

1 **Evaluation of Mei-yu Heavy-Rainfall Quantitative Precipitation**
2 **Forecasts in Taiwan by A Cloud-Resolving Model for Three**
3 **Seasons of 2012-2014**

4 Chung-Chieh Wang¹, Pi-Yu Chuang^{1*}, Chih-Sheng Chang¹, Kazuhisa Tsuboki²,
5 Shin-Yi Huang¹, and Guo-Chen Leu³

6 1 Department of Earth Sciences, National Taiwan Normal University, Taipei, Taiwan

7 2 Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan

8 3 Central Weather Bureau, Taipei, Taiwan

9 *Corresponding author:* Pi-Yu Chuang (giselle780507@hotmail.com)

10

11 **Abstract.**

12 In this study, the performance of quantitative precipitation forecasts (QPFs) by the Cloud-Resolving Storm Simulator
13 (CReSS) in Taiwan, at a horizontal grid spacing of 2.5 km and a domain size of $1500 \times 1200 \text{ km}^2$, at the ranges of 1-
14 3 days during three mei-yu seasons (May-June) of 2012-2014 is evaluated using categorical statistics, with an
15 emphasis on heavy rainfall events ($\geq 100 \text{ mm per 24 h}$). The categorical statistics are chosen because the main hazards
16 are landslides and floods in Taiwan, so predicting heavy rainfall at the correct location is important. The overall threat
17 scores (TSs) of QPFs for all events on day 1 (0-24 h) are 0.18, 0.15, and 0.09 at the threshold of 100, 250, and 500
18 mm, respectively, and indicate considerable improvements at increased resolution compared to past results and 5-km
19 models (TS < 0.1 at 100 mm and **TS ≤ 0.02** at 250 mm).

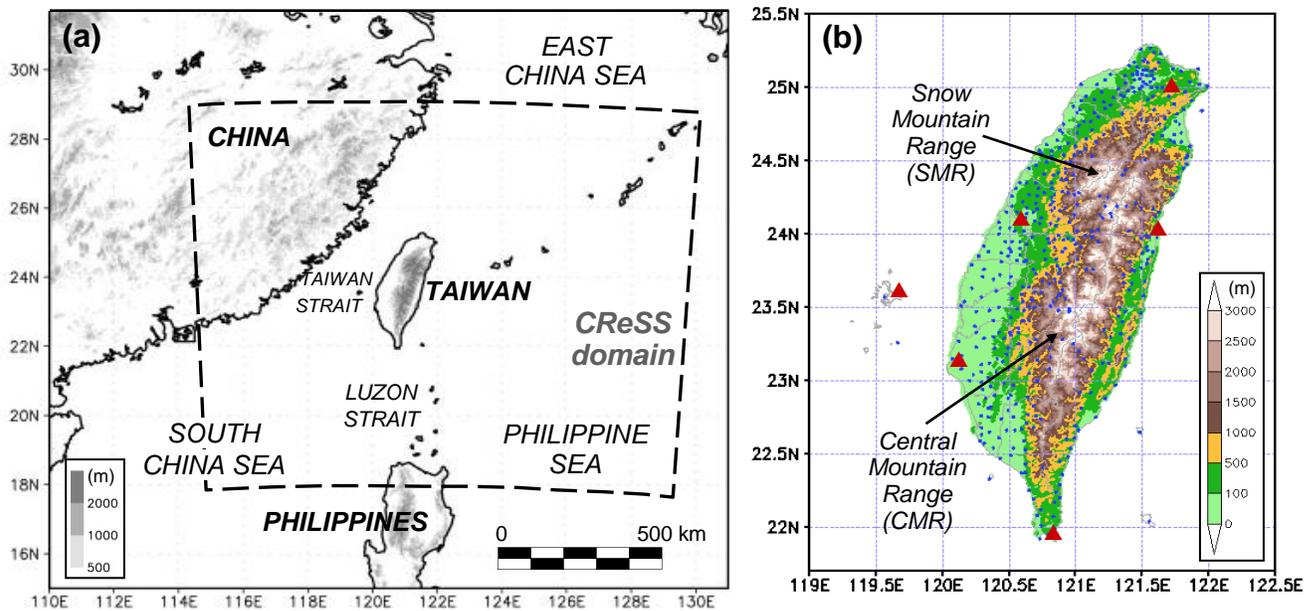
20 Moreover, the TSs are shown to be higher and the model more skillful in predicting larger events, in agreement with
21 earlier findings for typhoons. After classification based on observed rainfall, the TSs of day-1 QPFs for the largest 4%
22 of events by CReSS at 100, 250, and 500 mm (per 24 h) are 0.34, 0.24, and 0.16, respectively, and can reach 0.15 at
23 250 mm on day 2 (24-48 h) and 130 mm on day 3 (48-72 h). The larger events also exhibit higher probability of
24 detection and lower false alarm ratio than smaller ones almost without exception across all thresholds. With the
25 convection and terrain better resolved, the strength of the model is found to lie mainly in the topographic rainfall in
26 Taiwan rather than migratory events that are more difficult to predict. Our results highlight the crucial importance of
27 cloud-resolving capability and the size of fine mesh for heavy-rainfall QPFs in Taiwan.

