a background vertical motion. One-dimensional entrainment-detrainment plume mode is used in the KF scheme. The configuration of the CU_ON and CU_OFF experiments is the same except for the use of the cumulus scheme, and the implementation period covers the first day of 2005 to the last day of 2013.

A sensitivity test is conducted from June to August 2005 to determine the optimal WRF model physical scheme configuration. Eight combinations of microphysics schemes (WRF Single–moment 3–class (Hong et al. 2004), Goddard scheme (Tao et al. 1989)) and cumulus schemes (Kain–Fritsch scheme (Kain 2004), Betts–Miller–Janjic scheme (Betts and Miller 1986; Janjić 1994), without cumulus scheme (OFF)), which closely contribute to the rainfall simulation, are constructed (Table S1). The domain and physical parameterizations are the same as in Table 1 except for microphysics and cumulus schemes. As a result, the Case 6 experiment, which uses the Goddard scheme for microphysics and Kain–Fritsch scheme with and without the cumulus in Domain 1 and 2, respectively, shows the best performance because the magnitude of the root mean square error is the lowest (2.10), and the pattern correlation coefficient is the second highest (0.56) (Figure S1, Table S2). The Case 6 experiment is denoted as CU_OFF, representing the CPM experiment, and the Case 2 experiment is denoted as CU_ON.

Table 1 WRF model configuration

EXP	CU_ON	CU_OFF
Resolution (nlat × nlon)	Domain 1: 9 km (340 × 320) Domain 2: 3 km (168 × 150)	
Vertical levels	33 eta levels	
Reference position	Lon: 127.8°/lat: 36.5° (Lambert conformal)	
Microphysics	Goddard	
Long/short wave radiation	CAM	
Land surface	Noah	
PBL	YSU	
Cumulus	Domain 1: Kain–Fritsch Domain 2: Kain–Fritsch	Domain 1: Kain–Fritsch Domain 2:-

2.2 Verification data

The Automated Synoptic Observing System (ASOS) in situ rainfall data provided by the Korea Meteorological Administration (KMA) are used. The Domain 2 rainfall data simulated by WRF are interpolated bilinearly into 67 ASOS stations to evaluate its performance (Fig. 1b). The Korea Local Analysis and Prediction System (KLAPS) reanalysis data provided by KMA are used to check the mesoscale pattern related to South Korean rainfall. The KLAPS reanalysis data cover the Korean Peninsula with a 5 km horizontal resolution, which is available from 2005 to 2013. The data include regular observation data comprised of ASOS, radar, satellite, and civil aircraft data, and the upper-air data observed during the special observation period (NIMS 2014).

2.3 Definition of extreme rainfall event

An extreme rainfall event (ERE) is defined with reference to the Expert Team on Climate Change Detection and Indices (ETCCDI; http://etccdi.pacificclimate.org) and Zhao et al. (2017). First, the 90th percentile daily precipitation in the 2005–2013 summertime (MJJAS) is calculated for each ASOS station. The case is defined as the day when more than four stations record precipitation that is greater than the 90th percentile daily precipitation. According to Zhao et al. (2017), an extreme precipitation event is defined as the day when the number of grid points that record the above rainfall threshold exceeds 5% of the total number of grid points. Therefore, four stations are set as a threshold, which is more than 5% of the 67 ASOS stations. The days under the influence of typhoons are excluded, and the ERE is selected separately for CU_ON, CU_OFF, and ASOS data. The typhoon days are when the typhoon that influences South Korea is located above 25°N and below 140°E in MJJAS.

2.4 Classification of weather patterns

The clustering method is utilized to divide the weather patterns around South Korea that influence the ERE occurrence. Many climate studies have analyzed the temperature and precipitation using the clustering method (Fragoso and Tildes Gomes 2008; Zhao et al. 2017; Bae et al. 2019; Jo et al. 2019; Nguyen-Le and Yamada 2019; Abadi et al. 2020; Kim et al. 2021). Kim et al. (2021) described three clustered large-scale atmospheric circulation patterns related to the extreme temperature in South Korea using a self-organizing map (SOM). Abadi et al. (2020) performed climate regionalization in Bolivia by clustering precipitation and temperature separately using the k-means approach, and the consensus clustering method was then applied. Two types of clustering approaches exist in climate research: time clustering and space clustering. In this study, k-means clustering is conducted over time using the 850 hPa daily geopotential height anomaly data (MacQueen 1967; Chattopadhyay et al. 2020). Non-hierarchical k-means clustering is one of the unsupervised learning algorithms that groups data into k clusters with the nearest to the centroid by calculating the distance between the k centroid and individual data. K-means clustering uses the 850 hPa daily geopotential height anomaly of ERA5 reanalysis data for observation and the Domain 1 data for model experiments. The Domain 1 data are used for clustering because the two-way nesting technique allows an interaction between the inner and outer domains in the WRF model (Liu et al. 2012; Wang et al. 2012). Before clustering, an empirical orthogonal function (EOF) is performed on the geopotential height anomaly data to extract the major components of the weather pattern (Chattopadhyay et al. 2020). The first four modes that explain 83% of the variance for observation (81% for CU_ON and CU_OFF) are kept and applied to k-means clustering. The Euclidean distance, the most widely used method, is employed to calculate the distance between the centroid and individual data.

The optimal number of clusters is determined through the clustering validity indices, the sum of squared error (SSE), and the mean silhouette coefficient (Rousseeuw 1987; Syakur et al. 2018). The SSE is defined using the following equation:

$$\mathbf{SSE} = \sum_{i=1}^{K} \sum_{x \in X_i} d^2(c_i, x) \tag{1}$$

where K is the total number of clusters; c_i is the centroid of the i cluster; $d(c_i, x)$ is the distance between c_i and each data in the i cluster. Generally, SSE decreases rapidly as K increases, and the inflection point where the decrement rate of SSE becomes much lower denotes the optimal number of clusters. The silhouette coefficient indicates a similarity to its cluster compared to other clusters. The silhouette coefficient is defined as follows:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), \mathrm{b}(i))} \tag{2}$$

where i denotes the data point belonging to a certain cluster; $a(i)$ denotes the average of the distances between i and the data within the same cluster; $b(i)$ denotes the average of distances between i and data in the nearest cluster in which i does not belong. Silhouette coefficient exists in a range of − 1 to 1. Silhouette coefficient closer to 1 means more appropriate clustering. For the SSE calculated in this study, the inflection point is shown when the number of clusters (K) is four or five (Figure S2). Although the silhouette coefficient is largest when $K=2$, $K=4$ with the second largest silhouette coefficient is judged as the optimal number of clusters when considering SSE. The clustering domain is 25–50°N, 115–140°E, and the results do not change significantly when the region becomes broader or narrower.

2.5 Integrated vapor transport

In this study, integrated vapor transport (IVT) describes moisture transport over South Korea.

$$\mathbf{IVT} = \left(-\frac{1}{g}\int_{1000hPa}^{300hPa} uq\,dp\right)i + \left(-\frac{1}{g}\int_{1000hPa}^{300hPa} vqdp\right)j \quad (3)$$

where $\boldsymbol{g}$ denotes the gravitational acceleration; $\boldsymbol{u}$ and $\boldsymbol{v}$ denote zonal and meridional wind (m s^{-1}), respectively; $\boldsymbol{q}$ denotes the specific humidity (kg kg^{-1}); $\boldsymbol{p}$ denotes the pressure; $\boldsymbol{i}$ and $\boldsymbol{j}$ denote zonal and meridional unit vector, respectively.

2.6 Statistical assessment methods

The bias, mean absolute error (MAE), root mean square error (RMSE), and pattern correlation coefficients (PCC) are calculated to evaluate the performance of WRF model experiments quantitatively. The definition of the calculation formula is as follows:

$$\mathbf{Bias} = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{M}_i - \boldsymbol{O}_i) \quad (4)$$

$$\mathbf{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{M}_i - \boldsymbol{O}_i\right| \quad (5)$$

$$\mathbf{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{M}_i - \boldsymbol{O}_i)^2} \quad (6)$$

$$\mathbf{PCC} = \frac{\sum_{i=1}^{N}(\boldsymbol{M}_i - \overline{\boldsymbol{M}})(\boldsymbol{O}_i - \overline{\boldsymbol{O}})}{\sqrt{\sum_{i=1}^{N}(\boldsymbol{M}_i - \overline{\boldsymbol{M}})^2}\sqrt{\sum_{i=1}^{N}(\boldsymbol{O}_i - \overline{\boldsymbol{O}})^2}} \quad (7)$$

where $\boldsymbol{N}$ denotes the number of ASOS stations; M_i and O_i denote the data of the model and observation for each station, respectively. $\overline{M}$ and $\overline{O}$ denote the mean of the model and observation averaged over all stations, respectively.

3 Results

3.1 Evaluation of extreme rainfall event (ERE) simulation under different weather patterns

The identified EREs in observation, CU_ON and CU_OFF during the period (2005–2013) are 379, 370, and 355 days, respectively, indicating that WRF experiments provide a reasonable simulation of the number of ERE days. Figure 2 presents the mean ERE precipitation of the observation (ASOS) and two model experiments. Relatively high precipitation is distributed in the western and northern regions of South Korea for the observation. The overall precipitation around the central region is overestimated in CU_ON (Fig. 2b, d), and the station-averaged bias is 1.93. In contrast, CU_OFF shows negative precipitation biases in the northwest and south coast region (Fig. 2c, e), showing a − 1.20 station-averaged bias. In the central region, CU_OFF simulates a similar amount to the observation compared to CU_ON, although a weak wet bias is shown. In addition, CU_OFF represents a lower magnitude of bias, mean absolute error (MAE), and root mean square error (RMSE) than CU_ON for the total stations (e.g., the RMSE of CU_ON and CU_OFF is 3.56 and 3.32, respectively), representing a reasonable simulation (Table 2). CU_ON shows a higher pattern correlation coefficient (PCC) than CU_OFF (i.e., CU_ON: 0.72, CU_OFF: 0.54).

Clustering is conducted using ERA5 reanalysis data on the ERE occurrence days to analyze the ERE according to the weather patterns around South Korea. Figure 3 presents a composite of 850 hPa daily anomalies of geopotential height (GPT) and wind for each cluster. Both CU_ON (Fig. 3e–h) and CU_OFF (Fig. 3i–l) simulate similar patterns to the observation for all clusters, and the same results are shown in Fig. 4, which represents a composite of daily sea level pressure anomaly, 500 hPa GPT, and 200 hPa zonal wind.

The Cluster 1 pattern represents a negative anomaly of 850 hPa GPT and the sea level pressure north of the Korean Peninsula (Fig. 3a, e, i). The 500 hPa trough also appears in that region, suggesting that South Korea is affected by the negative GPT anomaly in the mid-to-lower level. In Cluster 1, ERE appears to occur as the cold air moves southward by cyclonic circulation and meets with relatively warm air in South Korea, resulting in atmospheric instability. Relatively cold air from the north passes through the Yellow Sea under the influence of cyclonic circulation, resulting in a westerly wind in South Korea. The WRF model can capture the location of the negative 850 hPa GPT anomaly core and shows the magnitude of westerly wind over South Korea, similar to the observations.

