Asia-Pacific Journal of Atmospheric Sciences
https://doi.org/10.1007/s13143-022-00310-4

Online ISSN 1976-7951
Print ISSN 1976-7633

ORIGINAL ARTICLE

Evaluation of Deep-Learning-Based Very Short-Term Rainfall Forecasts in South Korea

Seok-Geun Oh[1] · Chanil Park[1] · Seok-Woo Son[1] · Jihoon Ko[2] · Kijung Shin[2] · Sunyoung Kim[3] · Junsang Park[3]

Received: 26 July 2022 / Revised: 16 November 2022 / Accepted: 4 December 2022

Abstract

This study evaluates the performance of a deep learning model, Deep-learning-based Rain Nowcasting and Estimation (DEEPRANE), for very short-term (1–6 h) rainfall forecasts in South Korea. Rainfall forecasts and in-situ observations from June–September 2020, when record-breaking summer rainfall was observed in South Korea, are particularly considered. It is found that DEEPRANE adequately predicts moderate rainfall events (MREs; ≥ 1 mm h^{-1}) and strong rainfall events (SREs; ≥ 10 mm h^{-1}) with critical success indices of 0.6 and 0.4 at the 1-h lead time, respectively. The probability of detection scores of MRE forecasting is higher than the false alarm rates at all lead times, suggesting that DEEPRANE MRE forecast can be useful even at a long lead time. However, for SRE forecasting, the probability of detection scores becomes smaller than the false alarm rates at a lead time of 2 h. Localized heavy rainfall events (LHREs, ≥ 30 mm h^{-1}) are also reasonably well predicted only at a lead time of 2 h. Irrespective of their patterns, the forecast scores systematically decrease with lead time. This result indicates that DEEPRANE SRE forecast is useful only for nowcasting. DEEPRANE generally shows better performance in the early morning hours when rainfall events are more frequent than in other hours. When considering synoptic conditions, better performance is found when rainfall events are organized by monsoon rainband rather than caused by extratropical or tropical cyclones. These results suggest that DEEPRANE is especially useful for nowcasting early-morning rainfall events which are embedded in the monsoon rainband.

Keywords Deep learning model · Very short-term rainfall forecast · Localized heavy rainfall · South Korea

1 Introduction

Rainfall forecasts with accurate, reliable location information and timing are crucial for preventing life and financial losses. In June–September 2020, extraordinary heavy rainfall events have repeatedly occurred over East Asia (e.g., Hirockawa et al. 2020; Liu et al. 2020; Araki et al. 2021; Park et al. 2021a). South Korea was not exceptional. The 2020 summer rainfall in South Korea broke a record dating back to 1971 and resulted in destructive floods and thousands of flood victims across the country (Park et al. 2021a). Through these events, the importance of short-term rainfall forecasts for protecting public safety was highlighted.

The numerical weather prediction (NWP) model is a widely-used tool for rainfall forecasting (e.g., Shahrban et al. 2016; Pu and Kalnay 2018). With recent advances in computer technology, NWP models have been improved in terms of the dynamic core, physics parameterization, and data assimilation, resulting in considerable improvement in rainfall forecasts (e.g., Shuman 1989; Harper et al. 2007; Shahrban et al. 2016; Pu and Kalnay 2018). However, very short-term rainfall forecasts, which are generally considered to have lead times of 1–6 h, still remain challenging. For example, Shrestha et al. (2013) used four NWP models to evaluate rainfall forecasts over the Ovens catchment in southeast Australia. They reported large wet or dry biases, depending on the model resolution, which ultimately degraded the rainfall forecasting performance for short lead times on an

Responsible editor: Yong-Sang Choi

✉ Seok-Woo Son
seokwooson@snu.ac.kr

[1] School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea

[2] Kim Jaechul Graduate School of AI, KAIST, Seoul, South Korea

[3] National Institute of Meteorological Sciences, Jeju, South Korea

Published online: 08 December 2022

hourly scale. Wang et al. (2016) applied a storm-scale NWP model to five heavy rainfall events in summer over Jiangsu, China, in 2015. Their NWP model overestimated rainfall amounts at all lead times and failed to produce optimal forecasts at short lead times, i.e., 1–2 h. Yu et al. (2017) evaluated two NWP models for three summertime rainfall events over Sancheong Basin in South Korea. They reported high performance for wide-area rainfall such as that from tropical cyclones (TCs) but limited forecasting skill for localized rainfall, for example, along a quasi-stationary monsoon front.

Various attempts have been made to improve very short-term rainfall forecasts. Lin et al. (2005) and Sun et al. (2014), for example, reported that radar-based short-term statistical rainfall forecasts could outperform the NWP model at lead times of 2–3 h. Yoon (2019) proposed the use of blending techniques in statistical and NWP model-based rainfall forecasts. Another approach is artificial-intelligence-based forecasting (e.g., Shi et al. 2015; Agrawal et al. 2019; Ayzel et al. 2020). Shi et al. (2015) first used a deep learning model for radar-based short-term rainfall forecasting. Agrawal et al. (2019) demonstrated that a deep learning model, based on the U-Net convolutional neural network (CNN), has the potential to improve very short-term rainfall forecasts, specifically, nowcasting up to 2 h ahead, compared to the National Oceanic and Atmospheric Administration numerical 1-h High-Resolution Rapid Refresh model. Ayzel et al. (2020) reported that RainNet, which was inspired by U-Net-based deep learning models, significantly outperforms the benchmark models in terms of the forecast accuracy at lead times of up to 1 h for 11 summer rainfall events over Germany. Ravuri et al. (2021) proposed a deep generative model of radar for probabilistic nowcasting with lead times of 5–90 min. These studies suggest that deep learning models can be useful for predicting very short-term rainfall events and offer an alternative to the NWP model.

Several deep learning algorithms have been proposed to improve very short-term rainfall forecasts in Asia. Yen et al. (2019), in an evaluation of rainfall forecasting in southern Taiwan, proposed the Deep Echo State Network algorithm. Zhang et al. (2021) proposed a dual-input dual-encoder recurrent neural network, the Rainfall Nowcasting Network. Their deep learning model yields better threat scores in rainfall forecasts of 0.25 mm per 30 min in southeastern China at a lead time of 2 h than the Weather Research and Forecasting model. However, all of these studies indicate that further research is necessary to generalize and improve deep learning models in East Asia because East Asian summer monsoon rainfalls are complex in nature due to the topography, subseasonal variation, and other factors.

Although deep-learning-based rainfall forecasting has been attempted, such approach is still lacking in South Korea where more than half of the annual precipitation is concentrated in the monsoon season with complex meteorological factors (Park et al. 2021b; 2021c). Yoon et al. (2020) developed a CNN algorithm based on U-Net and SegNet by training with radar image data. However, they used spatially limited radar data covering only part of South Korea. The forecast lead time was also limited to only 1 h. Kim and Hong (2022) attempted very short-term prediction of the radar-based rainfall distribution and intensity using convolutional long-short-term memory and demonstrated its effectiveness compared to Korea's operational prediction model. However, the forecast lead time was still limited to a maximum of 2.5 h, which is insufficient to prepare for heavy-rainfall-related natural hazards.

Ko et al. (2022) recently proposed a novel deep learning model for very short-term rainfall forecasts, DEEP-learning-based RAin Nowcasting and Estimation (DEEPRANE), which is optimized for summer monsoon rainfalls in Korea, so-called Changma rainfalls. They particularly proposed a pre-training scheme and a new loss function and demonstrated their effectiveness in improving accuracy in rainfall forecasts, compared to models without a pre-training scheme and/or with conventional loss functions. They also proposed that DEEPRANE may be useful for real-time rainfall forecasts at lead times of up to 6 h.

As a companion work to Ko et al. (2022), the present study aims to thoroughly evaluate the performance of DEEPRANE in very short-term rainfall forecasting in summer monsoon season in South Korea. We evaluate the forecast skill of DEEPRANE at different lead times (1 to 6 h), and its regional, subseasonal, and diurnal dependences for two classes of rainfall events: moderate rainfall events (MREs, ≥ 1 mm h^{-1}) and strong rainfall events (SREs, ≥ 10 mmh^{-1}). The June–September 2020 season, when record-breaking summer rainfall was reported in South Korea (Park et al. 2021a), is of particular interest. The performance of DEEPRANE is also tested for five types of localized heavy rainfall events (LHREs; ≥ 30 mm h^{-1}) in June–September 2020.

The rest of this paper is organized as follows. Section 2 briefly describes the model and method used in this study. Section 3 presents the performance of the deep learning model, DEEPRANE, and its regional, subseasonal, and diurnal dependences. In Section 4, an overview of LHRE classification and the performance of DEEPRANE for each type are presented. The summary and discussion are presented in Section 5.

2 Model and Method

2.1 Deep Learning Model

This study uses DEEPRANE, in which a deep learning model is equipped with a pre-training scheme and a new loss function (Ko et al. 2022). This model is based on U-Net

(Ronneberger et al. 2015), which is a widely-used deep CNN. Therefore, the model consists of a contracting path, which reduces the resolution of the input images, and an expanding path, which expands again the contracted images, as illustrated in Fig. 1. In the contracting path, i.e., downsampling steps, the resolution of the input data is halved by merging each four incident pixels, while the number of latent features per pixel, which means the dimension of the vector representing each pixel, is doubled. In the expanding path, i.e., upsampling steps, the resolution of the input data is doubled, while the number of latent features per pixel is halved. In addition, skip connections are used between the contracting path and expanding path. These connections match the downsampling steps and upsampling steps one-to-one so that the intermediate output of each downsampling step, which indicates the vector representation of each pixel, is additionally considered to determine the output of the corresponding upsampling step.

Compared to the conventional U-Net, this model has two improved components (Ko et al. 2022). One is the pre-training process before fine-tuning to optimize the parameters for predicting the radar reflectivity used for rainfall forecasts. In the pre-training process, the model is trained for unsupervised tasks and/or supervised tasks with abundant labels. Through this process, the model is initialized with pre-trained parameters before the fine-tuning process. To pre-train as many parameters as possible, the same model architecture for pre-training and fine-tuning is used except for the last classifier part. The classifier part cannot be pre-trained since the number of output classes (i.e., the dimensionality of the classifier) in the pre-training process differs from that in the fine-tuning process. Ko et al. (2022) demonstrated that deep-learning-based rainfall forecasts obtained using the pre-training scheme significantly improve the critical success index (CSI) and F1 scores at lead times of 1–6 h compared to those without it.