28

29

30 **1 Introduction**

31 The Quantitative Precipitation Forecasting (QPF) is one of the most challenging areas in modern numerical weather
32 prediction (e.g., Golding, 2000; Fritsch and Carbone, 2004; Cuo et al., 2011), especially for extreme events that have
33 high potential for hazards. With its steep and complex topography, Taiwan over the western North Pacific (Fig. 1)
34 experiences extreme rainfall rather frequently, mainly during two periods: the typhoon (July-October) and mei-yu
35 (May-June) seasons (e.g., Kuo and Chen, 1990; Wu and Kuo, 1999; Jou et al., 2011; Chang et al., 2013). The landslides
36 and flash floods in/near the mountains and flooding over low-lying plains and urban areas are the main hazards (e.g.,
37 Wang et al., 2012b, 2013b, 2016b). In order to better prepare for these hazards and reduce their impacts, QPFs and
38 their verifications, especially over heavy-rainfall thresholds from large events, are thus very important for Taiwan. Of
39 course, to identify where the model can make significant improvements in QPFs and what approaches are effective to
40 achieve them are also crucial (e.g., Clark et al., 2011).



41
42 **Figure 1: (a) The geography and topography (m, shading) surrounding Taiwan and the domain of 2.5-km CReSS (thick**
43 **dashed box), and (b) the detailed terrain of Taiwan (m, color) and the locations of rain gauges (blue dots) and land-based**
44 **radars (scarlet triangles) used to produce the reflectivity composites by the Central Weather Bureau (CWB). The two major**
45 **mountain ranges in Taiwan, the Central Mountain Range (CMR) and Snow Mountain Range (SMR), are marked in (b).**

46 For the mei-yu season in Taiwan, earlier studies mainly employed the widely-used, standard categorical measures (see
47 Section 2.4) to evaluate the performance of models such as the Mesoscale Model version 5 (MM5) at thresholds up to
48 50 mm per 12 h (e.g., Chien et al., 2002, 2006; Yang et al., 2004). Their results show that the models at the time had
49 some ability in predicting rainfall occurrence at thresholds ≤ 2.5 mm, but little skill at 50 mm and above. In recent
50 years, several studies (e.g., Hsu et al., 2014; Li and Hong, 2014; Su et al., 2016; Huang et al., 2016) have also examined
51 the QPFs by the Weather Research and Forecasting (WRF) model (Skamarock et al., 2005) running at the Central
52 Weather Bureau (CWB) at 5-km grid spacing (Δx), including its ensembles. These studies indicate improvements over
53 earlier models at thresholds up to 50-100 mm (per 12 h) over the previous decade. However, the accuracy at 150-200

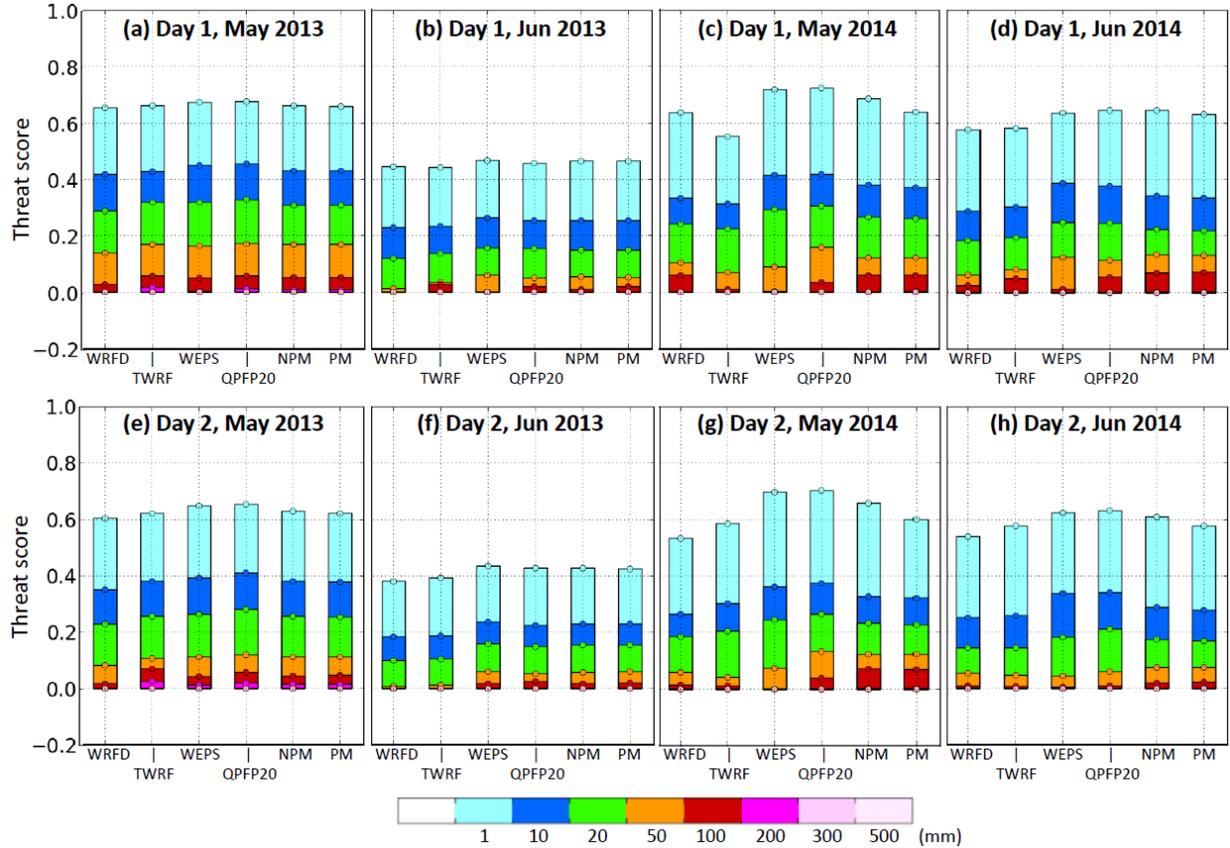
54 mm and beyond is still limited, even with probability-matching (e.g., Ebert, 2001) within the forecast range of 24 h
55 (see e.g., Figs. 9 and 10 of Huang et al., 2016).