The Cluster 2 pattern is characterized by a confrontation between negative anomaly over northeastern China

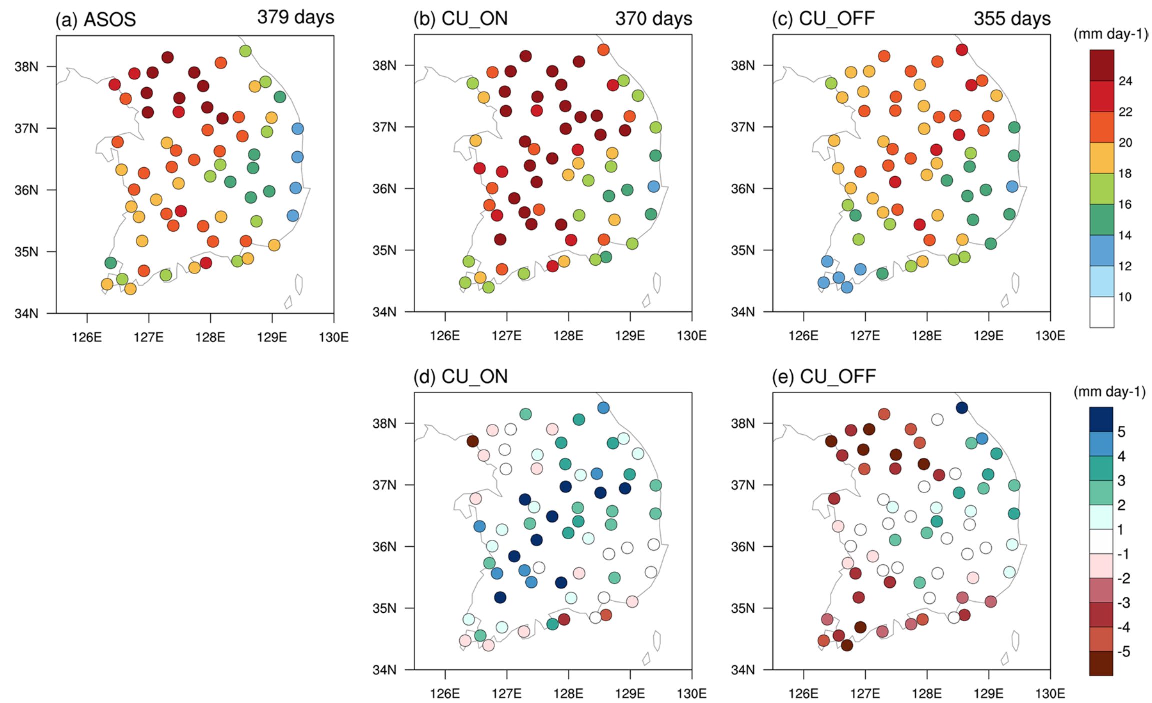

Fig. 2 Spatial distribution of the ERE precipitation mean (unit: mm day^{-1}) at the ASOS stations derived from **a** observation (ASOS), **b** CU_ON, and **c** CU_OFF and model biases for **d** CU_ON and **e** CU_OFF. The right upper number denotes the number of ERE days for each data

Table 2 Bias, RMSE, MAE, and PCC of the ERE precipitation mean derived from CU_ON and CU_OFF

	CU_ON	CU_OFF
Bias	1.93	−1.20
MAE	2.89	2.72
RMSE	3.56	3.32
PCC	0.72	0.54

and positive anomaly over the western North Pacific, forming a frontal boundary over South Korea (Fig. 3b, f, j). The GPT gradient is the largest over South Korea, and the strong southwesterly wind develops along the front. For the 500 hPa field, the 5880 gpm line is located at 32°N, suggesting an expansion of the western North Pacific subtropical high toward South Korea (Fig. 4b, f, j). This weather pattern is featured by the collision of two different air masses generating the rainfall band in South Korea during the East Asia summer monsoon period (Seo et al. 2015; Lee et al. 2017). A warm and humid southwesterly wind that blows into South Korea contributes to the development of ERE. In addition to Cluster 1, the fraction of observed Cluster 2 dates that overlap with the Changma period is higher than the other clusters (Cluster 1: 43.52%, Cluster 2: 36.26%, Cluster 3: 36.11%, Cluster 4: 30.53%). A comparison of the WRF model results with the observation shows that the two model experiments weakly simulate the meridional GPT gradient at 850 hPa. In addition, the center of the positive GPT anomaly in Japan is relatively biased to the southwest in the WRF experiments.

In Cluster 3, a distinct cyclonic anomaly appears in the southwestern region of South Korea, with counterclockwise flow occurring over the entire South Korea (Fig. 3c, g, k). South Korea, which is affected by negative sea level pressure anomaly, is located south of the upper-level jet entrance. Furthermore, a 500 hPa trough appears west of South Korea, indicating a baroclinic instability pattern (Fig. 4c, g, k). The development of an extratropical cyclone associated with baroclinic instability appears to cause ERE in Cluster 3. The model experiments can capture most of the characteristics in the Cluster 3 pattern, like the position of the jet stream. On the other hand, the simulated intensity of the extratropical cyclone is stronger than observed.

Lastly, the Cluster 4 pattern is described as a dominant positive GPT anomaly, unlike other clusters, which mostly show a negative anomaly (Fig. 3d, h, l). A relatively weak positive anomaly compared to the surrounding area penetrates the western region of South Korea, resulting in the southerly wind in South Korea. The baroclinic instability is also confirmed in Cluster 4 because South Korea is located east of the 500 hPa trough and south of the upper-level jet entrance (Fig. 4d, h, l). The EREs in Cluster 4 are considered localized heavy rainfall caused by diverse local factors (e.g., local instability, terrain effect) (Miyasaka et al. 2020; Park et al. 2021b). CU_ON and CU_OFF simulate the

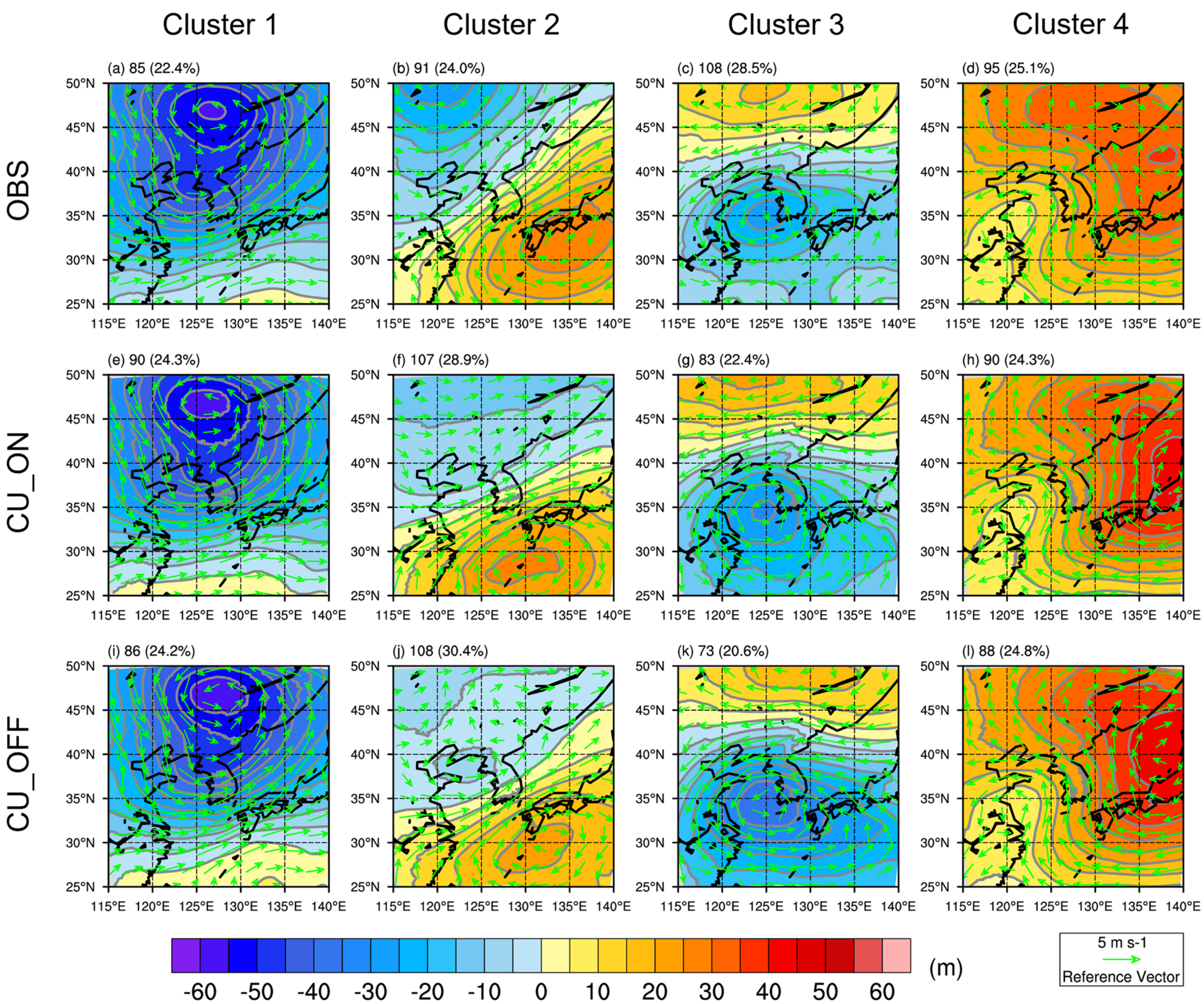

Fig. 3 Composite of the geopotential height daily anomaly (unit: m, shading) and wind daily anomaly (unit: m s^{-1}, vector) at 850 hPa for each cluster derived from **a–d** observation (ERA5), **e–h** CU_ON, and **i–l** CU_OFF. The upper left text denotes the number of occurrence days, and the text in parentheses denotes occurrence frequency

positive GPT anomaly core stronger than the observations. A stronger 850 hPa southerly wind appears over South Korea because the gradient of the 850 hPa GPT in WRF results is greater than the observed gradient.