The other component is a new loss function based on the CSI scores for fine-tuning, which could mitigate the performance degradation due to class imbalance in rainfall data. When typical classification loss functions (e.g., the cross-entropy loss) are used, classifiers are easily biased toward non-rainfall and light rainfall events during training since most rainfall events belong to the non- and light-rainfall classes. Thus, it is one of the challenges in improving

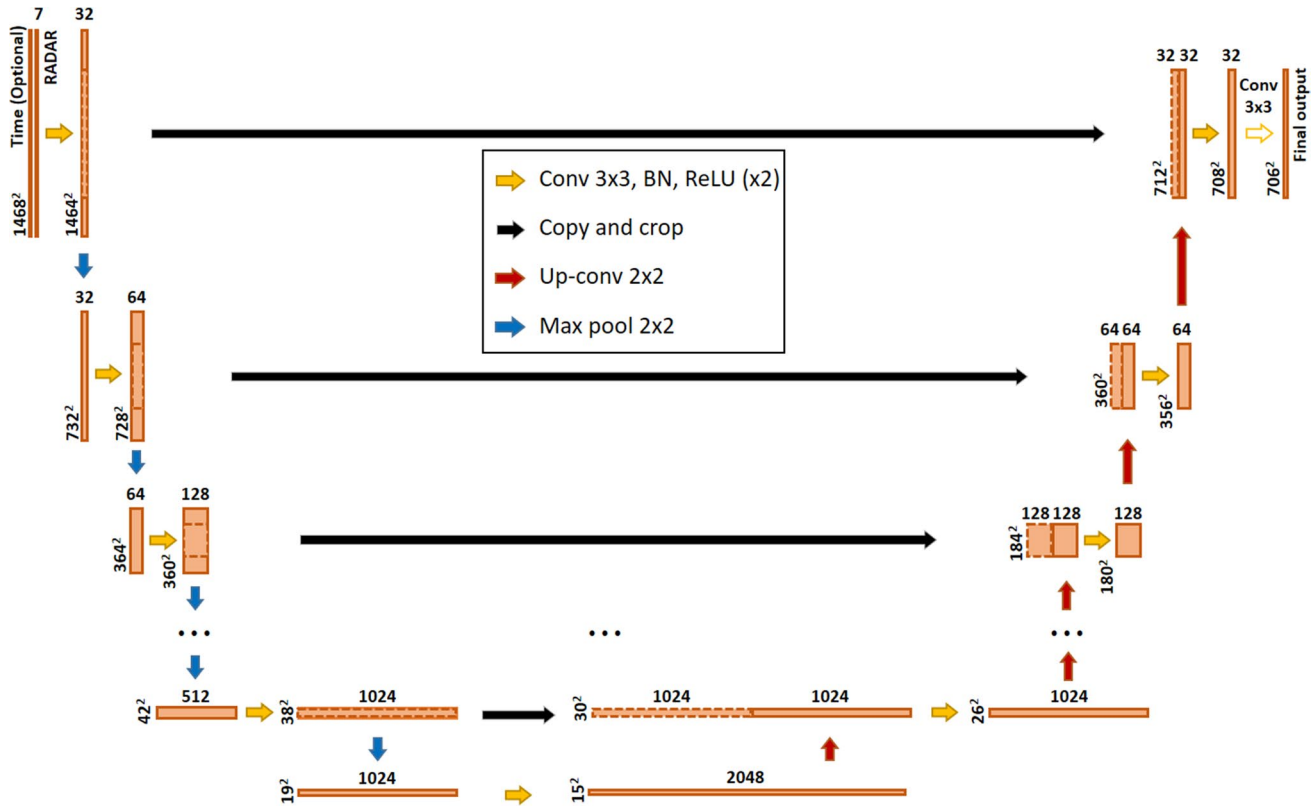

Fig. 1 Configuration of the deep-learning model, DEEPRANE, used in this study, where U-Net is tailored for the problems of the short-term rainfall forecast. This model consists of seven steps for downsampling (blue arrows) and upsampling (red arrows), respectively, to make the width and height of the coarsest feature map roughly. Input process has seven channels corresponding to radar images and additional six channels for encoding the target time. Note that the input and output dimensions are 1468×1468 and 706×706, respectively, at 1 km resolution

predictive performance for the heavy rainfall class which has a smaller proportion than non- and light-rainfall classes. To mitigate the issue, Ko et al. (2022) proposed a new loss function using the true positives (Hit in Table 1), false positives (False alarm), and false negatives (Miss) which are used to calculate the CSI score, based on the output probability distributions over the rain classes from the U-Net model. They showed that the new loss function allows the deep learning model to achieve better predictive performance at relatively long lead times, 3–6 h, than the widely used cross-entropy loss (Cox 1958) and focal loss (Lin et al. 2017). A detailed description of DEEPRANE can be found in Ko et al. (2022).

2.2 Model Design

Figure 1 shows the configuration of DEEPRANE. The input process consists of seven channels for the input radar reflectivity images and six additional channels for encoding the target times. For each timestamp t, seven latest consecutive radar images at 10-min intervals (i.e., radar images with timestamps $t-60, t-50, \ldots, t$) are used for the first seven channels. Six additional binary channels are then used to encode a target time t^* among $\{t+60, t+120, \ldots, t+360\}$. The input and output dimensions are set to 1468×1468 and 706×706, respectively, for radar images with 1-km resolution. The final output of the model is a probability distribution for the three rainfall classes, < 1 mm h^{-1}, 1–10 mm h^{-1}, and > 10 mm h^{-1}, for each pixel according to different lead times from 1 to 6 h.

As shown in Fig. 1, DEEPRANE processes the inputs using four main operations: i.e., 3×3 convolution, copy and crop, max pooling, and up-convolution. The 3×3 convolution operations compute output feature maps using a rectangular receptive field of size 3×3 around each pixel of the input feature maps. The copy and crop operations combine information from the output after each up-convolution operation in expansive paths and the corresponding intermediate feature map in contracting paths. Since the sizes of feature maps are different and the latter feature maps are always bigger, DEEPRANE crops the latter feature maps after aligning the centers of the feature maps. The max pooling and up-convolution operations change the size of intermediate feature maps. In contracting paths, max pooling operations are used to halve the size of feature maps by splitting the feature maps into 2×2 blocks and picking the maximum value in each block. In expanding paths, up-convolution operations are used for up-sampling feature maps so that the sizes of the feature maps are doubled.

To train the model, the radar reflectivity images around South Korea (roughly 120–138°E and 29–42°N) at 10 min intervals are collected for the period of 2014–2020 (Fig. 2a). The rainfall observations at approximately 714 automatic weather stations (AWSs) are also collected during the same period (Fig. 2b). All available datasets for the period of 2014–2018 are used for pre-training. Then, only the June–September rainfall datasets are used for fine-tuning to optimize the model parameters for summer rainfalls in Korea. Ko et al. (2022) applied this optimized model to the 2019 summer season for validation, and then tested it for four heavy rainfall events in 2020. They reported CSI scores of 0.61 to 0.35 in the summer rainfall forecasts for MREs with > 1 mm h^{-1} at lead times of 1 to 6 h in validation. The corresponding CSI scores for SREs with > 10 mm h^{-1} range from 0.39 to 0.12.

Table 1 A 2×2 contingency table showing four possible outcomes between forecasts and observations

Event forecast	Event Observed		
	Yes	No	Total
Yes	H (Hit)	F (False alarm)	H+F
No	M (Miss)	C (Correct negative)	M+C
Total	H+M	F+C	H+F+M+C

2.3 Evaluation Matrix and Rainfall Events

This study uses AWS hourly rainfall data for summer (June–September) 2020 to evaluate the performance of DEEPRANE trained by radar reflectivity images. For one-to-one comparison, the nearest grid of the DEEPRANE output (1-km resolution) is interpolated to each AWS location. Four verification indices, which are the probability of detection (POD), false alarm rate (FAR), CSI, and bias score (BS), are computed, as summarized in Tables 1 and 2. Here, POD denotes the fraction of correctly forecasted rainfall events to the observed events. A perfect score is 1, and the range is 0 to 1. The FAR is the fraction of incorrectly forecasted rainfall events to all forecasted rainfall events, ranging from 0 to 1, with a perfect score of 0. The CSI measures the fraction of correctly forecasted rainfall events to the total observed events except for correct negatives. This score is unaffected by the number of non-event forecasts and is therefore widely used to verify weather forecasts. It ranges from 0 to 1 with a perfect score of 1. The BS is estimated as the ratio of the total number of forecasted rainfall events to the observed rainfall events. A perfect score is 1, and its range is 0 to ∞.

We evaluate two rainfall classes: MREs, with rainfall intensity greater than 1 mm h^{-1}, and SREs, with intensity greater than 10 mm h^{-1}. Note that MREs include SREs to get more samples. The performance of DEEPRANE is quantified for these two rainfall classes at lead times of 1 to 6 h and their regional, subseasonal, and diurnal dependences using the evaluation matrix described above. A performance diagram (Roebber