56 Figure 2 shows the threat scores (TSs) of day-1 (0-24 h) and day-2 (24-48 h) QPFs for May and June of 2013 and
57 2014 by the CWB models as an example. The TS is defined as the fraction of hits among all verification points that
58 are either observed or predicted, or both, at the specified rainfall threshold, and thus $0 \leq TS \leq 1$ (further details in
59 Section 2.4). The CWB models include deterministic WRF and several products from their 20-member WRF ensemble
60 prediction system (WEPS, e.g., Hong et al., 2015), all with $\Delta x = 5$ km. These plots are produced for each month at the
61 CWB for routine verification (within a range of 48 h) since 2013, and are similar to those in Huang et al. (2015, 2016).
62 In addition to deterministic forecasts, the scores also include those using probability-matching techniques (PM and
63 NPM, e.g., Ebert, 2001; Fang and Kuo, 2013), which may provide some benefit over the ensemble mean (WEPS) for
64 thresholds between 50-200 mm (e.g., Su et al., 2016; Huang et al., 2016). In Fig. 2, one can see that the TSs are no
65 higher than 0.07 at 100 mm (per 24 h) and 0.03 at 200 mm (and $TS = 0$ at and above 300 mm) for either day 1 or day
66 2 in the two mei-yu seasons, in line with the review above. The scores in June also tend to be lower compared to May,
67 likely due to more events of thermally-driven, localized rainfall with low predictability (e.g., Chen and Chen, 2003;
68 Chen et al., 1999; Paul et al., 2018). Nevertheless, effective strategies and methods to improve the skill level at
69 thresholds near 100-150 mm and beyond are needed.

70 Wang (2015, hereafter referred to as W15) evaluated the QPFs, within 3 days, by a cloud-resolving model (CRM),
71 the Cloud-Resolving Storm Simulator (CReSS; Tsuboki and Sakakibara, 2002, 2007), for all 15 typhoons that hit
72 Taiwan in 2010-2012. With $\Delta x = 2.5$ km, a grid size more comparable to research studies (e.g., Wang et al., 2005,
73 2011; 2013a, also Bryan et al., 2003; Done et al., 2004; Clark et al., 2007; Roberts and Lean, 2007), these deterministic
74 forecasts showed superior performance in QPFs, with TSs of 0.38, 0.32, and 0.16 at thresholds of 100, 250, and 500
75 mm, respectively, for all typhoons on day 1 (0-24 h, cf. his Fig. 13). Even on day 3 (48-72 h), the corresponding TSs
76 are 0.21, 0.12, and 0.01. Thus, the accuracy of QPFs by this CRM over the thresholds of 100-500 mm is remarkably
77 higher for typhoon rainfall in Taiwan.

78 Moreover, as summarized in Wang (2016), W15 found a strong positive dependency of categorical scores on overall
79 rainfall amount (which represents event magnitude). That is, the larger the rain, the higher the scores, and the better
80 the model performs. For example, the TSs at the same thresholds (100, 250, and 500 mm) for his top-5 events (roughly
81 top 5%) on day 1 are 0.68, 0.49, and 0.24, respectively (Fig. 1 of Wang, 2016), all at least 1.5 times higher than their
82 counterparts for all typhoons. An important implication of this finding is that the model QPFs for extreme events *may*
83 *not* be accurately assessed through categorical statistics without proper classification to isolate them from ordinary
84 events, and particularly not by taking arithmetic mean of TSs of multiple forecasts. The study of W15 also predicts
85 the dependency, as a fundamental property, to exist in other rainfall regimes. Therefore, the main purpose of this study
86 is three-fold: 1) to assess the ability of the 2.5-km CReSS in predicting mei-yu rainfall at a higher resolution than
87 before, especially for heavy to extreme rainfall events, 2) to clarify whether the dependency property in categorical
88 scores also exists in the mei-yu regime in Taiwan? and 3) if the QPFs by CReSS prove to be improved, why or where
89 its strength lies?

90



91
 92 **Figure 2: The TS of 0-24-h QPFs (day 1) for (a) May and (b) Jun of 2013, and (c) May and (d) Jun of 2014, respectively, at**
 93 **selected thresholds over 1-500 mm (per 24 h, scale at bottom) by two deterministic forecasts from WRF (WRFD) and the**
 94 **Typhoon WRF (TWRF) and four ensemble forecasts from the 20-member WRF Ensemble Prediction System: ensemble**
 95 **mean (WEPS), top 20% (QFPF20), and WEPS employing the probability matching (PM) and new PM (NPM) techniques.**
 96 **(e)-(h) As in (a)-(d), but showing the TS of 24-48-h QPFs (day 2), respectively.**

97 In Section 2, the model, data, and methodology are described. In Section 3, the overall scores of QPFs for groups with
 98 different event magnitudes are presented and compared with previous results. Then in Section 4, examples are selected
 99 to illustrate how the CRM performs in real-time forecasts and where its strength lies. Aspects related to the dependency
 100 property are further discussed in Section 5, and our conclusions are given in Section 6.