Figure 5 shows the occurrence frequency of ERE in the 15-day bin period for each cluster. Cluster 1 has the highest frequency in the first (1st–15th) and second (16th–31st) half of July (Fig. 5a). CU_ON simulates the frequency of Cluster 1 well by appearing the highest in July with a similar frequency in the first and second half of July. CU_OFF shows only one strong peak in the second half of July, revealing a slightly different pattern from the observation, but it follows the observed variability well after August. Two frequency peaks appear in early July and early August for Clusters 2 and 3 of observation, which coincide with the summer monsoon period in South Korea (Fig. 5b, c). The WRF experiments show opposite results for Clusters 2 and 3. For Cluster 2, CU_ON can capture the timing of two peaks. Despite overestimating the occurrence frequency in the first half of July, it follows the observed variation in the overall period. CU_OFF represents the first peak but cannot simulate the second peak in the first half of August. For Cluster 3, both CU_ON and CU_OFF provide a good simulation of the first peak observed in early July. By contrast, CU_ON shows different variability from the observation after July, whereas CU_OFF depicts the second peak, albeit later than the observation. Cluster 4 shows a relatively higher occurrence frequency in late August than other clusters in observation (Fig. 5d). Therefore, Cluster 4 appears to occur mainly in late summer when the western North Pacific subtropical high

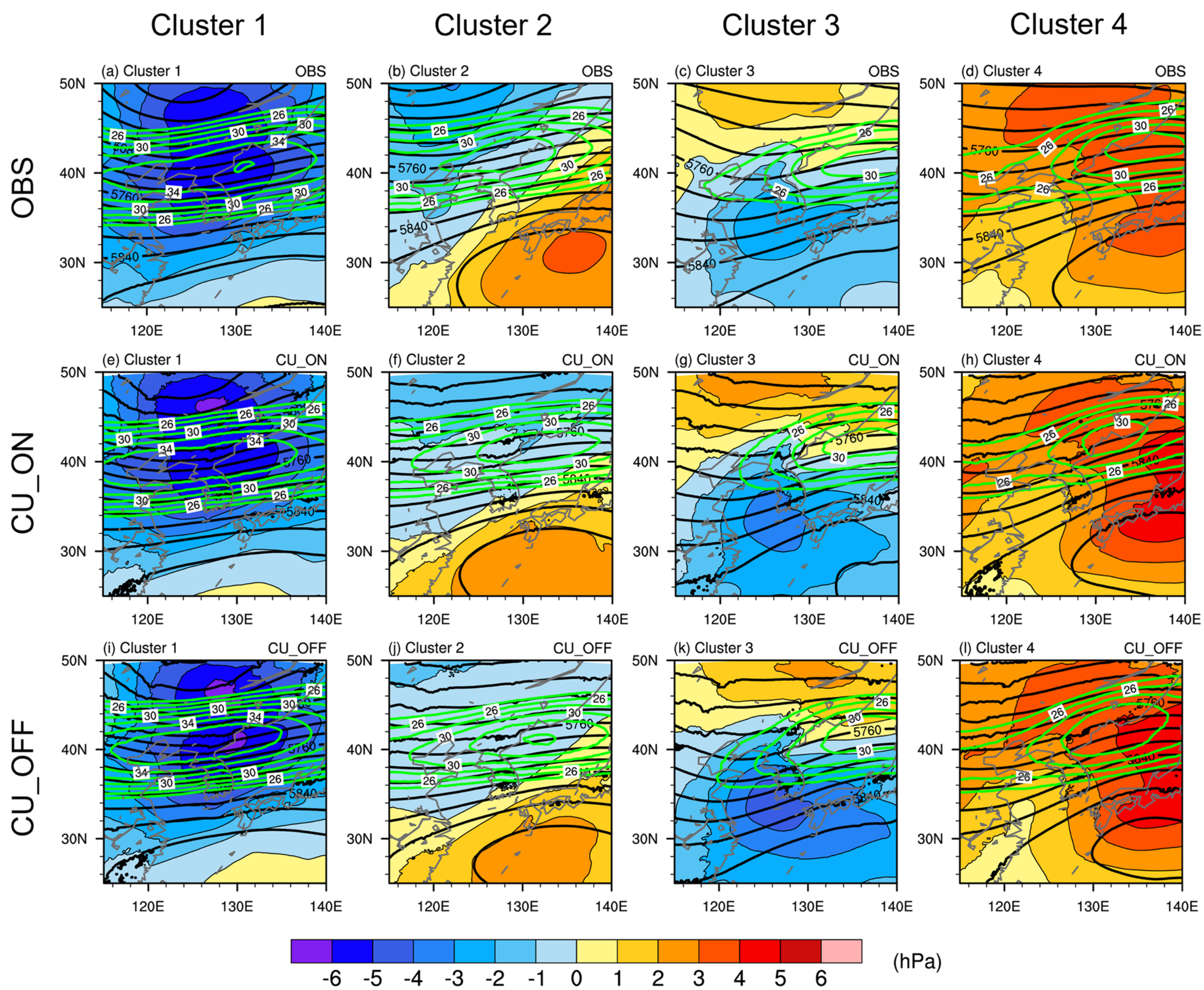

Fig. 4 Same as Fig. 3, but for the sea level pressure daily anomaly (unit: hPa, shading), 500 hPa geopotential height (unit: m, black solid thick line), and wind speed at 200 hPa (unit: m s^{-1}, solid green line)

develops, supporting the results above regarding the weather pattern. CU_ON and CU_OFF simulate two peaks in the same period as the observation, but CU_OFF shows better performance because the frequency is closer to observation than CU_ON.

Figure 6 presents the daily anomaly precipitation composite of the observation and WRF simulation results to evaluate the performance of the WRF rainfall simulation for each cluster. In the observation case, positive anomalies are shown in all stations with a different rainfall anomaly pattern for each cluster. For Cluster 1, more than 10 mm day^{-1} of rainfall is observed in more than half of the stations around the southwestern region in South Korea (Fig. 6a). For Clusters 2 and 3, significant rainfall is shown in a relatively narrow area compared to Cluster 1 (Fig. 6b–c). Cluster 4 has no distinct spatial pattern characteristics and shows lower rainfall than the other clusters (Fig. 6d), showing 11.2, 11.1, 9.7, and 9.3 mm day^{-1} station-averaged ASOS rainfall for Clusters 1, 2, 3, and 4, respectively. The amount of rainfall appears to be offset by averaging because the EREs in Cluster 4 occur locally. CU_ON can capture the major precipitation region of Clusters 1, 2, and 3, showing pattern correlation coefficients of 0.61, 0.64, and 0.57, respectively (Fig. 6e–h, Table 3). On the other hand, CU_ON tends to underestimate the Cluster 2 precipitation in the southern coastal region, which is also shown in CU_OFF. The reason can be found in the synoptic field. The center of the positive GPT anomaly, which is biased to the southwest in WRF simulation, leads to the orientation of the front more zonally than the observation (Fig. 3f, j). Therefore, the warm and humid wind blows toward the midwestern region rather than the southern region of South Korea, resulting

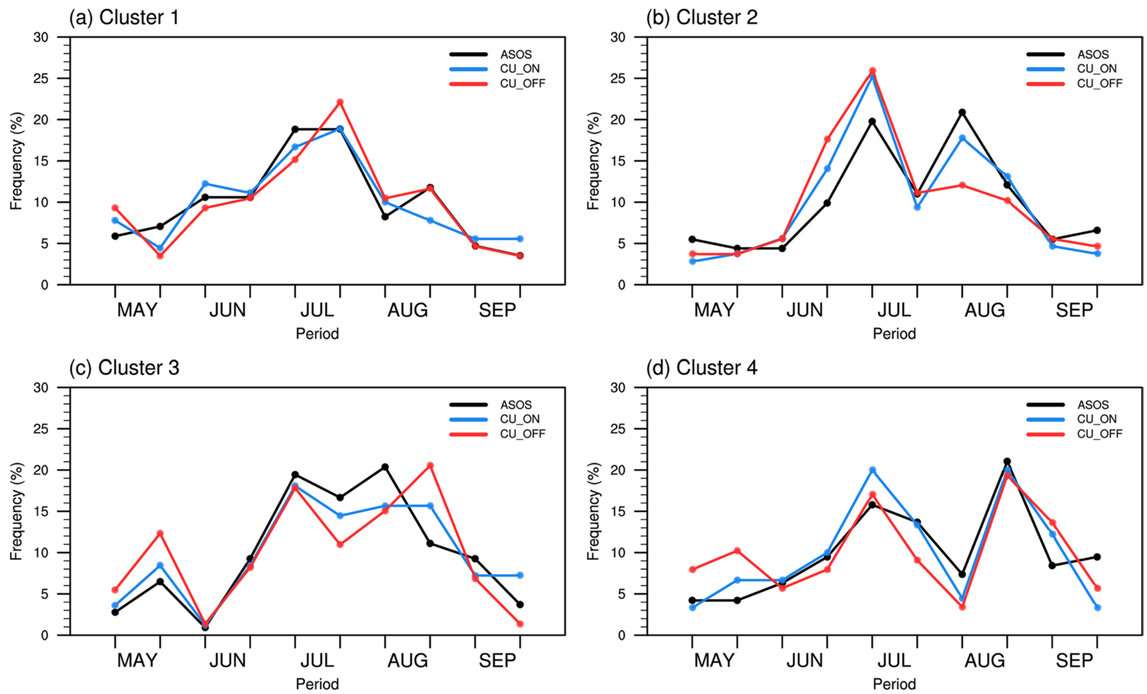

Fig. 5 Occurrence frequency (unit: %) of the ERE in the 15-day bin from May to September for each cluster

in dry biases over the southern coastal region. In Cluster 4, CU_OFF shows good performance regarding a quantitative evaluation, while CU_ON tends to overestimate the rainfall, showing 0.81 mm day^{-1} mean bias. CU_OFF represents the amount of rainfall core in Cluster 1, similar to the observation. In contrast, it underestimates the overall rainfall of the other clusters. For Cluster 3, CU_OFF has a limitation in describing the spatial pattern of observed Cluster 3 by overestimating the southwest precipitation.

ERE is also analyzed using the hourly precipitation data to evaluate the characteristics of ERE according to the weather pattern in detail (Figs. 7, 8). Figure 7 shows a scatter plot of the rainfall duration, intensity, and amount. The rainfall duration is defined by the consecutive hours of rainfall above 0.1 mm h^{-1} in ERE until below 0.1 mm h^{-1} rainfall occurs for two consecutive hours. The rainfall amount is defined by the accumulated rainfall for the duration, and the rainfall intensity is the rainfall amount divided by the duration. Figure 7 denotes the rainfall intensity, duration, and amount as the radius, angle, and dot color, respectively.