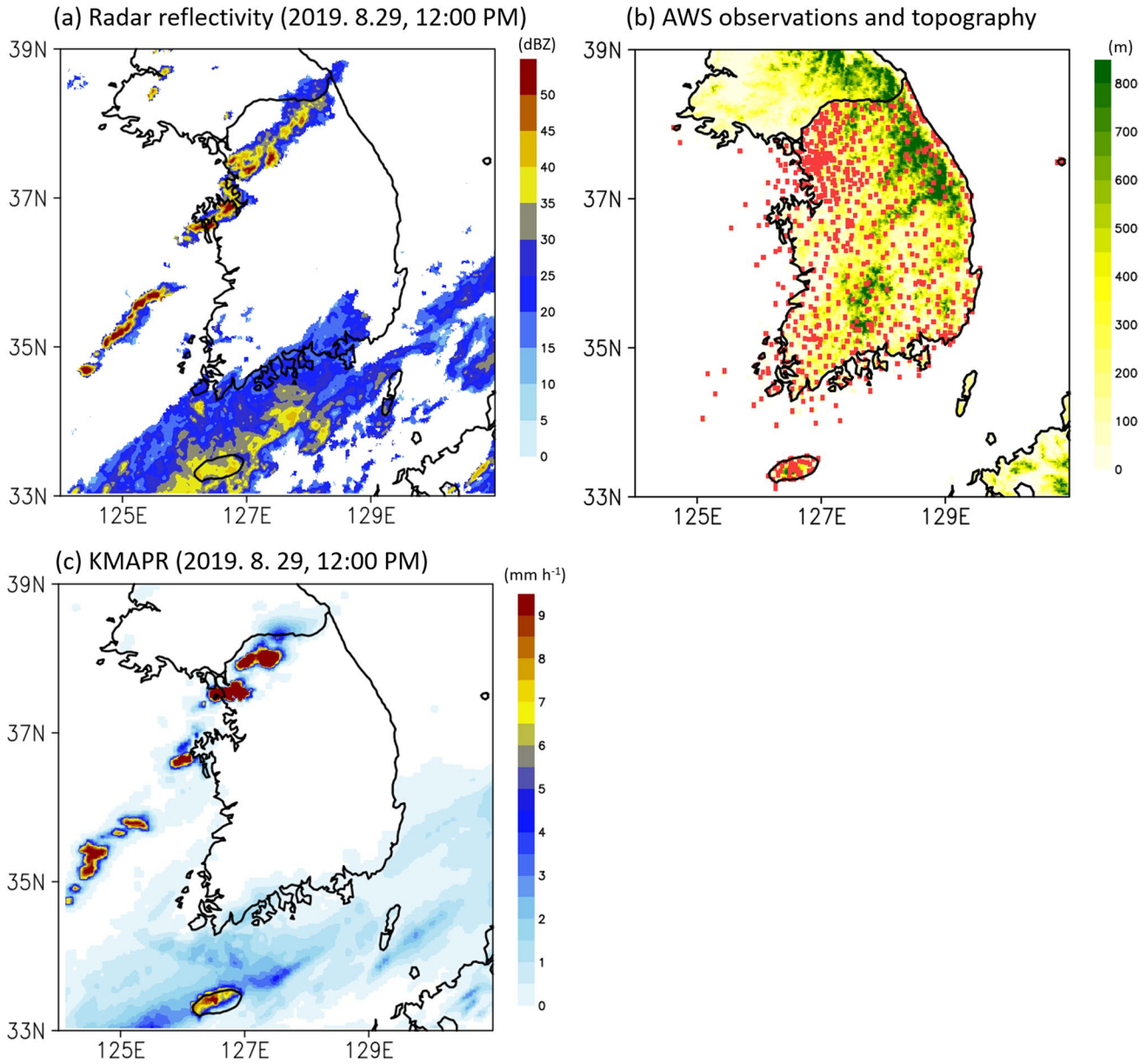

Fig. 2 (**a**) An example of radar reflectivity image used in training. (**b**) Spatial distribution of AWS observations in South Korea and topography (m). (**c**) An example of KMA precipitation reanalysis data (KMAPR). Note that radar reflectivity is a snapshot with a 10-min interval and KMAPR is the 1-h accumulated precipitation produced by synthesizing the AWS precipitation and radar reflectivity

Table 2 Four verification indices based on a 2×2 contingency table used in this study

Abbreviation	Full name	Formulation	Perfect score
POD	Probability of detection	$POD = \frac{H}{H+M}$	1
FAR	False alarm rate	$FAR = \frac{F}{H+F}$	0
CSI	Critical success index	$CSI = \frac{H}{H+M+F}$	1
BS	Bias score	$BS = \frac{H+F}{H+M}$	1

2009), which enables comprehensive performance evaluation by representing four verification indices in a single diagram, is also used. The bootstrapping method, which repeats 1000 resamplings and evaluations from the verification data, is applied to estimate the sampling uncertainty from the verification value in the performance diagram (Roebber 2009).

The LHREs, with rainfall intensity greater than 30 mm h^{-1} corresponding to the localized heavy rainfall

advisory standard of KMA, are also considered separately. All events satisfying this criterion are counted without considering event separation, indicating that multiple LHREs can occur under the same synoptic condition. This allows a total of 429 LHREs in the evaluation period.

The LHREs are further classified using the self-organization map (SOM) algorithm to evaluate the forecasting skills of LHREs depending on their types. The SOM is a sort of unsupervised neural network algorithm that compresses high-dimensional data into a manageable low-dimensional array of nodes based on the Euclidean distance by iterative training (Kohonen 1998; 2013) and has been used to investigate the complex characteristics of warm-season heavy rainfall in South Korea (Jo et al. 2020; Park et al. 2021b).

As a clustering property, we use the KMA hourly precipitation reanalysis (KMAPR) data corresponding to the LHREs. The KMAPR data are produced by synthesizing the precipitation data recorded at AWS stations and radar reflectivity (Fig. 2c; Roh et al. 2012; Jo et al. 2020). It provides high-resolution (5 km × 5 km) rainfall maps covering South Korea and its immediate surroundings every hour, allowing categorization of LHREs (e.g., Jo et al. 2020). Node arrangement is the most important parameter in SOM algorithm. After sensitivity tests with varying size of nodes, we find that 1 × 5 is most proper to summarize 429 LHREs. Other SOM parameters specified in this study follow the suggestion of Liu et al. (2016) and summarized in Table S1.

3 Moderate Rainfall Event (MRE) and Strong Rainfall Event (SRE) Forecastings

3.1 Regional Features

Figures 3A and 4a show the spatial distributions of MRE and SRE frequencies during June–September 2020 in South Korea. The validation results of DEEPRANE at lead times of 1, 3, and 6 h are also presented in Figs. 3b–g and 4b–g. The southern inland and northeast coastal regions, which are mountainous (see Fig. 2b), experienced more frequent MREs and SREs, possibly because of topographic forcing on the upwind side (Figs. 3a and 4a). In these regions, more than 350 MREs and 50 SREs (approximately 1.5 to 2 times more frequent than other regions) were observed. Compared to MREs, SREs were more concentrated in the western half of the country.

For MRE forecasting, DEEPRANE shows CSI scores greater than 0.6 at all AWS stations at a lead time of 1 h (Fig. 3b). These skill scores are higher than those of the deep learning algorithm of Yoon et al. (2020) for MRE forecasting. Although the CSI scores tend to decrease with increasing lead time, they are still 0.4 or higher at a lead time of 6 h (Fig. 3c and d). This result indicates that DEEPRANE can adequately predict MREs and their regionality. The BS results indicate that DEEPRANE tends to over-predict MREs by 1.2 to 1.6 times at all lead times. The over-prediction becomes larger with increasing lead time (Figs. 3e–g). This overprediction may be partly related to the model being set up to an aggressive blurring so that the loss function based on the CSI score could be optimized in the training process.

For SRE forecasting, the CSI scores have spatiotemporal distributions similar to those for MRE forecasting. However, they show poorer performance than MRE forecasting at all lead times. At a lead time of 1 h, the CSI scores are 0.3–0.5 (Fig. 4b). They decrease rapidly with lead time. At lead times of 3 h or longer, CSI scores of 0.1–0.3 are found at most AWS stations (Figs. 4c and d). Such a poor performance is due in part to the significant over-prediction at lead times of 3 h and longer (Figs. 4f and g). At these lead times, the BS score exceeds 2 at many metropolitan stations (northwestern part of South Korea), suggesting stronger over-prediction in this region. Note that the BS at a lead time of 1 h ranges from 0.5 to 0.8 at most AWS stations (Fig. 4e). This indicates that the performance of DEEPRANE is sensitive to the lead time when forecasting SREs, and DEEPRANE is applicable only for very short-term SRE forecasting.

This result is consistent with previous studies (e.g., Agrawal et al. 2019; Ayzel et al. 2020; Han et al. 2020) that reported that deep learning models have limitations in predicting the occurrence number and timing of SREs at relatively long lead times. Such limitations are largely attributable to the use of insufficient samples based on only radar data to train the algorithm (e.g., Schultz et al. 2021; Kim and Hong 2022). As summer 2020 was very unusual with record-breaking heavy rainfall in South Korea (Park et al. 2021a), a poorer performance of DEEPRANE for SRE forecasting which was trained with the 2014–2018 dataset is rather natural. This issue could be resolved by increasing the sample size based on various input data, such as radar, satellite, and NWP simulation data, in a future study.

3.2 Subseasonal and Diurnal Features

Figures 5A and d show the subseasonal variation of MREs and SREs aggregated across all AWS stations on a given day. The CSI and BS of DEEPRANE are also presented for MRE (Figs. 5b and c) and SRE forecastings (Figs. 5e and f) as a function of lead time. The days with no rainfall events, i.e., no rain across the country, are indicated by gray shading in the validation plots. The summer 2020 rainfall in South Korea resulted from consecutive rainfall systems under varying synoptic conditions (Park et al. 2021a). The rainfall events in mid-June–late July (red shading in Figs. 5a and d) were dominated by extratropical cyclones (ETCs)

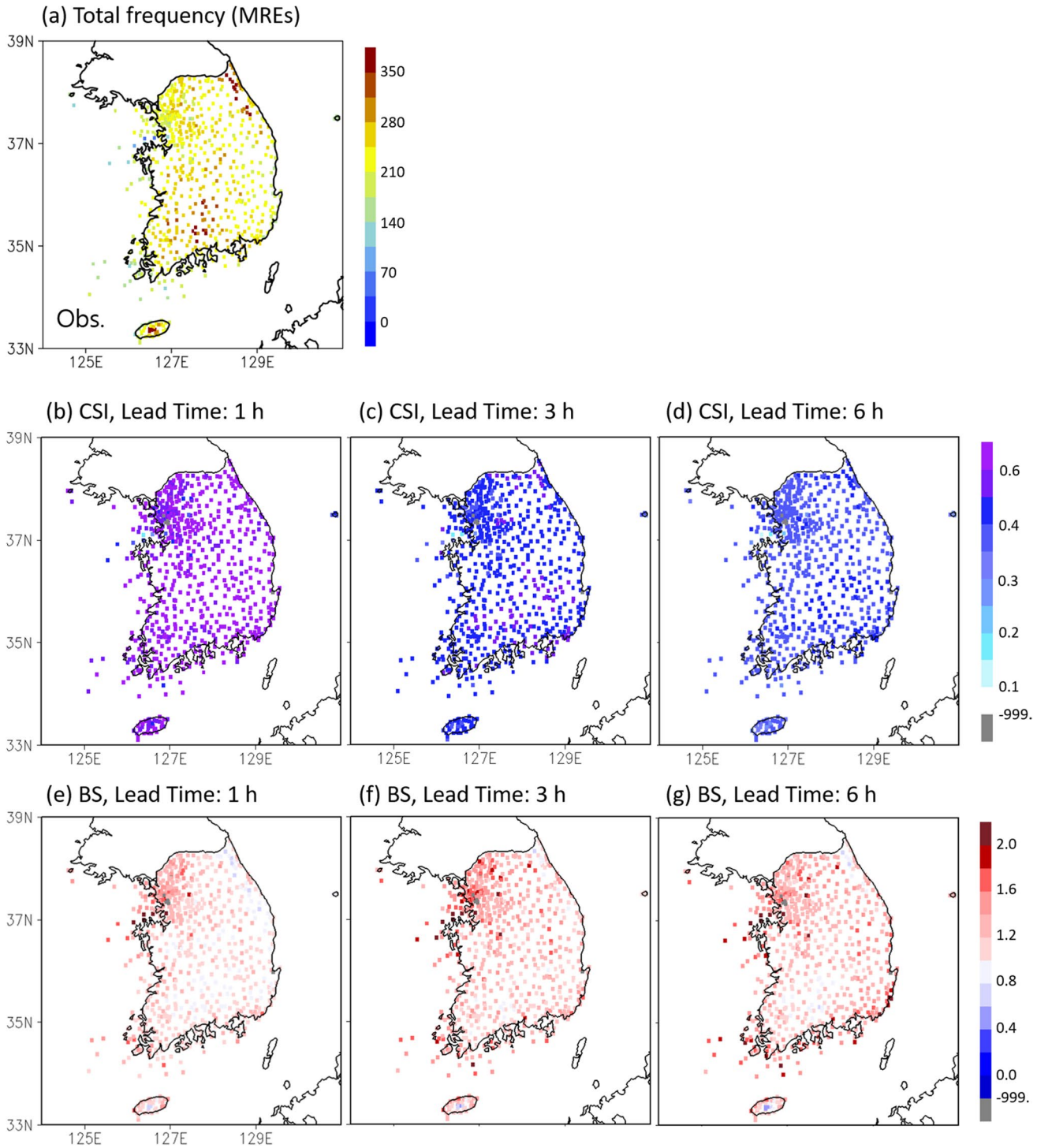

Fig. 3 Spatial distribution of (**a**) the frequency (unit: number) of moderate rainfall events (MREs; ≥ 1 mm h^{-1}) observed during June–September 2020 in South Korea. (**b–d**) CSI and (**e–g**) BS of DEEP-RANE rainfall forecasts at lead times of 1, 3, and 6 h. Note also that there are several observation points with CSI score of "-999" value in (**d**). These points have no hit (H = 0) which denotes no correctly forecasted rainfall events (Table 1). To avoid dividing by zero when calculating CSI score, we replaced it with a value of "-999"