101 2 Data and Methodology

102 2.1 The CReSS model and its forecasts

103 The CReSS model is a non-hydrostatic, compressible CRM with a single domain without intermediate nesting
 104 (Tsuboki and Sakakibara, 2002, 2007), and it has been used for weather forecasts in Taiwan since 2006
 105 (<http://cressfcst.es.ntnu.edu.tw/>, W15; Wang et al., 2013, 2016a). Starting from July 2010, a grid size of 2.5 km is
 106 utilized, with a domain of $1500 \times 1200 \text{ km}^2$ since May 2012 (Fig. 1a and Table 1). In CReSS, cloud formation,
 107 development, and all related processes are explicitly treated using a bulk cold-rain microphysical scheme with six
 108 species (Lin et al., 1983; Cotton et al., 1986; Murakami, 1990; Ikawa and Saito, 1991; Murakami et al., 1994): vapor,
 109 cloud water, cloud ice, rain, snow, and graupel (Table 1). Thus, no cumulus (or shallow convection) parameterization

110 is used. Other sub-grid scale processes parameterized in the model include turbulent mixing in the planetary boundary
 111 layer with a 1.5-order closure (Deardorff, 1980; Tsuboki and Sakakibara, 2007), as well as surface radiation and
 112 momentum/energy fluxes (Kondo, 1976; Louis et al., 1982; Segami et al., 1989). These physical options are identical
 113 to W15, and also given in Table 1.

Season	2012	2013 and 2014
Projection	Lambert Conformal (center at 120°E, secant at 10°N and 40°N)	
Grid spacing (km)	2.5 × 2.5 × 0.2-0.663 (0.5)*	
Grid dimension (x, y, z)	600 × 480 × 40	
Domain size (km)	1500 × 1200 × 20	
Forecast frequency	Every 6 h (at 0000, 0600, 1200, and 1800 UTC)	
Forecast range	72 h	78 h
IC/BCs (including sea surface temperature)	NCEP GFS analyses/forecasts (at 26 levels)	
	1° × 1°	0.5° × 0.5°
Topography	Real at (1/120)° spatial resolution (~0.9 km)	
Cloud microphysics	Bulk cold-rain scheme (Lin et al., 1983; Cotton et al., 1986; Murakami, 1990; Ikawa and Saito, 1991; Murakami et al., 1994)	
PBL/turbulence	1.5-order closure with prediction of turbulent kinetic energy (Deardorff, 1980; Tsuboki and Sakakibara, 2007)	
Surface processes	Energy/momentum fluxes, shortwave and longwave radiation (Kondo, 1976; Louis et al., 1982; Segami et al., 1989)	
Substrate model	41 levels, every 5 cm to 2 m	

114 **Table 1: The basic configuration, initial/boundary conditions (IC/BCs), and physical packages of the 2.5-km CReSS used**
 115 **for real-time operation in 2012-2014. *The vertical grid spacing of CReSS is stretched (smallest at bottom), and the averaged**
 116 **spacing is given in the parentheses.**

117 The operational analyses and forecasts by the Global Forecasting System (GFS, Kanamitsu, 1989; Kalnay et al., 1990;
 118 Moorthi et al., 2001; Kleist et al., 2009) of the National Centers for Environmental Prediction (NCEP), produced every
 119 6 h (at 26 levels), were used as initial and boundary conditions (IC/BCs) for CReSS (Table 1). The CReSS model is
 120 also run four times a day, each out to 72 h (now 78 h). At the lower boundary, terrain data at 30'' resolution (roughly
 121 900 m) and the NCEP analyzed sea surface temperature are also provided. With its limited domain size, the
 122 atmospheric evolution in CReSS is forced by the NCEP forecasts, especially at longer ranges. Note that since 2013,
 123 the IC/BCs from the GFS have increased the resolution from 1° × 1° to 0.5° × 0.5°, but all other settings are kept the
 124 same during our study period (Table 1).

125 2.2 Data

126 The observational data used include synoptic weather maps from the CWB, the vertical maximum indicator reflectivity
 127 composites every 30 min from land-based radars, and hourly rainfall data from about 440 gauges in Taiwan for QPF
 128 verification (Fig. 1b). Along with NCEP gridded final analyses (on 1° × 1° grid), the weather maps are used to identify

129 and synthesize the occurrence of favorable factors to heavy rainfall among events with different magnitude (to be
 130 described in Section 2.3). For selected heavy-rainfall cases, the radar composites are compared with the CReSS
 131 forecasts to assess the quality of the QPFs in Section 4.