For the observations, the overall rainfall intensity in Cluster 1 is distributed primarily below 15 mm h^{-1}, which is weaker than the other clusters, and more than 200 mm cases are rarely seen (Fig. 7a). In Cluster 2, the rainfall intensity is distributed evenly over the entire duration, indicating that cases above 10 mm h^{-1} of rainfall lasting for more than 8 h are more common than with other clusters (Fig. 7b). Rainfall events with long durations and large amounts are commonly seen in Cluster 2 because Cluster 2 is the representative pattern of the summer monsoon period. In Clusters 3 and 4, high-intensity rainfall events are distributed prominently in the range of durations less than 6 h (Fig. 7c, d). The rainfall cases with more than 20 mm h^{-1} intensity lasting less than 6 h, and the cases with above 250 mm lasting for more than 21 h appear in Cluster 3, showing that the high-intensity with a short-duration type and continuous precipitation with a large amount type coexist. Cluster 4 has the highest number of rainfall cases, with a high intensity lasting less than 6 h compared to the other clusters. Overall, the Cluster 1 and 2 types of rainfall are characterized by continuous low-intensity rainfall. By contrast, relatively high-intensity, short-duration rainfall mainly occurs in Clusters 3 and 4. The same results can be found through the profile of apparent heat source (Q1) and apparent moisture sink (Q2) area-averaged over South Korea (Figure S3). According to previous studies, frontal precipitation can occur if the maximum of Q1 and Q2 profiles are at the same height, and deep convective precipitation can occur if the maxima are at different heights (Lee et al. 2008; Geetha and Balachandran 2016). For observation, the peak of Q1 and Q2 appears at similar heights in Clusters 1 and 2, whereas the height of the Q2 maximum appears lower than Q1 in Clusters 3 and 4 (Figure S3a-d). The WRF experiments show similar results to the observations for each cluster, confirming that the EREs of Clusters 1 and 2 and those of Clusters 3 and 4 show different rainfall types.

The WRF experiments show different performances for each cluster. Most of the rainfall cases in CU_ON are

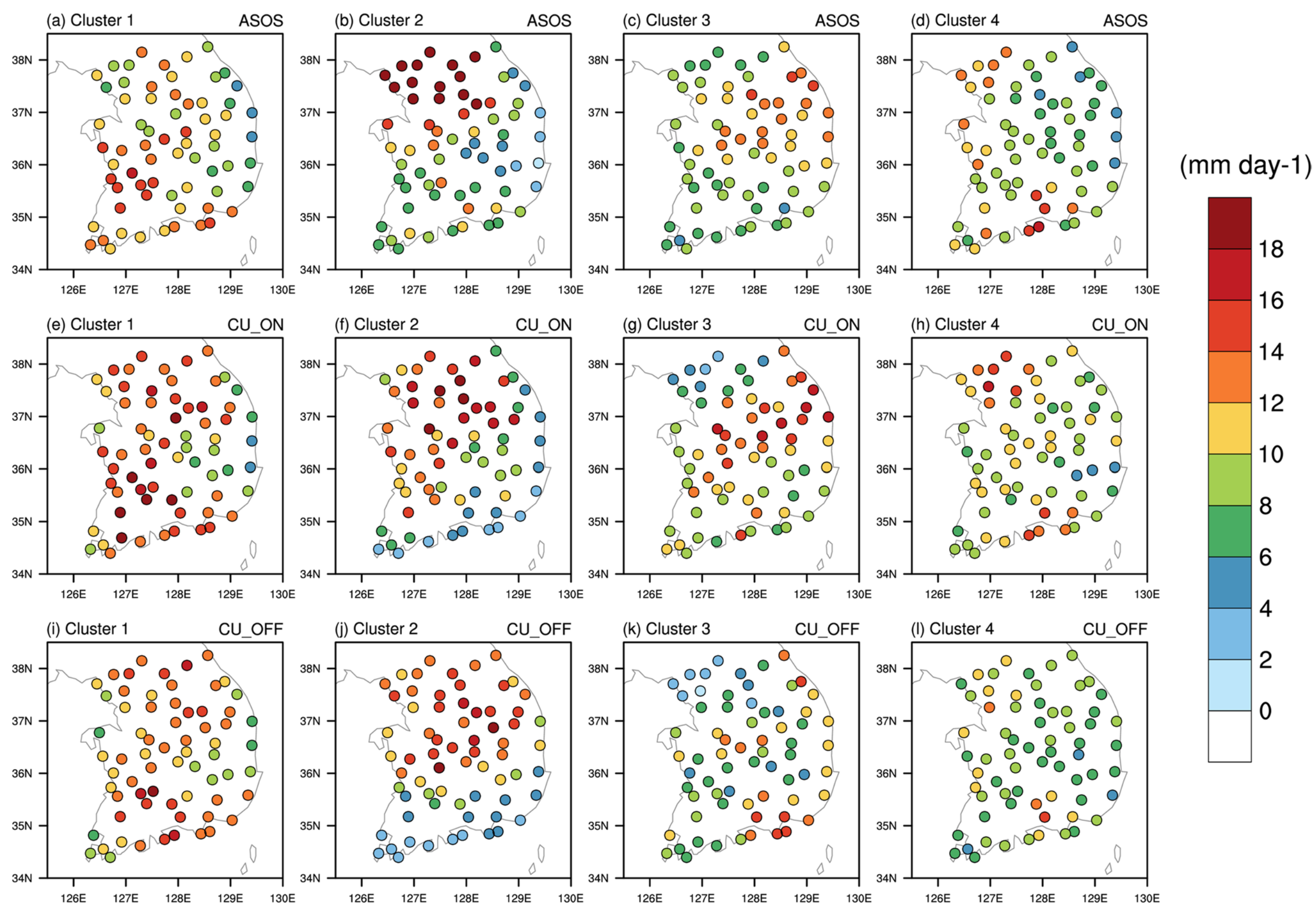

Fig. 6 Composite of the precipitation daily anomaly (unit: mm day^{-1}) at ASOS station for each cluster derived from **a–d** observation (ASOS), **e–h** CU_ON, and **i–l** CU_OFF

Table 3 Bias, MAE, RMSE, and PCC of the mean precipitation for each cluster derived from CU_ON and CU_OFF

	CU_ON				CU_OFF			
	Bias	MAE	RMSE	PCC	Bias	MAE	RMSE	PCC
Cluster 1	1.73	2.49	3.32	0.61	0.94	2.39	2.96	0.47
Cluster 2	−0.41	4.22	5.00	0.64	−0.35	5.48	6.27	0.39
Cluster 3	0.78	2.40	3.03	0.57	−1.08	3.39	4.03	0.11
Cluster 4	0.81	2.08	2.60	0.53	−0.74	1.78	2.28	0.58

distributed below a 15 mm h^{-1} intensity (Fig. 7e–h). Despite underestimating the rainfall intensity in CU_ON, the wet bias of the mean ERE precipitation appears because CU_ON overestimates the occurrence frequency of hourly precipitation (Figure S4). Regarding the distribution of occurrence frequency of more than 0.1 mm h^{-1} rainfall, CU_ON shows a frequency of more than 30% at most stations for all clusters. Otherwise, CU_OFF shows an occurrence frequency similar to the observation. RCM with an activating KF cumulus scheme tends to simulate low-intensity rainfall over a wide area frequently, and most of the simulated rainfall cases are characterized by continuous low-intensity precipitation (Pennelly et al. 2014; Mayor and Mesquita 2015; Konduru and Takahashi 2020). Therefore, rainfall with low intensity and long duration appears mainly in most clusters of CU_ON, even though rainfall with over 20 mm h^{-1} and a 3–6 h duration also appears occasionally in Cluster 2. CU_ON represents a similar distribution to observed Clusters 1 and 2, considering rainfall duration, intensity, and amount. On the other hand, it simulates rainfall intensity as relatively weak in Clusters 3 and 4, where intense short-term rainfall occurs in observation, which can be related to the drizzle problem of CU_ON. If precipitation continues and clouds are formed continuously, deep convection and heavy rainfall

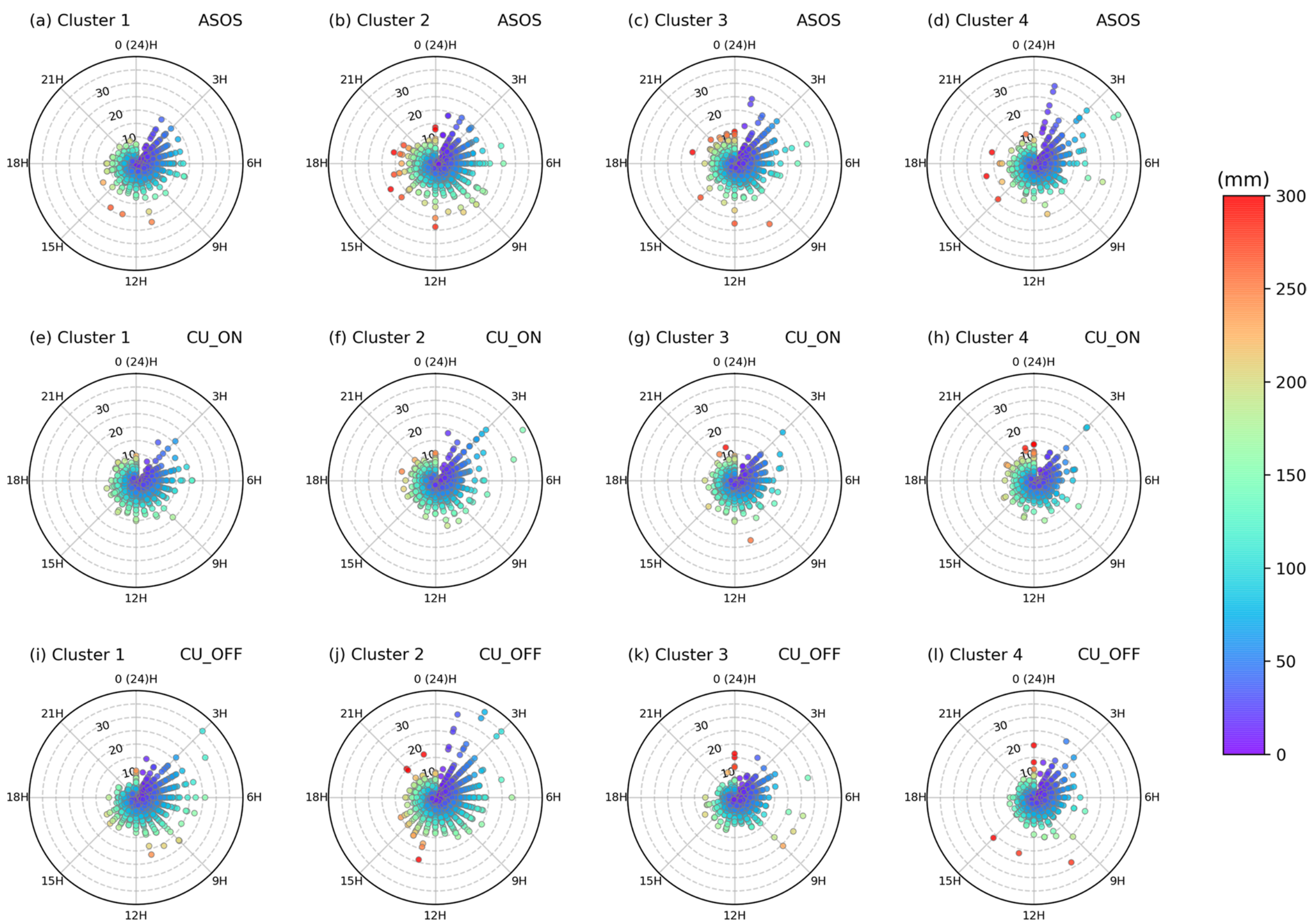

Fig. 7 Scatter plot of the ERE rainfall intensity (unit: mm h^{-1}, radius) and duration (unit: hours, angle) for each cluster derived from **a–d** observation (ASOS), **e–h** CU_ON, and **i–l** CU_OFF. The color of the dots denotes the ERE rainfall amount (unit: mm)