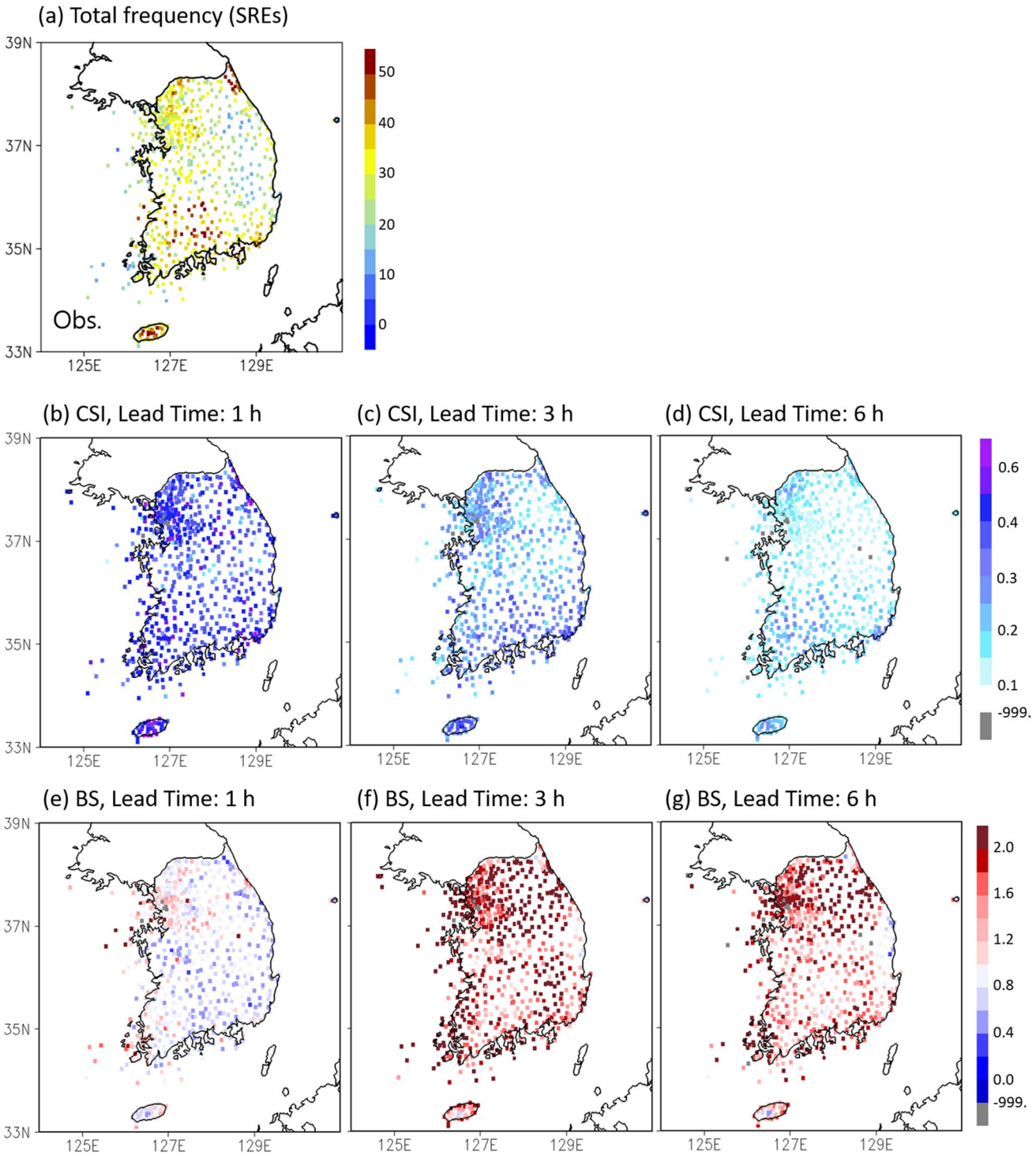

Fig. 4 Same as Fig. 3 but for strong rainfall events (SREs; ≥ 10 mm h^{-1}). Note that the color bar range in (**a**) is smaller than that in Fig. 3a

approaching from eastern China. The subsequent rainfall events in late July–mid-August (green shading) were mainly caused by quasi-stationary monsoon rainbands (MRBs), whereas those in late August–early September (yellow shading) were triggered by TCs.

For both MRE and SRE forecastings, DEEPRANE shows better performance during the MRB period than during the ETC and TC periods. The MRE forecasting during the MRB period (August 1 to 15) exhibits relatively high CSI scores of 0.5–0.6 even at a lead time of 6 h (Fig. 5b) with BSs of

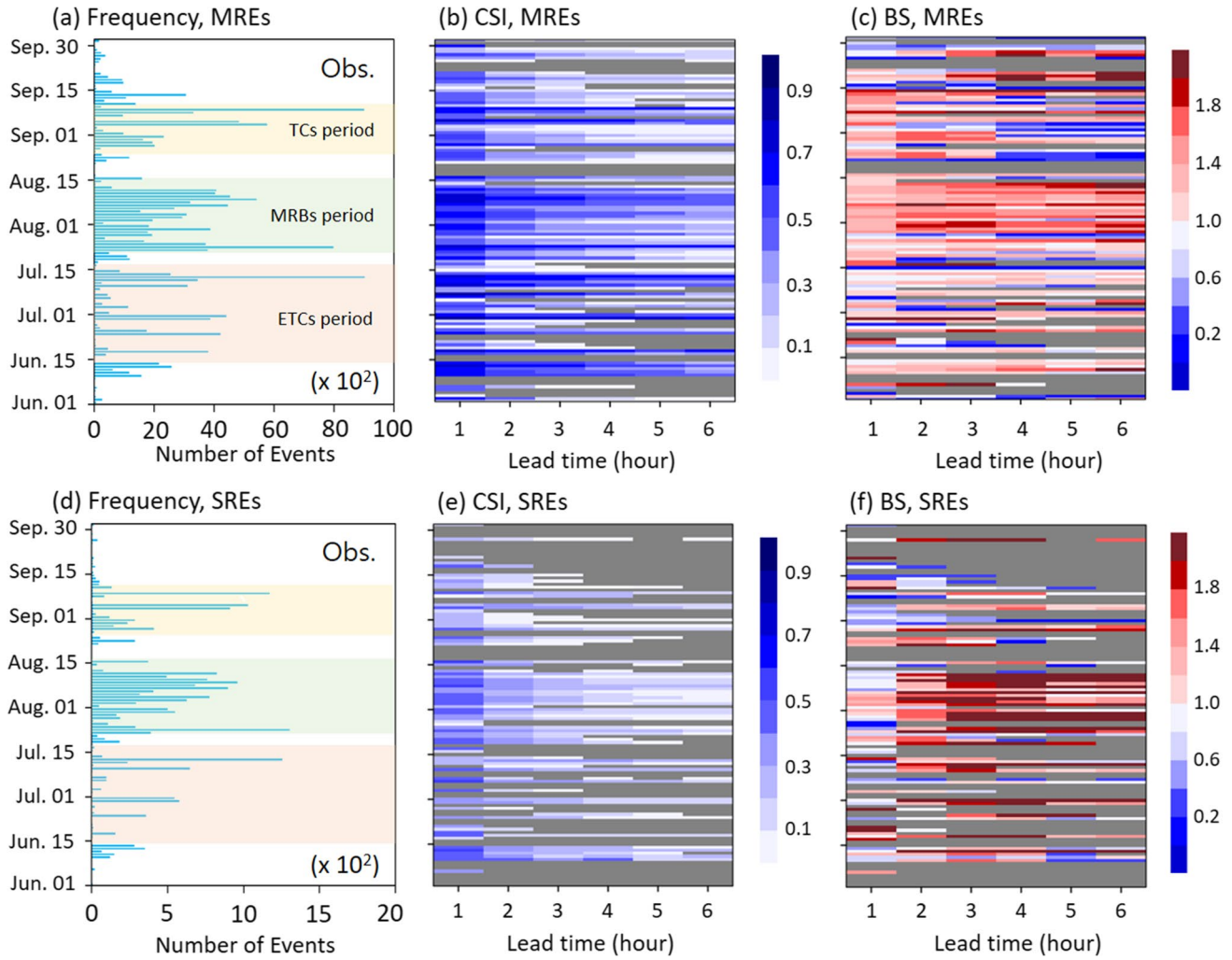

Fig. 5 Sub-seasonal variation of (**a**) MREs and (**d**) SREs, respectively, and DEEPRANE forecast skills in terms of (**b**, **e**) CSI and (**c**, **f**) BIAS scores according to the lead time. The periods separated by distinct synoptic weather systems, i.e., monsoon rainband (MRB), extratropical cyclone (ETC), and tropical cyclone (TC), are roughly indicated by shading in (**a**) and (**d**). Note that the ranges of the x-axis for the event frequency in (**a**) and (**d**) are different. The days with no observed and forecasted events, i.e., no rain day across South Korea, are shaded in gray in (**b**, **c**, **e**, **f**)

1.4–1.8 (Fig. 5c). For SRE forecasting, the CSI scores of 0.4–0.5 are maintained for 2 h (Fig. 5e). The skill scores then rapidly decrease with BSs greater than 2 at lead times of 3 h and longer (Fig. 5f). By contrast, during the ETC and TC periods, DEEPRANE tends to under-predict both MREs and SREs. The under-prediction is evident even at short lead times, i.e., 1–2 h, in SRE forecasting especially during the TC period (Fig. 5f). This result implies that relatively low performance of DEEPRANE in SRE forecasting compared to MRE forecasting results in part from migrating weather systems, i.e., ETCs and TCs, instead of slowly moving systems, i.e., MRBs.