Group	Criterion (of 10% gauges)	No. of segments (%)	No. of all points (<i>N</i>)	No. of points (<i>H + M + FA</i>) at threshold (mm)			
				50	100	250	500
A+	≥ 130 mm (a subset of A)	13 (3.9)	5622	3807	2453	490	32
A	≥ 50 mm	61 (18.1)	26826	11000	4889	675	47
B	≥ 25 mm, but not A	75 (22.3)	33018	4279	1078	98	4
C	≥ 10 mm, but not B	88 (26.1)	38583	1675	281	10	3
D	≥ 1 mm, but not C	67 (19.9)	29267	266	32	0	0
X	< 1 mm	46 (13.6)	20067	59	20	4	0
All	A through D plus X	337 (100.0)	147761	17279	6300	787	54

132 **Table 2: The classification criteria using (at least) 10% of rain gauges with highest 24-h accumulated rainfall (0000-2400 or**
 133 **1200-1200 UTC) over Taiwan, and the results in the number of 24-h segments (and percentage) and total points (sites) of *H***
 134 **+ *M* + *FA* at selected rainfall thresholds (mm) for the different groups. During the mei-yu seasons of 2012-2014, the total *N***
 135 **is 148776 and on average there are 442 rain gauges per segment. The points of *H + M + FA* are based on the statistics of**
 136 **day-1 (0-24 h) QPFs, and *N* is also given (with no threshold).**

137 2.3 Verification period classification

138 In this study, objective categorical statistics (e.g., Schaefer, 1990; Wilks, 2011) are used to verify QPFs mainly because:
 139 1) the ability of models to predict the heavy rainfall at the correct location is imperative in Taiwan, since its primary
 140 hazards are landslides and floods, and 2) our results can be easily compared with earlier studies. Here, 24-h QPFs are
 141 chosen because: 1) the bulk rainfall accumulation from mei-yu events, as for typhoons, is our main concern rather
 142 than the rain over shorter periods, especially at longer ranges (days 2-3), and 2) the issue of double penalty on high-
 143 resolution QPFs (e.g., Ebert and McBride, 2000) is less serious using a longer accumulation period. Although the
 144 CReSS forecasts are made four times a day, only those from 0000 and 1200 UTC are evaluated in this study.
 145 A total of 366 target segments (0000-2400 and 1200-1200 UTC) in May-June, 2012-2014 are classified into several
 146 groups based on the observed rainfall using the following criteria, as summarized in Table 2. Groups A, B, C, and D
 147 are those periods with at least 10% rain gauges reaching 50, 25-50, 10-25, and 1-10 mm, respectively, while group X
 148 is the remaining periods with little or no rain. The full classification results (Table 3) give a total of 337 segments,
 149 excluding those under typhoon influence. Groups A-D individually account for about 18-26% and are comparable in
 150 sample size, while the driest group X is about 14% (Tables 2 and 3). These five groups are exclusive to each other,
 151 and the results without classification will be referred to as the “all” group. From group A, a subset of A+ that has ≥
 152 10% sites reaching 130 mm is identified and represents the most-rainy 4% in our sample with the highest hazard
 153 potential. The spatial distribution of mean mei-yu rainfall per season in 2012-2014, with a peak amount of about 1700
 154 mm, is shown in Fig. 3 and resembles the climatology (e.g., Yeh and Chen, 1998; Chien and Jou, 2004; Chi, 2006;
 155 Wang et al., 2017).

Year	Month	Time (UTC)	Date in month			Segments included (A-D, X)
			1-10	11-20	21-31 (or 21-30)	
2012	May	0000	XAABDXXXBB	CCXXCBAAAB	XXDXDCAAABC	31
		1200	CAADXXXBCC	CXXCBBABAD	XXXXDBAABB	31
	Jun	0000	CCBCCCCBAA	AAAAABBT	TCCDDCTTTD	23
		1200	DBCDCDBAAA	AAAABBT	TCDDDTTTD	21
2013	May	0000	CCCCBBCCDA	AACCCABAAA	CBCCDDDDDX	31
		1200	CCDCACDDA	BBCCABABA	BBCCDDDDDX	31
	Jun	0000	XXCBCCXXCB	BBABCDDDX	BDBCCXXXXD	30
		1200	XDBCCDXDCB	BABDDDXB	CBBCXXXXD	30
2014	May	0000	CCBBBBCCD	DCBDABXBAA	ADDCCCBBC	31
		1200	CBDABCDCDD	CBDAAXBAAA	BDCCDBBABC	31
	Jun	0000	XDACAABBC	TTTTTTTDD	DCBCCDXCAC	24
		1200	XACBAABDC	TTTTTTTDBD	CCBCDXDBAC	23
Total			A+: 13, A: 61, B: 75, C: 88, D: 67, X: 46 (T: 29)			337

156 **Table 3: The full classification result for all the 24-h verification periods during the three mei-yu seasons of 2012-2014. For**
157 **each month, the first (second) row gives the results of 0000-2400 (1200-1200) UTC. While the groups of A-D and X are**
158 **denoted by their corresponding letter, a bold A indicates group A+ (a subset of A) and T marks the periods influenced by**
159 **tropical cyclones and thus excluded from study.**