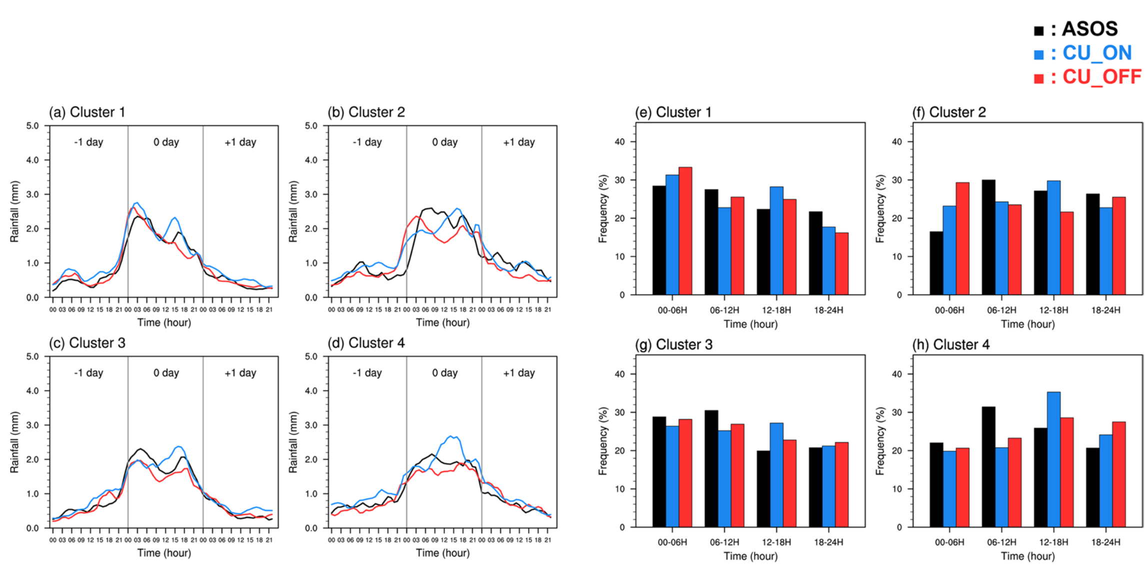

Fig. 8 **a–d**) Composite of the hourly rainfall time series for 3 days (−1 to +1 day of ERE) and **e–h** occurrence frequency of the maximum hourly rainfall in 6 h bin for each cluster

are inhibited by not allowing the increase in temperature and humidity (Fosser et al. 2015). This problem can be solved to some extent using the CPM. The CPM performs well, particularly in extreme precipitation and hourly time-scale rainfall (Prein et al. 2013; Ban et al. 2014; Fosser et al. 2015; Meredith et al. 2015). CPM, which explicitly simulates convection, shows improvements for short-term high-intensity rainfall compared to the RCM using the cumulus scheme because deep convection plays a vital role in mid-latitude extreme precipitation. Clusters 3 and 4 of CU_OFF show high-intensity EREs with less than 6 h duration, showing better performance than CU_ON (Fig. 7k, l). For Cluster 3, CU_OFF shows similar rainfall intensity to observations for 3–6 h duration, despite underestimating the rainfall intensity for less than 3 h duration. The numerous rainfall cases with more than 15 mm h^{-1} intensity are distributed in the range of the short duration for Cluster 4, indicating the capability to capture high-intensity rainfall with a short duration. In Clusters 1 and 2, CU_OFF overestimates the rainfall intensity with less than 9 h duration, resulting in a different distribution from the observation.

Figure 8 shows the diurnal rainfall variations and occurrence frequency of the maximum hourly precipitation for each cluster. For the observation of Clusters 1 and 3, rainfall peak in the diurnal variation appears late at night (00–06 local standard time (LST)) and in the afternoon (12–18 LST) (Fig. 8a, c). The maximum rainfall occurs mainly late at night (00–06 LST) and in the morning (06–12 LST) for Clusters 1 and 3, respectively (Fig. 8e, g). Clusters 2 and 4 have no diurnal variation features on the ERE day (Fig. 8b, d). The frequency of maximum precipitation for Cluster 2 appears to be more than 25% in all periods after 06 LST, whereas a relatively high frequency appears in the daytime for Cluster 4. Considering the diurnal variation and frequency of maximum precipitation, CU_ON (CU_OFF) performs well for Clusters 1 and 2 (3 and 4). CU_ON can capture the observed peaks in the diurnal variations and the peak time of the frequency of maximum precipitation in Cluster 1, despite overestimating the rainfall. A high frequency during the daytime affects the second rainfall peak in the diurnal cycle because relatively low-intensity rainfall is frequently simulated in CU_ON. Thus, the diurnal cycle of CU_ON is similar to the observations, even though the frequency at 12–18 LST is higher than observed.

For the frequency of maximum rainfall in Cluster 2, CU_ON shows a similar result to the observation for the entire time. In contrast to the observation, the highest frequency is shown late at night in CU_OFF. In Cluster 3, CU_OFF follows the observed diurnal variation in which two peaks appear with similar rainfall. In addition, CU_OFF well simulates the observed frequency distribution of maximum rainfall in Cluster 3 and 4 by presenting closer to the observation than CU_ON.

3.2 Mesoscale features linked to ERE type

Precipitation is formed by a large circulation field and interaction between other meteorological factors (Park et al. 2021a; Zhao et al. 2022). Therefore, variables significantly affecting rainfall are analyzed for each cluster using the KLAPS reanalysis data and model data. Figure 9 shows the daily anomaly moisture transport composite over South Korea. Figure 9 presents the daily anomaly IVT as a vector and the magnitude of daily anomaly IVT as shading. For the observation of Clusters 1 and 2, the overall magnitude of the IVT anomaly in South Korea is above 120 kg m^{-1} s^{-1}, showing a larger IVT magnitude than the other clusters (Fig. 9a, b). Warm and humid air from the low latitude region moves into South Korea through the westerly and southwesterly wind in Clusters 1 and 2, inducing ERE (Song and Sohn 2015). For CU_ON, the magnitude of IVT simulated in Cluster 1 is similar to the observation (Fig. 9e). CU_ON captures the observed IVT core in the western region near 36°N and simulates the direction of IVT vectors in Cluster 2 well. The CU_OFF overestimates (underestimates) the IVT magnitude of Cluster 1 (Cluster 2). In the observed Clusters 3 and 4, the magnitude of IVT anomaly is lower than 100 kg m^{-1} s^{-1}, showing opposite results to Clusters 1 and 2 (Fig. 9c, d). Hence, no significant relationship appears between IVT and rainfall in Cluster 3 and 4 because most of the IVT vectors over South Korea are not statistically significant.

Figure 10 presents a latitude-height cross-section of the equivalent potential temperature (EPT) daily anomaly for the 126–129°E land area averaged where South Korea is located. The observed EPT distribution shows different patterns for each cluster (Fig. 10a–d). More than 5 K EPT appears in the low level of South Korea for observed Clusters 1 and 2, which is relatively larger than Clusters 3 and 4 (Fig. 10a, b). This is related to the large inflow of IVT from the warm low latitude region to South Korea (Fig. 9). In Cluster 3, a large EPT gradient with a poleward-tilted structure is observed in South Korea (Fig. 10c). This feature indicates a baroclinicity structure, which aligns with Figs. 3 and 4. CU_ON shows good performance simulating a distribution of more than 5 K in South Korea, similar to the observation for Clusters 1 and 2 (Fig. 10e, f). In Cluster 2, CU_ON represents the observed pattern well in that the positive anomaly extends to the low level of the 42°N region, even though it underestimates the EPT in the high latitude region by one degree. On the other hand, CU_OFF overestimates the EPT gradient in South Korea, showing relatively low skill compared to CU_ON in Cluster 1 (Fig. 10i). Nevertheless, it can describe the observed poleward tilted structure of Cluster 3, showing a negative anomaly extending to 36°N (Fig. 10k). Furthermore, CU_OFF shows a similar spatial pattern to the observations of Cluster 4 by simulating approximately 4 K EPT