Figure 6 shows the diurnal variation of MREs and SREs aggregated across all AWS stations during summer 2020. The corresponding CSI and BS of DEEPRANE are also displayed as a function of lead time. The monsoon rainfall climatology is characterized by a distinct diurnal variation with an early morning peak (approximately 4–8 local time) and a late afternoon–early evening peak (approximately 15–20 local time; Jung and Suh 2005; Lee and Seo 2008; Oh and Suh 2018; Jo et al. 2020). The early morning peak is typically associated with nocturnal cloud-top radiative cooling, whereas the late afternoon–early evening peak is caused by daytime surface heating by solar insolation (Jung and Suh 2005; Zhou et al. 2008). In summer 2020, the morning peak was pronounced in both MREs and SREs, while the late afternoon–early evening peak was not evident.

The DEEPRANE shows better performance in the early morning hours than in other hours for both MRE

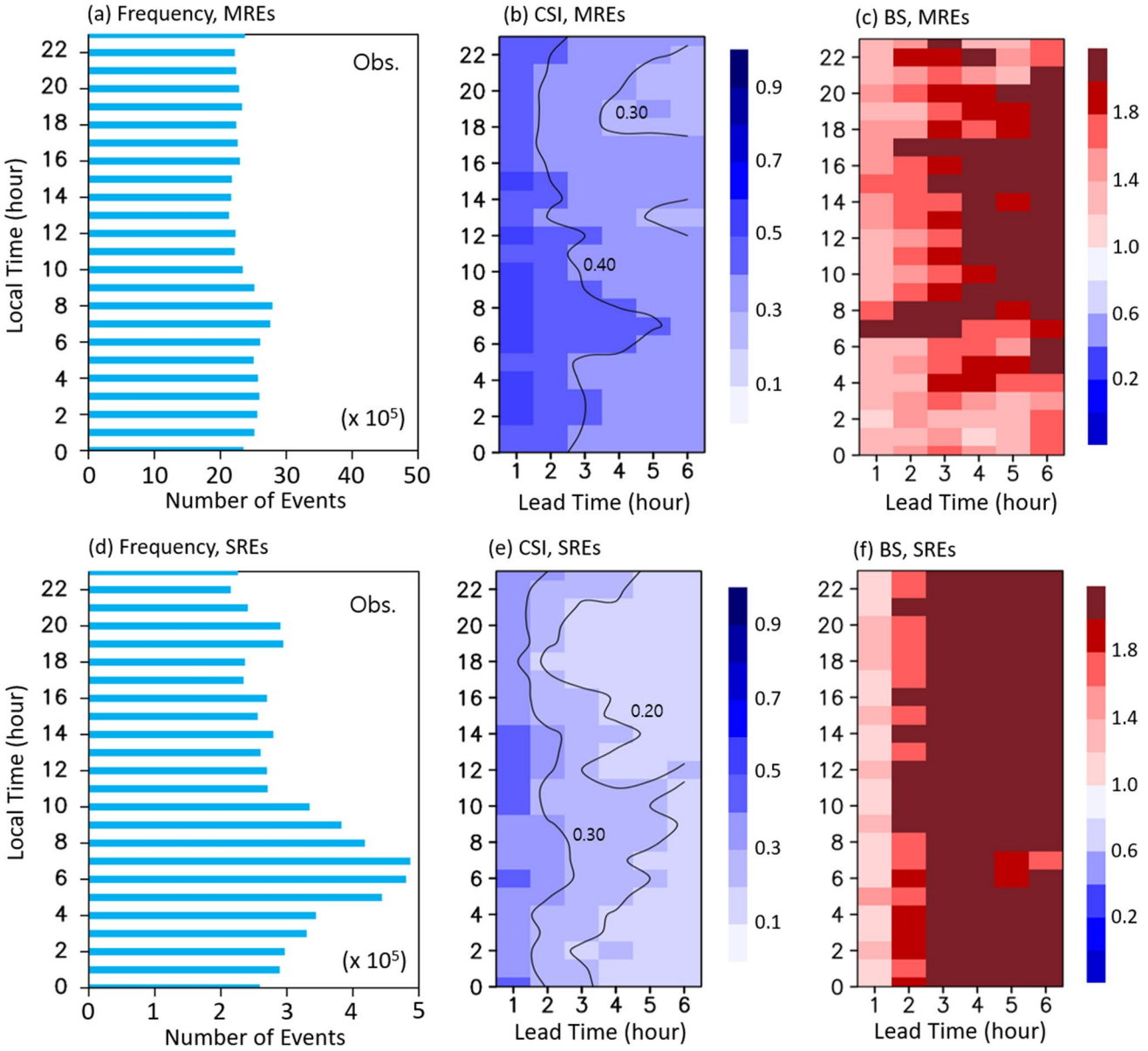

Fig. 6 Same as Fig. 5 but for the diurnal variation for all MREs and SREs in June–September 2020

and SRE forecastings. The CSI scores of MRE forecasting are greater than 0.40 in the early morning hours at lead times of 1–5 h (Fig. 6b), although the BSs are large (Fig. 6c). Likewise, the CSI scores of early-morning SRE forecasting are greater than 0.30 at lead times of up to 2 h and maintained above 0.20 until a 5-h lead time (Fig. 6e). The better performance in the early morning than in the afternoon may result from more samples to train the model as the early-morning rainfall events are most frequent (Oh and Suh 2018; Jo et al. 2020). This result indicates that DEEPRANE may provide more reliable forecast information for rainfall events triggered by nocturnal radiative cooling.

Figure 7a presents the overall CSI score versus lead time. Here, the total number of samples used at each lead time is 2,064,240 (122 days × 24 h × 705 AWS stations). As mentioned above, the CSI scores range from 0.60 to 0.38 for MRE forecasting and from 0.40 to 0.10 for SRE forecasting. They decrease with increasing lead time. A qualitative comparison with the deep-learning-based rainfall nowcasting in previous studies (e.g., Ravuri et al. 2021) reveals that DEEPRANE works well even at long lead times.

The POD and FAR are further illustrated in Fig. 7b. For a reliable forecast, the former should be higher than the latter. The POD score of MRE forecasting decreases from 0.81

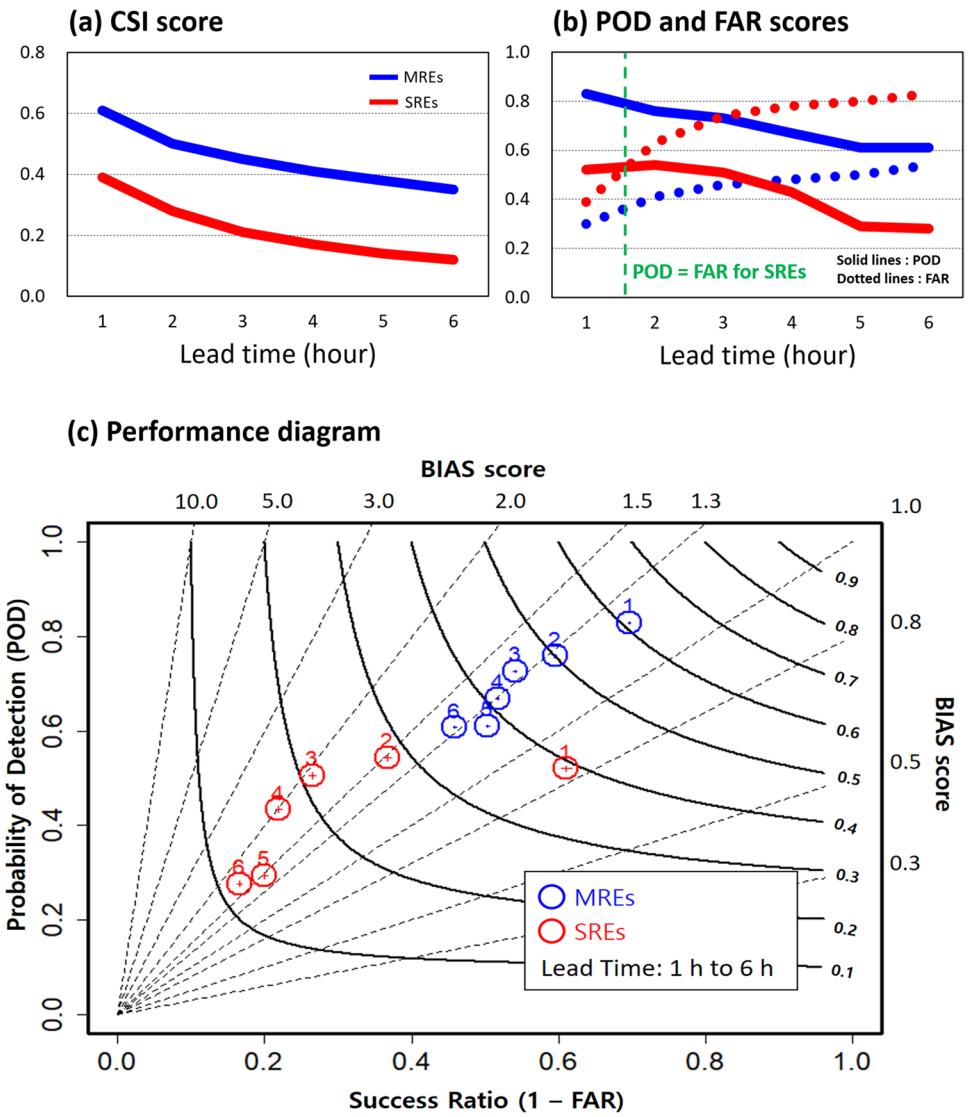

Fig. 7 An integrated validation of the DEEPRANE forecasts during June–September 2020 over South Korea. (**a**) CSI and (**b**) POD and FAR scores according to the lead times. (**c**) Performance diagram. Note that the dashed lines indicate BIAS scores and solid contours are CSI. Sampling uncertainty is represented by the crosshairs about the verification point and is estimated using bootstrapping method

to 0.60 as the lead time increases from 1 to 6 h. The FAR score shows the opposite pattern, increasing from 0.30 to 0.58. Here it is important to note that POD scores are always higher than FAR scores. As shown in the performance diagram (Fig. 7c), with increasing lead time, the MRE forecasting scores decrease but they are distributed relatively closer to the upper right (POD > FAR), with BS ranging from 1.2 to 1.5. In addition, the sampling uncertainty is also reasonably small (see crosshairs within each symbol in Fig. 7c). This result indicates that DEEPRANE-based very short-term MRE forecasting could be useful even at lead times of up to 6 h. However, this is not the case for SRE forecasting. The POD score decreases from 0.50 to 0.30, while its FAR score increases from 0.40 to 0.81 with lead time. The POD becomes smaller than the FAR at lead times of 2 h or longer (Fig. 7b), resulting in most SRE forecasting scores being distributed relatively closer to the lower left in the performance diagram except for it at 1 h lead time (Fig. 7c). In addition, BS is also showing relatively greater sensitivity depending on the lead time than that of MRE forecasting. This result indicates that DEEPRANE-based SRE forecasting at these lead times should be used with caution.