160 2.4 Categorical measures of model QPFs

161 As mentioned, the 24-h QPFs by CReSS are verified against the rain gauge data, at three different ranges of 0-24, 24-
162 48, and 48-72 h (days 1-3). For this purpose, objective scores computed from the standard 2×2 contingency table (or
163 the categorical matrix) at a wide range of 14 thresholds from 0.05 to 750 mm are adopted. These measures include the
164 TS (also called critical success index), bias score (BS), probability of detection (POD), and false alarm ratio (FAR),
165 respectively defined as (e.g., Schaefer, 1990; Wilks, 2011; Ebert et al., 2003; Barnes et al., 2009)

$$166 \quad TS = H/(H + M + FA), \quad (1)$$

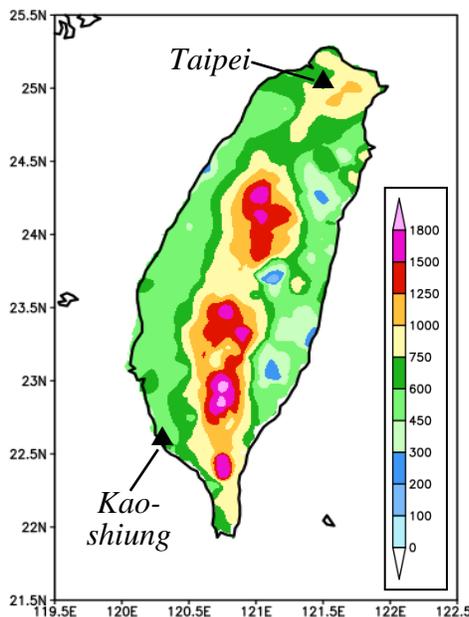
$$167 \quad BS = (H + FA)/(H + M) = F/O, \quad (2)$$

$$168 \quad POD = H/(H + M) = H/O, \text{ and} \quad (3)$$

$$169 \quad FAR = FA/(H + FA) = FA/F, \quad (4)$$

170 where H , M , and FA are the counts of hits (both observed and predicted), misses (observed but not predicted), and
171 false alarms (predicted but not observed), respectively, among a total number of N verification points. Here, $N = H +$
172 $M + FA + CN$, where CN is the correct negatives (neither observed nor predicted), and the total counts in observation
173 (O) and forecast (F) are simply $O = H + M$ and $F = H + FA$. The values of TS, POD, and FAR are all bounded by 0
174 and 1, and the higher (lower) the better for TS and POD (FAR). For BS, its value can vary from 0 to ∞ (or $N - 1$ in

175 practice), but unity is the most ideal and implies no bias. Also, $BS > (<) 1$ implies overestimation (underestimation) of
 176 the events. By interpolating the model QPFs onto the gauge sites that serve as verification points (i.e., $N \approx 440$ per
 177 segment) using the bi-linear method, the counts of H , M , FA , and CN at any given threshold can be easily obtained
 178 for each segment. Although the density of rain gauges varies to some extent (roughly every 5-10 km in the plains and
 179 ≥ 10 -20 km in the mountains, cf. Fig. 1b), their weights are assumed equal (e.g., Wang, 2014). For any group (e.g.,
 180 A+) at a given threshold, the scores are obtained from a single 2×2 table that combines the entries from all segments,
 181 so that the sample sizes are maximized (cf. Table 2, e.g., W15). This practice also remedies the issue of sampling in-
 182 homogeneity and increases the stability of results, especially toward the high thresholds, as long as the points involved
 183 in the matrix are not too few in number (cf. Table 2). Since neither the observation nor the forecast ever reached 750
 184 mm (per 24 h) during the study period, results for 13 thresholds from 0.05 up to 500 mm (the next highest threshold)
 185 are presented in later sections. Also, only 24-h QPFs are evaluated in the current study. Except for the categorical
 186 matrix, subjective visual verification is also used in the selected examples (Section 4).



187
 188 **Figure 3: Spatial distribution of mean total rainfall (mm) per mei-yu season (1 May through 30 Jun) in 2012-2014. The**
 189 **cities of Taipei and Kaoshiung are marked.**