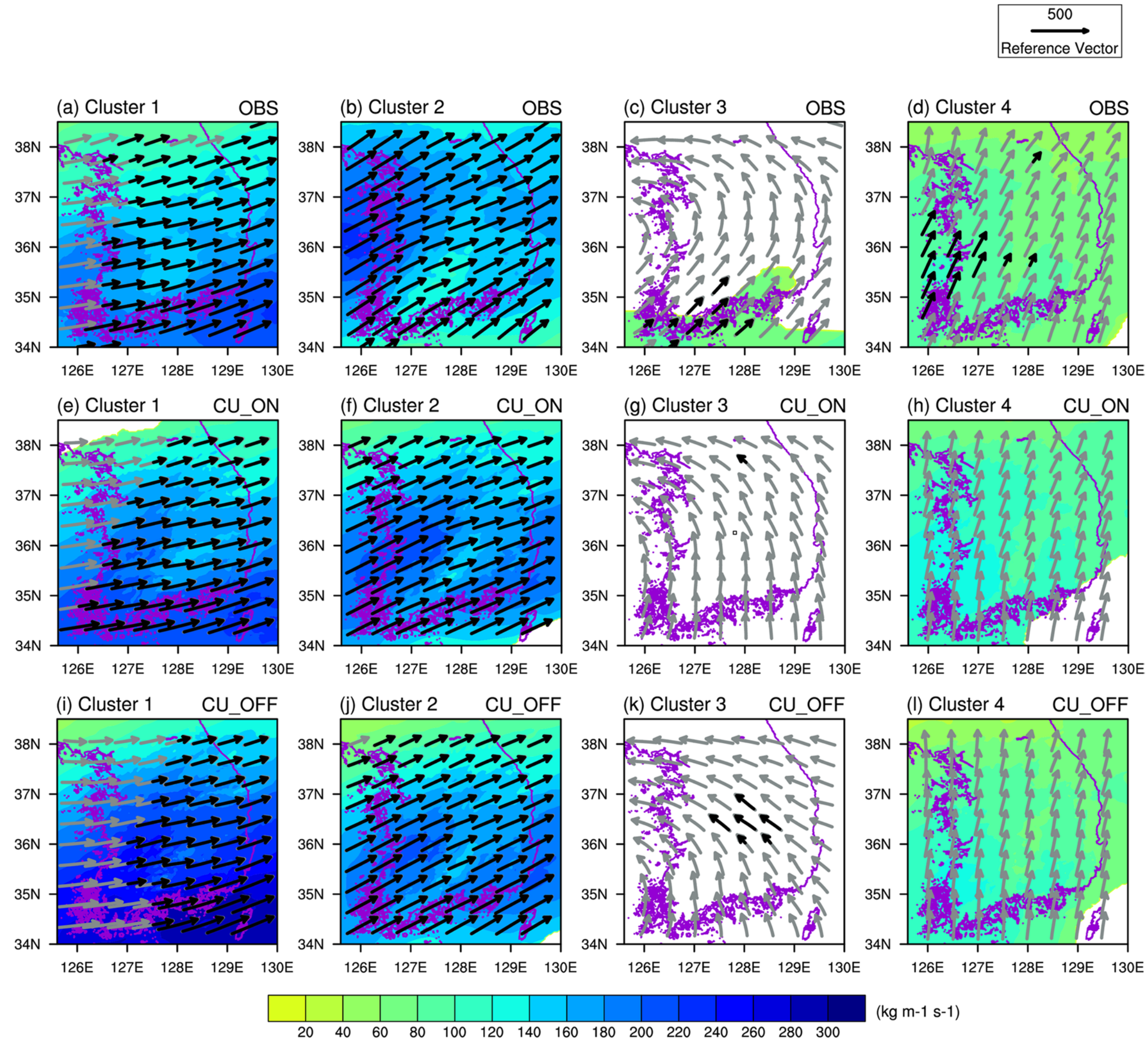

Fig. 9 Composite of IVT daily anomaly (unit: kg m^{-1} s^{-1}, vector) and magnitude of IVT daily anomaly (shading) for each cluster derived from **a**–**d** observation (KLAPS), **e**–**h** CU_ON, and **i**–**l** CU_OFF. Shadings are shown only for the statistically significant values at the 95% confidence level. Black vectors denote that the zonal and meridional wind anomalies are statistically significant at the 95% confidence level

in the mid-lower level over South Korea, whereas CU_ON overestimates it (Fig. 10h, l). Consequently, the good performance for simulating rainfall-related variables according to the two WRF experiments seems to contribute to the advantage in their ERE precipitation simulation.

4 Discussion

The WRF experiments results show that the CU_ON and CU_OFF experiments have different advantages in simulating ERE rainfall, but the reasons for this are slightly ambiguous. One possibility is that the trigger function of the KF scheme appears to interfere with the reproduction of high-intensity rainfall. The Fritsch–Chappell trigger function used in the KF scheme has a drawback: it tends to simulate widespread light rainfall in marginally unstable environments (Warner and Hsu 2000; Yu and Lee 2011; Mayor and Mesquita 2015). Fosser et al. (2015) argued that a lower lifting condensation level (LCL) and level of free convection (LFC) can lead to easier cloud formation, generating drizzle. Therefore, additional analysis is conducted for the Cluster 4 rainfall, which has distinct sub-daily characteristics and shows different performance between CU_ON and CU_OFF. Figure 11 presents the diurnal variation of LCL and LFC (Fig. 11a) and the Hovmöller diagram of the p-velocity (ω) in CU_ON and CU_OFF (Fig. 11b, c) for Cluster 4. CU_ON generally shows a lower LCL and LFC than CU_OFF in Cluster 4, indicating more favorable conditions for cloud formation. For the distribution of ω according

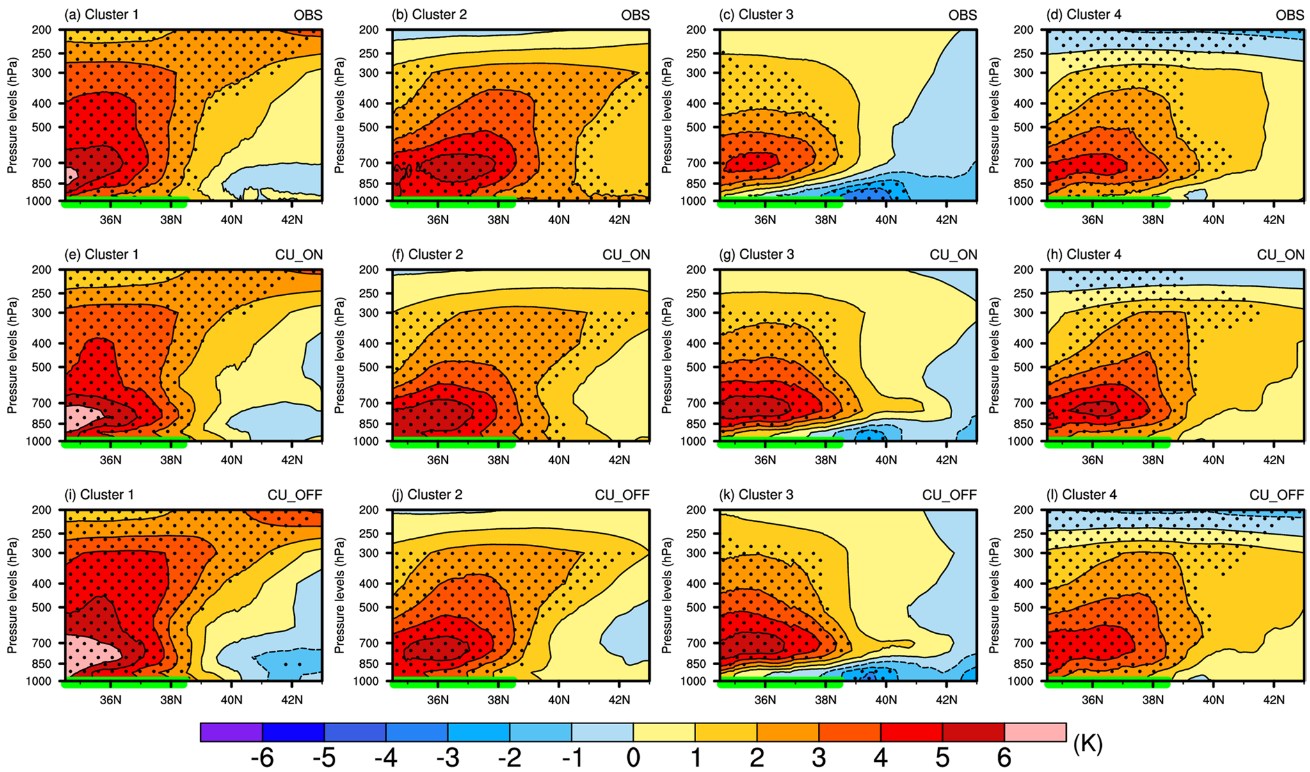

Fig. 10 Composite of latitude–height cross section of equivalent potential temperature daily anomaly (unit: K) for each cluster derived from **a–d** observation (KLAPS), **e–h** CU_ON, and **i–l** CU_OFF. Black dots denote the statistically significant values at the 95% confidence level. Green line denotes the region where South Korea is located

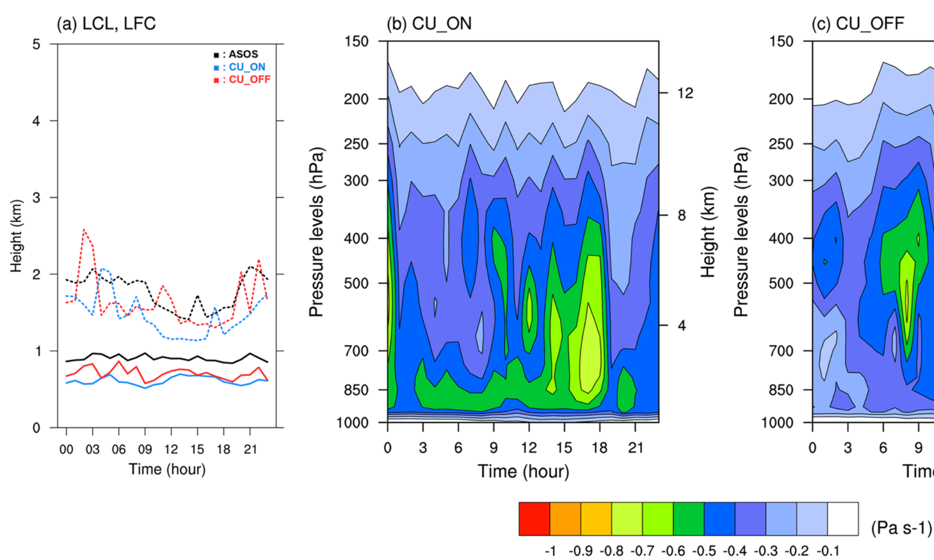

Fig. 11 **a** Diurnal variation of lower lifted condensation level (LCL) and level of free convection (LFC), and the Hovmöller diagram of p-velocity (unit: Pa s^{-1}) derived form **b** CU_ON and **c** CU_OFF for Cluster 4

to time, CU_ON shows continuous rising motion with values below -0.5 pa s^{-1} since 09 LST (Fig. 11b). These results are consistent with previous results in that continuous rainfall mainly occurs in CU_ON. This is because the continuous cloud formation suppresses the deep convection in Cluster 4 of CU_ON, resulting in a relatively weak continuous precipitation occurrence instead of short-duration, high-intensity rainfall. On the other hand, CU_OFF shows low ω values occasionally, indicating an environment where the precipitation patterns in observed Cluster 4 appear. Therefore, these environmental simulation differences, including IVT and EPT, are responsible for the different rainfall simulation performances of the WRF models. In other clusters, there may be differences in simulation performance because of environmental characteristics specific to each precipitation feature. Therefore, further in-depth analysis of the convection process and surrounding synoptic environment for each cluster will be needed.

5 Summary and conclusion

The characteristics of simulated EREs in South Korea by the summer weather patterns are compared according to the use of cumulus parameterization for the WRF model. Two experiments (CU_ON and CU_OFF) are conducted depending on whether the Kain–Fritsch cumulus scheme is activated for a 3 km resolution mesh domain. Dynamical downscaling with the ERA5 reanalysis data as boundary condition is conducted for 2005–2013.

The ERE is analyzed according to the weather pattern, and the ERE simulation performance of the WRF model is evaluated using the clustering method. The four clusters of 850 hPa geopotential height daily anomaly pattern, which is linked to ERE, are identified, and the WRF experiments capture the observed patterns well. In Cluster 1, the EREs occur as the cold air moves southward by cyclonic circulation and meets with relatively warm air in South Korea, developing atmospheric instability. The ERE of Cluster 2 is caused by the formation of a front when two different air masses meet over South Korea during the summer monsoon period. Cluster 3 ERE occurs due to the extratropical cyclone associated with baroclinic instability. The ERE in Cluster 4 appears to be caused by local instability factors.