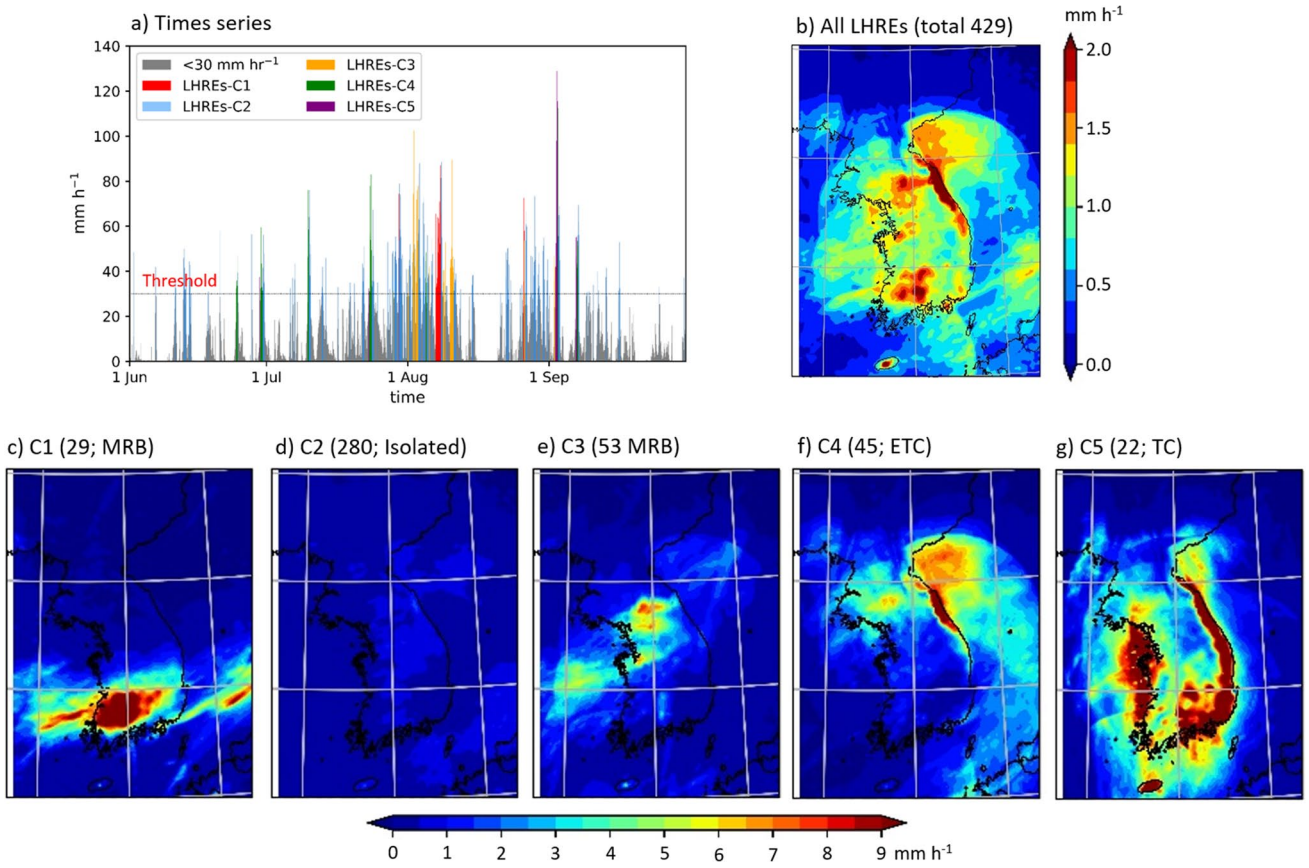

Fig. 8 (**a**) Time series of the observed 1-h accumulated rainfall in June–September 2020 over South Korea. The maximum rainfall amounts among 705 stations operated in 2020 are shown. The localized heavy rainfall events (LHREs; ≥ 30 mm h^{-1}) are colored by referring their cluster number. (**b**–**f**) Spatial distribution of rainfall (mm h^{-1}) averaged for all LHREs and each cluster obtained from the self-organizing map (SOM) analysis. The MRB, ETC, and TC are monsoon rainband, extratropical cyclone, and tropical cyclones, respectively

4 Localized Heavy Rainfall Event (LHRE) Forecasting

This section focuses on the evaluation of DEEPRANE for LHREs. The LHREs were observed roughly 429 times in South Korea in summer 2020. Figure 8a presents the time series of the observed 1-h maximum rainfall amounts among 705 AWSs in June–September 2020 over South Korea. The KMAPR-based spatial distribution of rainfall amount averaged across all LHREs is also illustrated in Fig. 8b. They are concentrated mainly in the southern inland region, the western coastal and middle inland regions, and the northeast coastal region.

The LHREs in 2020 are classified into five types using the SOM method (see Section 2d for details). Their occurrence days are indicated in Fig. 8a with the spatial distribution of each LHRE type in Figs. 8b–g. Most frequent LHREs are cluster 2 (C2 hereafter; Fig. 8d). They are the isolated LHREs (Jo et al. 2020) caused by very localized rainstorms with no synoptic-scale systems. They occurred sporadically throughout summer 2020 (Fig. 8a) and accounted for approximately 65% of the total LHREs (Fig. 8d). This result is consistent with Jo et al. (2020). The C1 and C3 are LHREs organized by MRBs in the southern and central regions of the Korean Peninsula, respectively (Figs. 8c and e). They occurred in early to mid-August (Fig. 8a; see also Fig. 5a), accounting for approximately 6.8% (C1) and 12.3% (C3) of total LHREs. The C4 and C5 consist of LHREs mainly resulting from ETCs (or extratropical transitions of TCs) and TCs, respectively (Figs. 8f and g). Although most of C4 LHREs occurred from mid-June to late July, C5 LHREs occurred in early September. They accounted for approximately 10.5% (C4) and 5.1% (C5) of total LHREs.

The performance of DEEPRANE in predicting the spatial distribution of LHREs is first examined for the representative cases in Fig. 9. The details for each case selected, such as case number, date, maximum rainfall, and synoptic system, are summarized in Table S2. As shown in the first

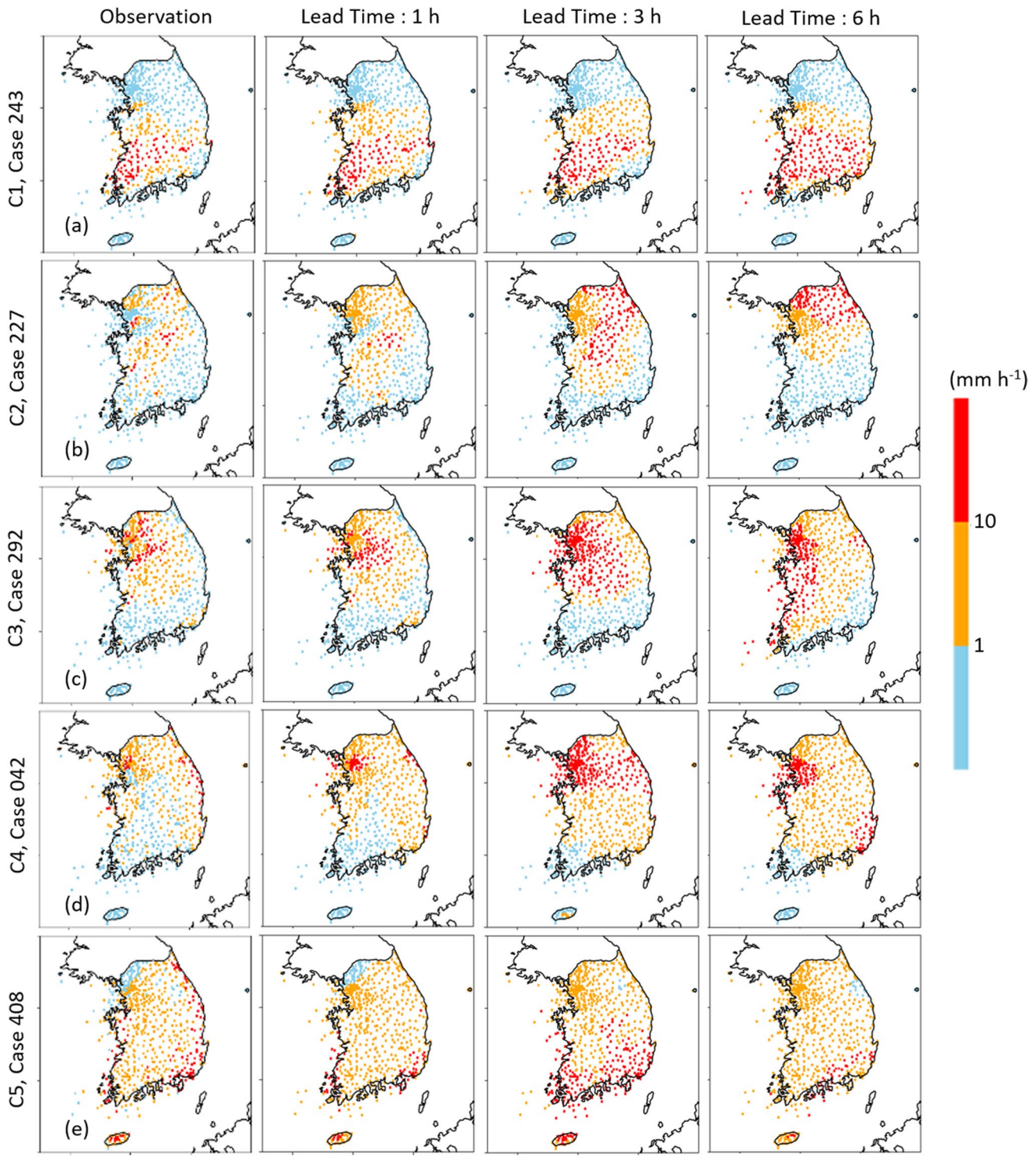

Fig. 9 **a**) Spatial distribution of the (first column) observed and (second to fourth columns) predicted rainfalls at lead times of 1, 3, and 6 h, for the selected LHRE. The details of the selected case, such as case number, date, maximum rainfall, synoptic condition, etc., are summarized in Table S3

column of Fig. 9, the spatial distribution of hourly rainfall well matches with each LHRE cluster shown in Figs. 8c–g. The spatial distribution of DEEPRANE forecasts at a 1-h lead time is quantitatively similar to the observations regardless of LHRE type (see the second column of Fig. 9). This result indicates that DEEPRANE can accurately predict LHREs in advance of 1 h. However, DEEPRANE tends to over-predict LHREs with increasing lead time, as in MRE and SRE forecastings. The location of LHREs is also inaccurately predicted at long lead times. For instance, C4 LHRE (Fig. 9d) is predicted not only in Seoul/Gyeonggi region but also in the southeastern coastal region at a lead time of 6 h. The significant over-prediction in the northeast regions is also evident in C2 LHRE at a lead time of 6 h (Fig. 9b). This result is consistent with a case study by Kim and Hong (2022). They showed that the deep learning model has a relatively poor skill in very localized rainfall events compared to MRB- and TC-related rainfall events. The poor forecasting skill of C2-related rainfall events can be attributable to the short lifetime and small spatial scale (e.g., Lee et al. 1998; Lee and Kim 2007). It is noteworthy that DEEPRANE reasonably well predicts C1 LHRE in the southwestern coastal and inland regions even at a lead time of 6 h. This result is consistent with good forecasting skill during the MRB period, as shown in Fig. 5.