190 3 Mei-yu QPFs in 2012-2014

191 3.1 Overall performance by the 2.5-km CReSS

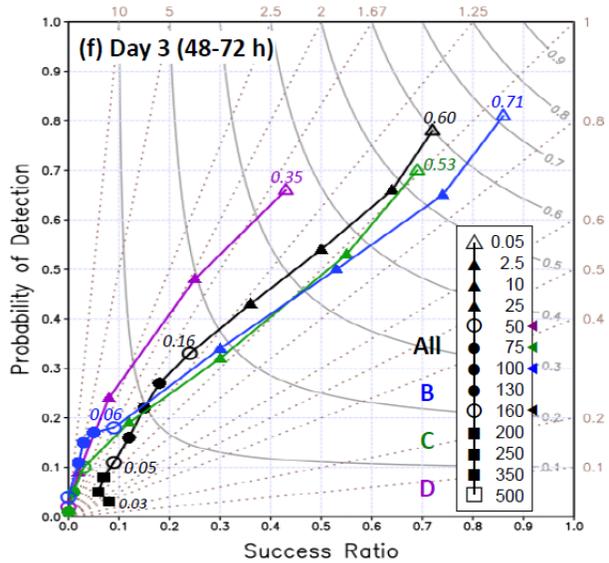
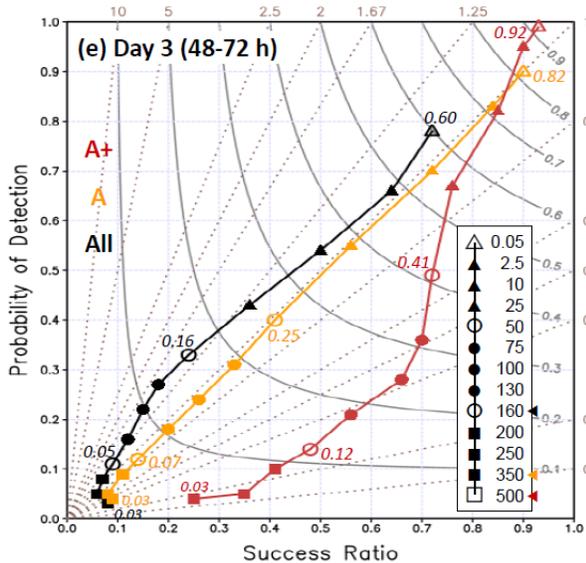
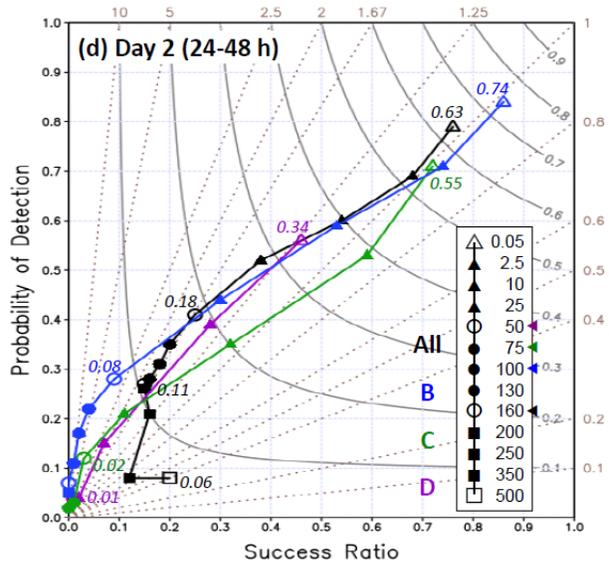
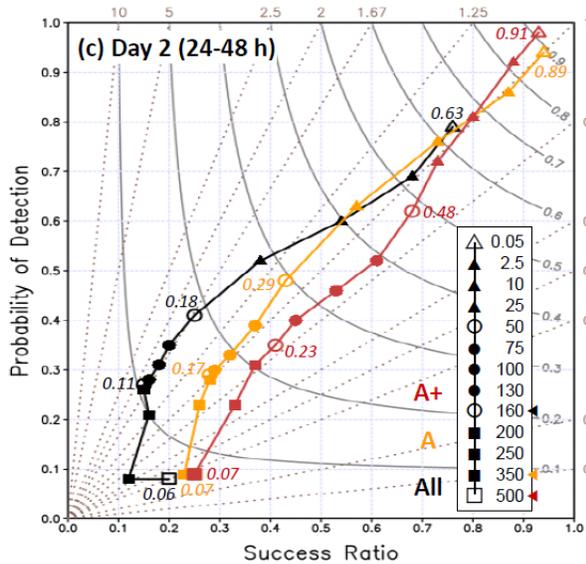
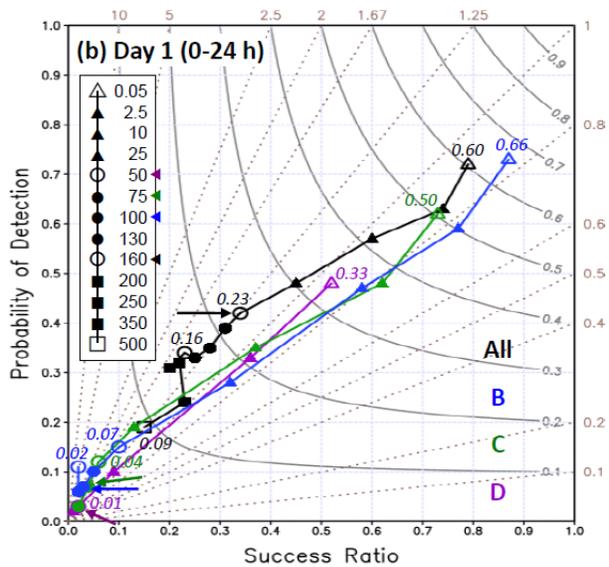
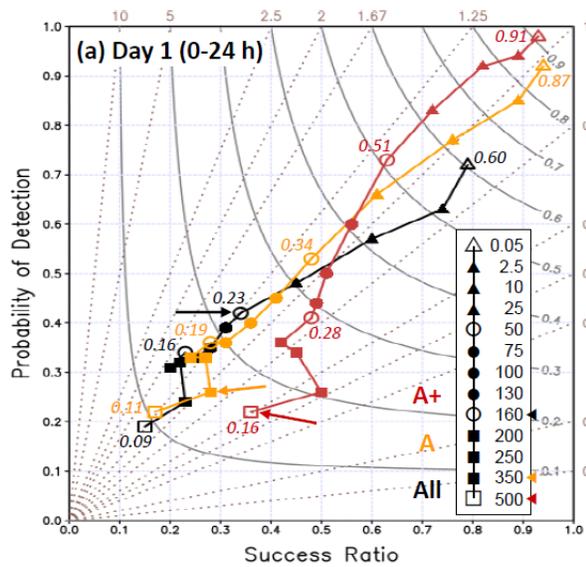
192 Following the method described above, the categorical matrices across the thresholds are obtained and the overall
 193 ability of CReSS in mei-yu QPFs during 2012-2014 is shown in Fig. 4 using the performance diagram. Proposed by
 194 Roebber (2009), the diagram uses the success ratio ($SR = 1 - FAR = H/F$) and POD as its two axes, and can also
 195 depict the TS (gray curved isopleths, higher toward upper-right) and BS (brown dotted lines) simultaneously. In Fig.
 196 4, the scores from forecasts at both 0000 and 1200 UTC for segments (of 24 h) in groups A+, A to D, and all periods
 197 (A-D plus X, cf. Table 2) at 13 thresholds are shown for ranges of day 1, 2, and 3, respectively. The “all” group (black)