The spatial and temporal characteristics of the EREs are analyzed to evaluate the rainfall simulation in various aspects for each cluster. Considering the overall results, CU_ON depicts the rainfall type of Clusters 1 and 2 well, which is featured by continuous low-intensity rainfall. By contrast, CU_OFF can capture the rainfall type of Clusters 3 and 4 where more than 20 mm h^{-1} intensity rainfall with less than 6 h duration occurs relatively frequently. CU_ON shows a relatively high pattern correlation coefficient of ERE and similar rainfall intensity distributions of long-duration ERE cases to the observation for Clusters 1 and 2. CU_ON frequently simulates low-intensity rainfall over a wide area, simulating the ERE of Clusters 1 and 2 better than CU_OFF. CU_OFF shows good performance for the hourly time-scale rainfall of Clusters 3 and 4 compared to CU_ON. CU_OFF can capture the observed features that high-intensity EREs with less than 6 h duration occur more frequently in Clusters 3 and 4 than the other clusters because of its ability to simulate the deep convection explicitly. In the diurnal cycle of Cluster 3, the timing of two peaks and variation are relatively well simulated in CU_OFF compared to CU_ON. Furthermore, the composite of IVT and EPT, which are closely related to rainfall, are analyzed for each cluster. Relatively high IVT and EPT daily anomalies are distributed over South Korea for Clusters 1 and 2, while no statistically significant IVT appears in Clusters 3 and 4. CU_ON (CU_OFF) simulates a similar IVT and EPT pattern to the observations for Clusters 1 and 2 (3 and 4), indicating that the rainfall simulation performance is closely related to the synoptic pattern.

This study focused on analyzing the precipitation characteristics according to the weather patterns and the WRF performance to simulate them depending on cumulus parameterization. The characteristics of each type of ERE described in this study concur with Song and Sohn (2015), where the heavy rainfall over East Asia is classified as warm type or cold type. In addition to identifying the ERE precipitation features according to the weather patterns, it is meaningful to found out that the high-resolution WRF performance varies according to the ERE type depending on the use of the cumulus scheme. CU_OFF would be a more suitable option after a bias correction for the precipitation amount if the model is to be used for real-time forecasting. Because the CU_ON model tends to simulate relatively weak precipitation continuously that is almost identical across all clusters, it is difficult to distinguish the characteristics of precipitation according to the weather pattern. On the other hand, CU_OFF clearly simulates the unique features of precipitation in each cluster despite underestimating the precipitation. Nevertheless, the model used in the comparative experiments is WRF, and the results according to whether only the KF cumulus scheme is applied to the model are compared. The climate model outputs are sensitive to the physical process parameterization and the type of RCM (Giorgi and Shields 1999; Frei et al. 2003). Therefore, more study is needed to determine how the cumulus scheme will improve ERE in the high-resolution climate model simulation by employing other cumulus schemes with different RCMs and a multi-model ensemble approach.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00704-023-04603-0.

Author contribution Ga-Yeong Seo performed model run, analyzed the data, and wrote the manuscript. Joong-Bae Ahn conceptualized the research and revised the manuscript writing. All authors contributed to the manuscript review and editing.

Funding This work was carried out with the support of Rural Development Administration Cooperative Research Program for Agriculture Science and Technology Development under Grant Project No. PJ01475503, Republic of Korea.

Data availability The ASOS station data are available at https://www.data.go.kr/en/data/15043648/fileData.do. The ERA5 data can be downloaded at https://cds.climate.copernicus.eu. The typhoon data are available at https://www.weather.go.kr/w/typhoon/typ-history.do. The Changma period data are available at https://data.kma.go.kr/climate/rainySeason/selectRainySeasonList.do.

Code availability Data analysis and graphics are conducted using Python 3 (https://www.python.org/) and NCAR command language (NCL; https://www.ncl.ucar.edu/). The codes used in this study are available at https://github.com/seogoing/WeatherClust.

Declarations

Competing interests The authors declare no competing interests.

Conflict of interest The authors declare no competing interests.

References

Abadi AM, Rowe CM, Andrade M (2020) Climate regionalization in Bolivia: a combination of non-hierarchical and consensus clustering analyses based on precipitation and temperature. Int J Climatol 40:4408–4421. https://doi.org/10.1002/joc.6464

Afrizal T, Surussavadee C (2018) High-resolution climate simulations in the tropics with complex terrain employing the CESM/WRF model. Adv Meteorol 2018:1–15. https://doi.org/10.1155/2018/5707819

Bae D-H, Jung I-W, Chang H (2008) Long-term trend of precipitation and runoff in Korean river basins. Hydrol Process 22:2644–2656. https://doi.org/10.1002/hyp.6861

Bae H, Ji H, Lim Y-J, Ryu Y, Kim M-H, Kim B-J (2019) Characteristics of drought propagation in South Korea: relationship between meteorological, agricultural, and hydrological droughts. Nat Hazards 99:1–16. https://doi.org/10.1007/s11069-019-03676-3

Ban N, Schmidli J, Schär C (2014) Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. J Geophys Res: Atmos 119:7889–7907. https://doi.org/10.1002/2014JD021478

Betts AK, Miller MJ (1986) A new convective adjustment scheme. Part II: single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. Q J R Meteorol Soc 112:693–709. https://doi.org/10.1002/qj.49711247308

Chattopadhyay A, Hassanzadeh P, Pasha S (2020) Predicting clustered weather patterns: a test case for applications of convolutional neural networks to spatio-temporal climate data. Sci Rep 10:1317. https://doi.org/10.1038/s41598-020-57897-9

Chen F, Dudhia J (2001) Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: model implementation and sensitivity. Mon Weather Rev 129:569–585. https://doi.org/10.1175/1520-0493(2001)129%3c0569:Caalsh%3e2.0.Co;2

Ciarlo JM, Coppola E, Fantini A, Giorgi F, Gao X, Tong Y et al (2021) A new spatially distributed added value index for regional climate models: the EURO-CORDEX and the CORDEX-CORE highest resolution ensembles. Clim Dyn 57:1403–1424. https://doi.org/10.1007/s00382-020-05400-5

Collins WD, Hackney JK, Edwards DP (2002) An updated parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model. J Geophys Res 107:D22. https://doi.org/10.1029/2001jd001365

Deng A, Stauffer DR (2006) On improving 4-km mesoscale model simulations. J Appl Meteorol Climatol 45:361–381. https://doi.org/10.1175/jam2341.1

Diro GT, Tompkins AM, Bi X (2012) Dynamical downscaling of ECMWF Ensemble seasonal forecasts over East Africa with RegCM3. J Geophys Res: Atmos 117:D16103. https://doi.org/10.1029/2011jd016997

Fantini A, Raffaele F, Torma C, Bacer S, Coppola E, Giorgi F et al (2018) Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-CORDEX and EURO-CORDEX experiments against high resolution observations. Clim Dyn 51:877–900. https://doi.org/10.1007/s00382-016-3453-4

Fosser G, Khodayar S, Berg P (2015) Benefit of convection permitting climate model simulations in the representation of convective precipitation. Clim Dyn 44:45–60. https://doi.org/10.1007/s00382-014-2242-1

Fragoso M, Tildes Gomes P (2008) Classification of daily abundant rainfall patterns and associated large-scale atmospheric circulation types in Southern Portugal. Int J Climatol 28:537–544. https://doi.org/10.1002/joc.1564

Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. J Geophys Res: Atmos 108(1–4):73–86. https://doi.org/10.1029/2002JD002287

Gao X, Shi Y, Song R, Giorgi F, Wang Y, Zhang D (2008) Reduction of future monsoon precipitation over China: comparison between a high resolution RCM simulation and the driving GCM. Meteorol Atmos Phys 100:73–86. https://doi.org/10.1007/s00703-008-0296-5

Geetha B, Balachandran S (2016) Diabatic heating and convective asymmetries during rapid intensity changes of tropical cyclones over North Indian Ocean. Trop Cyclone Res Rev 5:32–46. https://doi.org/10.6057/2016TCRRh1.04

Giorgi F, Mearns LO (1991) Approaches to the simulation of regional climate change: a review. Rev Geophys 29:191–216. https://doi.org/10.1029/90RG02636

Giorgi F, Shields C (1999) Tests of precipitation parameterizations available in latest version of NCAR regional climate model (RegCM) over continental United States. J Geophys Res: Atmos 104:6353–6375. https://doi.org/10.1029/98JD01164

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049. https://doi.org/10.1002/qj.3803

Hohenegger C, Brockhaus P, Schär C (2008) Towards climate simulations at cloud-resolving scales. Meteorol Z 17:383–394. https://doi.org/10.1127/0941-2948/2008/0303

Hong J-Y, Ahn J-B (2015) Changes of early summer precipitation in the Korean Peninsula and nearby regions based on RCP simulations. J Clim 28:3557–3578. https://doi.org/10.1175/jcli-d-14-00504.1

Hong S-Y, Dudhia J, Chen S-H (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. Mon Weather Rev 132:103–120. https://doi.org/10.1175/1520-0493(2004)132%3c0103:Aratim%3e2.0.Co;2

Hong S-Y, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev 134:2318–2341. https://doi.org/10.1175/mwr3199.1

Im E-S, Ha S, Qiu L, Hur J, Jo S, Shim K-M (2021) An evaluation of temperature-based agricultural indices over Korea from the high-resolution WRF simulation. Frontiers in Earth Science 9. https://doi.org/10.3389/feart.2021.656787

Ishida K, Tanaka K, Hama T (2019) Sensitivity analysis of convective parameterizations of a regional climate model in higher-resolution domains for long-term precipitation reconstruction. J Water Clim Chang 11:1467–1480. https://doi.org/10.2166/wcc.2019.069

Janjić ZI (1994) The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon Weather Rev 122:927–945. https://doi.org/10.1175/1520-0493(1994)122%3c0927:Tsmecm%3e2.0.Co;2

Jo E, Park C, Son S-W, Roh J-W, Lee G-W, Lee Y-H (2019) Classification of localized heavy rainfall events in South Korea. Asia-Pac J Atmos Sci 56:77–88. https://doi.org/10.1007/s13143-019-00128-7

Jung H-S, Lim G-H, Oh J-H (2001) Interpretation of the transient variations in the time series of precipitation amounts in Seoul, Korea. Part I: diurnal variation. J Clim 14:2989–3004. https://doi.org/10.1175/1520-0442(2001)014%3c2989:Iottvi%3e2.0.Co;2

Jung I-W, Bae D-H, Kim G (2011) Recent trends of mean and extreme precipitation in Korea. Int J Climatol 31:359–370. https://doi.org/10.1002/joc.2068

Kain JS (2004) The Kain-Fritsch convective parameterization: an update. J Appl Meteorol 43:170–181. https://doi.org/10.1175/1520-0450(2004)043%3c0170:Tkcpau%3e2.0.Co;2