All LHREs are comprehensively evaluated in a performance diagram (Fig. 10) in the same manner as in Fig. 7. With increasing lead time, CSI score sharply decreases with BS ranging from 0.5 to 3.0 (Figs. 10a and c). For all LHRE clusters, POD is greater than 0.5 at lead times of up to 3 h, and is higher than FAR at lead times of up to 2 h (Figs. 10b and c). This result demonstrates that DEEPRANE LHRE forecasting can be meaningful up to lead times of at least 2 h. More importantly, the performance of DEEPRANE is

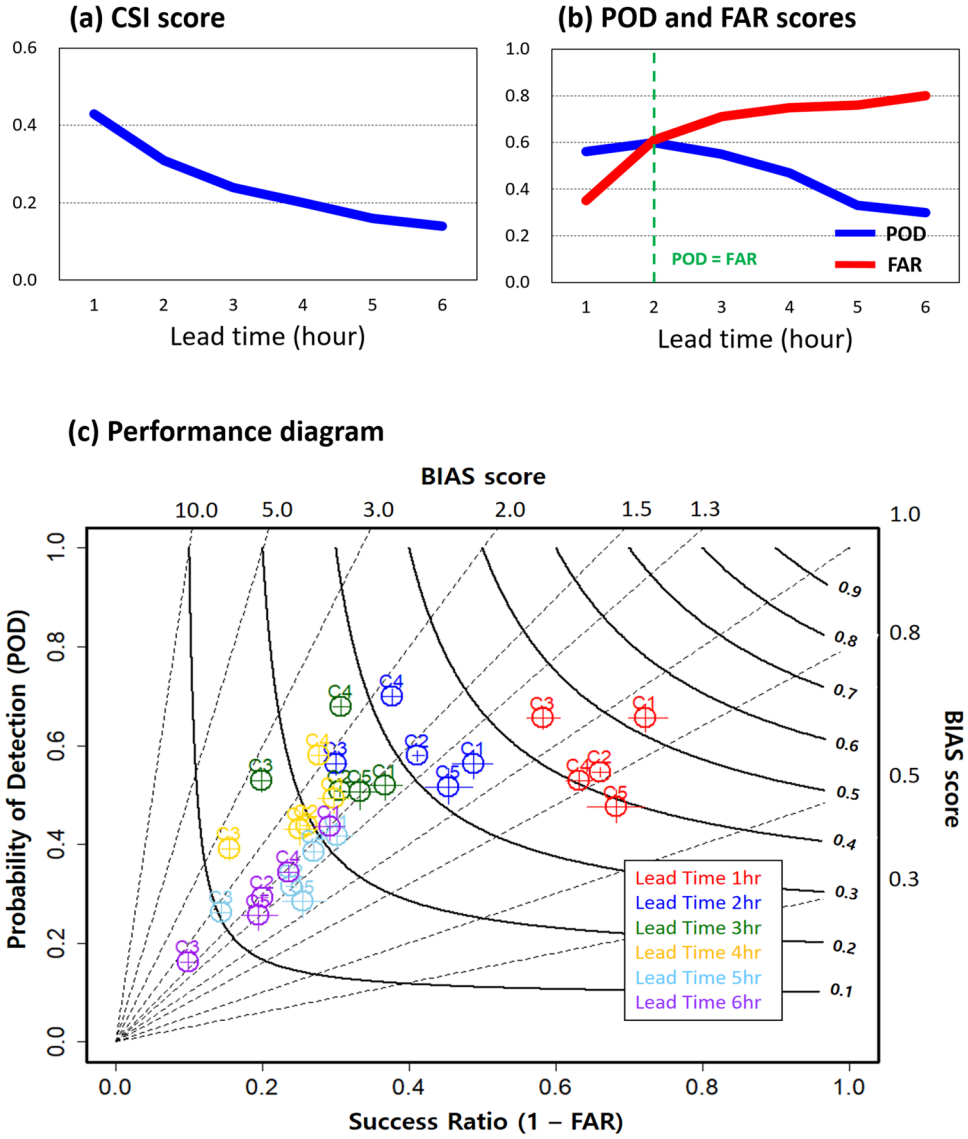

Fig. 10 Same as Fig. 7 except for DEEPRANE LHRE forecast. In (**a**) and (**b**), the scores are calculated using all LHREs. In (**c**), each LHRE cluster is presented above the circle symbol

comparatively for each LHRE cluster at lead times shorter than 3 h (Fig. 10c). This result suggests that DEEPRANE has meaningful performance for predicting LHREs at shorter lead times, irrespective of LHRE type.

5 Summary and Discussion

This study evaluates the performance of the deep learning model DEEPRANE (Ko et al. 2022) for very short-term rainfall forecasts during the summer monsoon season in South Korea. The June–September 2020, when record-breaking summer rainfall occurred, is particularly tested. The DEEPRANE can adequately predict MREs (rainfall ≥ 1 mm h^{-1}), with CSI scores of 0.6 to 0.4 at lead times of 1 to 6 h, respectively. However, the model tends to over-predict intensity of MREs by 1.2 to 1.6 times, increasing with lead time. For SREs (rainfall ≥ 10 mm h^{-1}), CSI scores of 0.3–0.5 are found at a lead time of 1 h. However, they rapidly decrease with lead time. The rainfall amounts are also significantly over-predicted at lead times of 2 h or longer. When synoptic conditions are considered, higher performance is found when MREs and SREs are associated with MRBs. The rainfall events associated with ETCs or TCs are rather poorly predicted. The DEEPRANE also shows better performance in the early morning hours, when rainfall events are more frequent than at other times of day.

The MRE forecasting is meaningful for 6 h as POD is higher than FAR. However, SRE forecasting is meaningful only at a lead time of 1 h. This result indicates that although DEEPRANE is useful for SRE nowcasting, it should be used with caution at lead times of 2 h and longer. When considering LHREs (rainfall ≥ 30 mm h^{-1}), DEEPRANE provides meaningful forecast at lead times of up to 2 h irrespective of LHRE types.

Here it should be stated that this study is not the first to attempt deep-learning-based rainfall forecasts in South Korea. Yoon et al. (2020), for instance, have tested very short-term rainfall forecasts in South Korea using a deep learning model. The present study extended the spatial coverage and forecasting lead times of Yoon et al. (2020) by using a new model DEEPRANE. The predictability of LHRE and its dependency to LHRE type are also evaluated. Although a direct comparison is not possible, DEEPRANE in this study, which has a CSI score of 0.60 at the lead time of 1 h for the MRE forecast, shows better performance than Yoon et al. (2020) with a CSI score of 0.33. It also shows a reliable performance of LHREs above a CSI of 0.14 at longer lead times of 4–6 h than the model developed by Kim and Hong (2022) with a CSI of 0.09 at a lead time of 2.5 h, allowing more time to protect against LHRE-related natural hazards. These results indicate that DEEPRANE has potential for use as an alternative to the current NWP-based operational short-term forecast.

Several issues need to be addressed to improve DEEPRANE. The DEEPRANE has a relatively poorer performance for SRE (and LHRE) forecasting than MRE forecasting, in particular when transient weather systems such as ETCs and TCs are active. The model also tends to over-predict rainfall amounts at long lead times. These are common in deep-learning-based models (e.g., Agrawal et al. 2019; Ayzel et al. 2020; Han et al. 2020).

A relatively poor prediction skill of SREs at long lead times can be largely attributed to the insufficient sample size used based on only radar data for model training. The diversity, as well as the quantity of training data, is also critical to improve deep-learning performance across different conditions, such as extreme events (O et al. 2020; Meyer and Pebesma 2021). To resolve this issue, additional input data (i.e., satellite image, NWP simulation, etc.) could be used. Moraux et al. (2021) for instance utilized thermal infrared satellite imagery to train a deep learning model. They showed that the addition of satellite images provides more accurate rainfall detection and estimation than those obtained when only radar images are used. The NWP model output could be also useful to increase the sample size. The long-term NWP model simulation, which well captures ETCs and TCs, would be particularly useful for training the model (e.g., Han et al. 2020; Kashinath et al. 2021). In addition, the consideration of various variables (i.e., temperature, moisture, geopotential height, etc.) could potentially improve the rainfall forecasts (Liyew and Melese 2021; Endalie et al. 2022).

Another possible approach to improve DEEPRANE is to detect synoptic weather pattern in advance and apply it to rainfall forecasts. Kumler-Bonfanti et al. (2020) attempted to detect ETCs and TCs using the U-Net models trained from NWP and satellite water vapor data. They showed that the deep learning models can extract cyclone location information faster than conventional methods with relatively high accuracy. With improved ETC and TC forecasts, a short-term rainfall forecasting could be improved.

Finally, it may be worthwhile to attempt combining information (Xiao et al. 2018; Mohammed and Kora 2021) predicted by various deep-learning models (e.g., CNN, SegNet, LSTM, U-Net, etc.), like multi-model ensembles often used in climate models, making probabilistic rainfall forecasting, which is more beneficial for longer lead times.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13143-022-00310-4.

Acknowledgements This study is supported by the Korea Meteorological Administration Research and Development Program under Grant KMI2020-01010-1. This work also was funded by the Korea Meteorological Administration Research and Development Program "Enhancement of Convergence Technology of Analysis and Forecast on Severe Weather" under Grant (KMA2018-00121).

Data Availability The software of deep-learning model, DEEPRANE version 1.0, used in this study is released under the GPL-3.0 license at https://github.com/jihoonko/DeepRaNE. To run the model, Python 3.6.9 or later, NumPy 1.19 or later, and Pytorch 1.6.0 or later are required to be installed in the system. Contact S.-G. Oh (seokgeunoh@snu.ac.kr) if you have interested in the analysis codes based on Fortran, Grads, R, etc. utilized for the results of this study. The radar reflectivity and AWS observation across South Korea used in this study can be available on Korea Meteorological Administration data released website "https://data.kma.go.kr/cmmn/main.do". Contact "https://www.kma.go.kr/nmc/html/main/index.jsp" if you have interested in the KMAPR data.

Declarations

Competing Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agrawal, S., Barrington, L., Bromberg, C., Burge, J., Gazen, C., Hickey, J.: Machine learning for precipitation nowcasting from radar images. Arxiv preprint, (2019). https://arxiv.org/abs/1912.12132. Accessed 1 June 2021

Araki, K., Kato, T., Hirockawa, Y., Hashiko, W.: Characteristics of atmospheric environments of quasi-stationary convective bands in Kyushu, Japan during the July 2020 heavy rainfall event. SOLA (2021). https://doi.org/10.2151/sola.2021-002

Ayzel, G., Scheffer, T., Heistermann, M.: RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting. Geosci. Model Dev. **13**, 2631–2644 (2020)

Cox, D.R.: The regression analysis of binary sequences. J. R. Stat. Soc. **20**(2), 215–242 (1958)

Endalie, D., Haile, G., Taye, W.: Deep learning for daily rainfall prediction: case study of Jimma. Ethiopia. Water Supply **22**(3), 3448–3461 (2022)

Han, Y., Zhang, G.J., Huang, X., Wang, Y.: A moist physics parameterization based on deep learning. J. Adv. Model. Earth Syst. **12**(9), (2020). https://doi.org/10.1029/2020MS002076

Harper, K., Uccellini, L.W., Kalnay, E., Carey, K., Mornone, L.: 50th anniversary of operational numerical weather prediction. Bull. Am. Meteorol. Soc. **88**(5), 639–650 (2007)

Hirockawa, Y., Kato, T., Araki, K., Mashiko, W.: Characteristics of an extreme rainfall event in Kyushu district, southwestern Japan in early July 2020. SOLA **16**, 265–270 (2020). https://doi.org/10.2151/sola.2020-044

Jo, E., Park, C., Son, S.W., Roh, J.W., Lee, G.W., Lee, Y.H.: Classification of localized heavy rainfall events in South Korea. Asia-Pac. J. Atmos. Sci. **56**, 77–88 (2020)

Jung, J.H., Suh, M.S.: Characteristics and type of the diurnal variation of hourly precipitation during rainy season over South Korea. J. Korean Meteorol Soc. **41**(4), 533–546 (2005)

Kashinath, K., Mustafa, M., Albert, A., Wu, J.L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H.A., Marcus, P., Anandkumar, A., Hassanzadeh and Prabhat, P.: Physics-informed machine learning: case studies for weather and climate modelling. Phil. Trans. R. Soc. A **379**, (2021). https://doi.org/10.1098/rsta.2020.0093

Kim, Y., Hong, S.: Very Short-term prediction of weather radar-based rainfall distribution and intensity over the Korean Peninsula using convolutional long short-term memory network. Asia-Pac. J. Atmos. Sci. (2022). https://doi.org/10.1007/s13143-022-00269-2

Ko, J., Lee, K., Hwang, H., Oh, S.G., Son, S.W., Shin, K.: Effective training strategies for deep-learning-based preciptiation nowcasting and estimation. Comput. Geosci. (2022). https://doi.org/10.1016/j.cageo.2022.105072

Kohonen, T.: The self-organizing map. Neurocomputing **21**, 1–6 (1998)

Kohonen, T.: Essentials of the self-organizing map. Neural Netw. **37**, 52–65 (2013). https://doi.org/10.1016/j.neunet.2012.09.018

Kumler-Bonfanti, C., Stewart, J., Hall, D., Govett, M.: Tropical and Extratropical cyclone detection using deep darning. J. Appl. Meteorol. Climatol. **59**, 1971–1985 (2020)

Lee, T.Y., Kim, Y.H.: Heavy precipitation systems over the Korean peninsula and their classification. Asia-Pac. J. Atmos. Sci. **43**, 367–396 (2007)

Lee, G.H., Seo, K.H.: Analysis of diurnal and semidiurnal cycles of precipitation over South Korea. Atmos. **18**(4), 475–483 (2008)

Lee, D.K., Kim, H.R., Hogn, S.Y.: Heavy rainfall over Korea during 1980–1990. Korean J. Atmos. Sci. **1**, 32–50 (1998)

Lin, C., Vasic, S., Kilambi, A., Turner, B., Zawadzki, I.: Precipitation forecast skill of numerical weather prediction models and radar nowcasts. Geophys. Res. Lett. **32**, (2005), https://doi.org/10.1029/2005GL023451

Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P.: Focal loss for dense object detection. IEEE international conference on compute vision (ICCV), (2017), https://arxiv.org/abs/1708.02002

Liu, Y., Weisberg, R.H., Mooers, C.N.K.: Performance evaluation of the self-organizing map for feature extraction. J. Geophys. Res. **111**, C05018 (2016)

Liu, B., Yan, Y., Zhu, C., Ma, S., Li, J.: Record-breaking Meiyu rainfall around Yangtze River in 2020 regulated by the subseasonal phase transition of North Atlantic Oscillation. Geophys. Res. Lett. **47**, e2020GL090342 (2020). https://doi.org/10.1029/2020GL090342

Liyew, C.M., Melese, H.A.: Machine learning techniques to predict daily rainfall amount. J. Big Data **8**, 153 (2021). https://doi.org/10.1186/s40537-021-00545-4

Meyer, H., Pebesma, E.: Predicting into unknown space? Estimating the area of appicability of spatial prediction models. Methods Ecol. Evol. **12**(9), 1620–1633 (2021)

Mohammed, A., Kora, R.: An effective ensemble deep learning framework for text classification. J. King Saud Univ. – Comput. Inf. Sci. **10** (2021) https://doi.org/10.1016/j.jksuci.2021.11.001

Moraux, A., Dewitte, S., Cornelis, B., Munteanu, A.: A deep learning multimodeal method for preciptiation estimation. Remote Sens. **13**, 3278 (2021). https://doi.org/10.3390/rs13163278

O, S., Dutra, E., Orth, R.: Robustness of process-based versus data-driven modeling in changing climatic conditions. J. Hydrometeorol. **21**, 1929–1944 (2020)

Oh, S.G., Suh, M.S.: Change in seasonal and diurnal precipitation types during summer over South Korea in the late twenty-first century (2081–2100) projected by the RegCM4.0 based on four RCP scenarios. Clim. Dyn. **51**, 3041–3060 (2018)

Park, C., Son, S.W., Kim, H., Ham, Y.G., Kim, J., Cha, D.H., Chang, E.C., Lee, G.W., Kug, J.S., Lee, W.S., Lee, Y.Y., Lee, H.C., Lim, B.: Record-breaking summer rainfall in South Korea in 2020: Synoptic characteristics and the role of large-scale circulations. Mon. Wea. Rev. Published Onlne (2021a). https://doi.org/10.1175/MWR-D-21-0051.1

Park, C., Son, S.W., Kim, J., Chang, E.C., Kim, J.H., Jo, E., Cha, D.H., Jeong, S.: Diverse synoptic weather patterns of warm-season heavy rainfall events in South Korea. Mon. Wea. Rev. **149**, 3875–3893 (2021b)

Park, C., Son, S.W., Kim, J.H.: Role of baroclinic trough in triggering vertical motion during summertime heavy rainfall events in Korea. J. Atmos. Sci. **78**, 1687–1702 (2021c)

Pu, Z., Kalnay, E.: Numerical weather prediction basics: models, numerical methods, and data assimilation. In: Duan, Q., Pappenberger, F., Thielen, J., Wood, A., Cloke, H., Schaake, J. (eds.) Handbook of Hydrometeorological Ensemble Forecasting. Springer, Berlin, Heidelberg (2018)

Ravuri, S., Lenc, K., Willson, M., Kangin, K., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., Mohamed, S.: Skilful precipitation nowcasting using deep generative models of radar. Nature **597**, (2021). https://doi.org/10.1038/s41586-021-03854-z

Roebber, P.J.: Visualizing multiple measures of forecast quality. Wea. Forecast. **24**, 601–608 (2009)

Roh, J.W., Lee, Y.H., Nam, J.E., Chung, K.Y.: Diurnal variations of summertime precipitation in South Korea in 2009 using precipitation reanalysis data. SOLA **8**, 151–159 (2012)

Ronneberger, O., Fischer, P., Brox, T.:. U-Net: Convolutional networks for biomedical image segmentation. Interantional Conference on Medical image computing and computer-assisted intervention, (2015). https://arxiv.org/abs/1505.04597. Accessed 5 June 2021

Schultz, M.G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L.H., Mozaffari, A., Stadtler, S.: Can deep learning beat numerical weather prediction? Phil. Trans. R. Soc. A **379**, (2021). https://doi.org/10.1098/rsta.2020.0097

Shahrban, M., Walker, J.P., Wang, Q.J., Seed, A., Steinle, P.: An evaluation of numerical weather prediction based on rainfall forecasts. Hydrolog. Sci. J. **16**, 2704–2717 (2016)

Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K., Woo, W.C.: Convolutional LSTM network: a machine learning approach for precipitation. Comput. Sci. (2015). https://arxiv.org/abs/1506.04214. Accessed 27 May 2021

Shrestha, D.L., Robertson, D.E., Wang, Q.J., Pagano, T.C., Hapuarachchi, H.A.P.: Evaluation of numerical weather prediction model precipitation forecasts for shor-term streamflow forecasting purpose. Hydrol. Earth Syst. Sci. Discuss. **17**, 1913–1931 (2013)

Shuman, F.G.: History of numerical weather prediction at the National Meteological Center. Wea. Forecasting **4**, 286–296 (1989)

Sun, J., Xue, M., Wilson, J.W., Zawadzki, I., Ballard, S.P., Onwlee-Hooimeyer, J., Joe, P., Barker, D.M., Li, P.W., Golding, B., Xu, M., Pinto, J.: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. Bull. Am. Meteorol. Soc. **95**(3), 409–426 (2014)

Wang, G., Wang, E., Yang, J., Liu, L.: Evaluation and correction of quantiative preciptiation forecast by storm-scale NWP model in Jiangsu. China. Adv. Meteorol. (2016). https://doi.org/10.1155/2016/8476720

Xiao, Y., Wu, J., Lin, Z., Zhao, X.: A deep learning-based multi-model ensemble method for cancer prediction. Comput. Methods Programs Biomed. **153**, 1–9 (2018)

Yen, M.H., Liu, D.W., Hsin, Y.C., Lin, C.E., Chen, C.C.: Application of the deep learning for the precipitation of rainfall in Southern Taiwan. Sci. Rep, **9**, (2019). https://doi.org/10.1038/s41598-019-49242-6

Yoon, S., Park, H., Shin, H.: Very short-term rainfall prediction based on radar image learning using deep neural network. J. Korea Water Resour. Assoc. **53**(12), 1159–1172 (2020)

Yoon, S.: Adaptive blending method of radar-based and numerical weather prediction QPFs for urban flood forecasting. Remote Sens. **11**(6), (2019). https://doi.org/10.3390/rs11060642

Yu, W., Yoon, S., Choi, M., Jung, K.: Performance comparison of rainfall and flood forecasts using short-term numerical weather prediction data from Korea and Japan. J. Korea Water Resour. Assoc. **50**(8), 537–549 (2017)

Zhang, F., Wang, X., Guan, J., Wu, M., Guo, L.: RN-Net: a deep learning approach to 0–2 h rainfall nowcasting based on radar and automatic weather station data. Sens. **21**, (2021). https://doi.org/10.3390/s21061981

Zhou, T., Yu, R., Chen, H., Dai, A., Pan, Y.: Summer precipitation frequency, intensity, and diurnal cycle over China: a comparison of satellite data with rain gauge observation. J. Clim. **21**(16), 3997–4010 (2008)