Karmacharya J, Jones R, Moufouma-Okia W, New M (2017) Evaluation of the added value of a high-resolution regional climate model simulation of the South Asian summer monsoon climatology. Int J Climatol 37:3630–3643. https://doi.org/10.1002/joc.4944

Kim W, Jhun J-G, Ha K-J, Kimoto M (2011) Decadal changes in climatological intraseasonal fluctuation of subseasonal evolution of summer precipitation over the Korean Peninsula in the mid-1990s. Adv Atmos Sci 28:591–600. https://doi.org/10.1007/s00376-010-0037-9

Kim H-K, Moon B-K, Kim M-K, Park J-Y, Hyun Y-K (2021) Three distinct atmospheric circulation patterns associated with high temperature extremes in South Korea. Sci Rep 11:12911. https://doi.org/10.1038/s41598-021-92368-9

Konduru RT, Takahashi HG (2020) Effects of convection representation and model resolution on diurnal precipitation cycle over the Indian monsoon region: toward a convection-permitting regional climate simulation. J Geophys Res: Atmos 125:e2019JD032150. https://doi.org/10.1029/2019JD032150

Kotroni V, Lagouvardos K (2004) Evaluation of MM5 high-resolution real-time forecasts over the urban area of Athens, Greece. J Appl Meteorol 43:1666–1678. https://doi.org/10.1175/jam2170.1

Lee D-K, Park J-G, Kim J-W (2008) Heavy rainfall events lasting 18 days from July 31 to August 17, 1998, over Korea. Journal of the Meteorological Society of Japan. Ser II 86:313–333. https://doi.org/10.2151/jmsj.86.313

Lee S-W, Lee D-K, Chang D-E (2011) Impact of horizontal resolution and cumulus parameterization scheme on the simulation of heavy rainfall events over the Korean Peninsula. Adv Atmos Sci 28:1–15. https://doi.org/10.1007/s00376-010-9217-x

Lee J-Y, Kwon M, Yun K-S, Min S-K, Park I-H, Ham Y-G et al (2017) The long-term variability of Changma in the East Asian summer monsoon system: a review and revisit. Asia-Pac J Atmos Sci 53:257–272. https://doi.org/10.1007/s13143-017-0032-5

Lee M-H, Bae D-H, Im E-S (2019) Effect of the horizontal resolution of climate simulations on the hydrological representation of extreme low and high flows. Water Resour Manage 33:4653–4666. https://doi.org/10.1007/s11269-019-02359-9

Liu J, Bray M, Han D (2012) Sensitivity of the Weather Research and Forecasting (WRF) model to downscaling ratios and storm types in rainfall simulation. Hydrol Process 26:3012–3031. https://doi.org/10.1002/hyp.8247

Lucas-Picher P, Argüeso D, Brisson E, Tramblay Y, Berg P, Lemonsu A et al (2021) Convection-permitting modeling with regional climate models: latest developments and next steps. WIREs Clim Chang 12:e731. https://doi.org/10.1002/wcc.731

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Oakland, CA, USA, 281–297.

Mayor YG, Mesquita MDS (2015) Numerical simulations of the 1 May 2012 deep convection event over Cuba: sensitivity to cumulus and microphysical schemes in a high-resolution model. Adv Meteorol 2015:973151. https://doi.org/10.1155/2015/973151

Meredith EP, Maraun D, Semenov VA, Park W (2015) Evidence for added value of convection-permitting models for studying changes in extreme precipitation. J Geophys Res: Atmos 120:12500–12513. https://doi.org/10.1002/2015JD024238

Miyasaka T, Kawase H, Nakaegawa T, Imada Y, Takayabu I (2020) Future projections of heavy precipitation in Kanto and associated weather patterns using large ensemble high-resolution simulations. Sola 16:125–131. https://doi.org/10.2151/sola.2020-022

Nguyen-Le D, Yamada TJ (2019) Using weather pattern recognition to classify and predict summertime heavy rainfall occurrence over the Upper Nan River Basin, northwestern Thailand. Weather Forecast 34:345–360. https://doi.org/10.1175/waf-d-18-0122.1

NIMS (2014) Improvement of prediction skills in very short, short and mid-term for severe weather. National Institute of Meteorological Research, South Korea. https://doi.org/10.23000/TRKO201500014009

Park C, Min S-K, Lee D, Cha D-H, Suh M-S, Kang H-S et al (2016) Evaluation of multiple regional climate models for summer climate extremes over East Asia. Clim Dyn 46:2469–2486. https://doi.org/10.1007/s00382-015-2713-z

Park C, Son S-W, Kim J-H (2021a) Role of baroclinic trough in triggering vertical motion during summertime heavy rainfall events in Korea. J Atmos Sci 78:1687–1702. https://doi.org/10.1175/jas-d-20-0216.1

Park C, Son S-W, Kim J, Chang E-C, Kim J-H, Jo E et al (2021b) Diverse synoptic weather patterns of warm-season heavy rainfall events in South Korea. Mon Weather Rev 149:3875–3893. https://doi.org/10.1175/mwr-d-20-0388.1

Park C, Shin S-W, Cha D-H, Suh M-S, Hong S-Y, Ahn J-B et al (2022a) Future projections of precipitation using bias–corrected high–resolution regional climate models for sub–regions with homogeneous characteristics in South Korea. Asia-Pac J Atmos Sci 58:715–727. https://doi.org/10.1007/s13143-022-00292-3

Park H, Kim G, Cha D-H, Chang E-C, Kim J, Park S-H et al (2022b) Effect of a scale-aware convective parameterization scheme on the simulation of convective cells-related heavy rainfall in South Korea. J Adv Model Earth Syst 14:e2021MS002696.https://doi.org/10.1029/2021MS002696

Pennelly C, Reuter G, Flesch T (2014) Verification of the WRF model for simulating heavy precipitation in Alberta. Atmos Res 135–136:172–192. https://doi.org/10.1016/j.atmosres.2013.09.004

Prein AF, Gobiet A, Suklitsch M, Truhetz H, Awan NK, Keuler K et al (2013) Added value of convection permitting seasonal simulations. Clim Dyn 41:2655–2677. https://doi.org/10.1007/s00382-013-1744-6

Prein AF, Langhans W, Fosser G, Ferrone A, Ban N, Goergen K et al (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. Rev Geophys 53:323–361. https://doi.org/10.1002/2014RG000475

Qiu L, Im E-S, Hur J, Shim K-M (2020) Added value of very high resolution climate simulations over South Korea using WRF modeling system. Clim Dyn 54:173–189. https://doi.org/10.1007/s00382-019-04992-x

Rauscher SA, Coppola E, Piani C, Giorgi F (2010) Resolution effects on regional climate model simulations of seasonal precipitation over Europe. Clim Dyn 35:685–711. https://doi.org/10.1007/s00382-009-0607-7

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Seo GY, Ahn JB (2020) Sensitivity analysis of cumulus parameterization in WRF model for simulating summer heavy rainfall in South Korea. J Clim Res 15:243–256. https://doi.org/10.14383/cri.2020.15.4.243

Seo KH, Son JH, Lee JY, Park HS (2015) Northern East Asian monsoon precipitation revealed by airmass variability and its prediction. J Clim 28:6221–6233. https://doi.org/10.1175/jcli-d-14-00526.1

Skamarock C, Klemp B, Dudhia J, Gill O, Liu Z, Berner J et al. (2019) A description of the advanced research WRF model version 4. https://doi.org/10.6084/m9.figshare.7369994.v4

Song H-J, Sohn B-J (2015) Two heavy rainfall types over the Korean Peninsula in the humid East Asian summer environment: a satellite observation study. Mon Weather Rev 143:363–382. https://doi.org/10.1175/mwr-d-14-00184.1

Syakur MA, Khotimah BK, Rochman EMS, Satoto BD (2018) Integration K-means clustering method and elbow method for identification of the best customer profile cluster. IOP Conf Ser: Mater Sci Eng 336:012017. https://doi.org/10.1088/1757-899x/336/1/012017

Tao W-K, Simpson J, McCumber M (1989) An ice-water saturation adjustment. Mon Weather Rev 117:231–235. https://doi.org/10.1175/1520-0493(1989)117%3c0231:Aiwsa%3e2.0.Co;2

Tao Y, Cao J, Lan G, Su Q (2016) The zonal movement of the Indian-East Asian summer monsoon interface in relation to the land–sea thermal contrast anomaly over East Asia. Clim Dyn 46:2759–2771. https://doi.org/10.1007/s00382-015-2729-4

Torma C, Giorgi F, Coppola E (2015) Added value of regional climate modeling over areas characterized by complex terrain-precipitation over the Alps. J Geophys Res: Atmos 120:3957–3972. https://doi.org/10.1002/2014JD022781

Wang S, Yu E, Wang H (2012) A simulation study of a heavy rainfall process over the Yangtze River valley using the two-way nesting approach. Adv Atmos Sci 29:731–743. https://doi.org/10.1007/s00376-012-1176-y

Warner TT, Hsu H-M (2000) Nested-model simulation of moist convection: the impact of coarse-grid parameterized convection on fine-grid resolved convection. Mon Weather Rev 128:2211–2231. https://doi.org/10.1175/1520-0493(2000)128%3c2211:NMSOMC%3e2.0.CO;2

Weusthoff T, Ament F, Arpagaus M, Rotach MW (2010) Assessing the benefits of convection-permitting models by neighborhood verification: examples from MAP D-PHASE. Mon Weather Rev 138:3418–3433. https://doi.org/10.1175/2010mwr3380.1

Wootten A, Bowden JH, Boyles R, Terando A (2016) The sensitivity of WRF downscaled precipitation in Puerto Rico to cumulus parameterization and interior grid nudging. J Appl Meteorol Climatol 55:2263–2281. https://doi.org/10.1175/jamc-d-16-0121.1

Xu Z, Han Y, Yang Z (2019) Dynamical downscaling of regional climate: a review of methods and limitations. Sci China Earth Sci 62:365–375. https://doi.org/10.1007/s11430-018-9261-5

Yu X, Lee T-Y (2011) Role of convective parameterization in simulations of heavy precipitation systems at grey-zone resolutions — case studies. Asia-Pac J Atmos Sci 47:99–112. https://doi.org/10.1007/s13143-011-0001-3

Zhao S, Deng Y, Black RX (2017) A dynamical and statistical characterization of U.S. extreme precipitation events and their associated large-scale meteorological patterns. J Clim 30:1307–1326. https://doi.org/10.1175/JCLI-D-15-0910.1

Zhao Y, Zhou T, Li P, Furtado K, Zou L (2021) Added value of a convection permitting model in simulating atmospheric water cycle over the Asian Water Tower. J Geophys Res: Atmos 126:e2021JD034788. https://doi.org/10.1029/2021JD034788

Zhao Y, Deng L, Li Z, Wang Y (2022) Quantitative attribution of vertical motion responsible for summer heavy rainfall over North China. J Geophys Res: Atmos 127:e2021JD035765. https://doi.org/10.1029/2021JD035765