198 shows the overall accuracy for all mei-yu rainfall without classification, and its TS for day-1 QPFs decreases slowly
 199 from 0.6 at 0.05 mm to 0.18 at 100 mm, 0.15 at 250 mm, and 0.09 at 500 mm (Fig. 4a). Over heavy-rainfall thresholds
 200 ≥ 160 mm, the TSs of 0.09-0.16 are considerably higher than those reviewed in section 1. Even on day 2, the TSs
 201 remain at 0.11 to 0.06 over 160-500 mm, and above 0.03 up to 350 mm on day 3 (Figs. 4c,e).
 202 When all segments are stratified by the observed event magnitude, the TSs are higher and the skill better for larger
 203 events than smaller ones, following the order of A+ then A to D for all thresholds at all three ranges without any
 204 exception (Fig. 4), while each individual curve mostly decreases with threshold when rain areas reduce in size (as
 205 shown in Fig. 5). Thus, the positive dependency of categorical measures on rainfall amount is also strong and evident
 206 in mei-yu QPFs in Taiwan, as predicted by W15. Linked to this dependency, the TSs for large events are also higher
 207 than those for the “all” group from the entire sample. For the most hazardous group A+, the TS on day 1 is 0.34 at
 208 100 mm, 0.24 at 250 mm, and 0.16 at 500 mm (per 24 h, Fig. 4a). On days 2 and 3, the corresponding TSs are 0.32,
 209 0.15, and 0.07 (Fig. 4c), and 0.25, 0.05, and 0.00 (Fig. 4e), respectively, all higher than their counterparts for the all
 210 group (except day 3 at 500 mm). Similar to some earlier studies (e.g., Chien et al., 2002, 2006; Chien and Jou, 2004;
 211 Yang et al., 2004) and based on experience, if the value of $TS \geq 0.15$, perhaps somewhat arbitrary, is used to indicate
 212 some level of accuracy, then the QPFs by the 2.5-km CReSS can reach it all the way up to 500 mm (per 24 h) on day
 213 1, 250 mm on day 2, and 130 mm on day 3. Also, for A+, A, and all groups, the TSs of day-2 QPFs stay quite close
 214 to the values on day 1, and some are even identical, from low thresholds up to 200-250 mm. For day-3 QPFs compared
 215 to day 2, the same is true up to about 130 mm (Fig. 4, left column). Such results that some skill of heavy-rainfall QPFs
 216 still exists on days 2-3 are very encouraging. On the other hand, at thresholds ≥ 50 mm, the model’s ability for B-D
 217 events (Fig. 4, right column) are limited ($TS \leq 0.08$) when the rain areas are relatively small (with $O/N \leq 6\%$, Fig. 5),
 218 but as discussed in W15, this is not important due to low hazard potential.
 219 Another fairly subtle but important feature in Fig. 4 is that the TSs of all, A, and A+ groups decrease only marginally
 220 for some heavy-rainfall thresholds, particularly on days 1-2, despite the reduction in rain-area size (left column). Some
 221 examples include the TSs for group A+ over 100-350 mm on day 1 (drops from 0.34 to 0.21), and those for group A
 222 over the same thresholds on day 1 (from 0.23 to 0.15) and over 100-250 mm on day 2 (from 0.20 to 0.14). Even on
 223 day 3, the decrease of A and “all” curves from 160 to 350 mm is rather slow, although the TSs there are only 0.03-
 224 0.07 (Fig. 4e). Such a slow decline in TSs with thresholds indicates that in a relative sense, the model is more capable
 225 to produce hits toward the rainfall maxima, which occur more frequently in the mountains (cf. Figs. 1b and 3).
 226 By definition, both POD and SR cannot be lower than the TS [cf. Eqs. (1), (3), and (4)], and the ratio of POD/SR
 227 equals to the BS (thus, $POD < SR$ if $BS < 1$ and vice versa). In Fig. 4, the PODs start at 0.05 mm from nearly perfect
 228 values of 0.98-0.99 for days 1-3 for group A+, at least 0.9 for A, and ≥ 0.72 for all segments (left column). For these
 229 three groups, the PODs at 250 mm remain at least 0.32 on day 1, 0.21 on day 2, and 0.05 on day 3. Like the TS, the
 230 POD for mei-yu rainfall indeed decreases quite significantly with forecast range (lead time), particularly toward high
 231 thresholds, mainly due to error growth and the reduction in predictability. However, even on day 3, POD and TS can
 232 still reach 0.16 and 0.07 at 130 mm (for all segments), respectively. The SR values (and thus FAR) are again the best
 233 for group A+ and ≥ 0.36 across all thresholds on day 1, including 500 mm (Fig. 4a). On day 2, the SRs for A+ over
 234 130-500 mm decrease but not by too much, and the values over 10-250 mm even increase on day 3 (Figs. 4c,e). Often,

235 the SR for A+ is considerably higher than those for A and all events regardless of forecast range, particularly over
236 heavy-rainfall thresholds. Overall, the model also produces higher POD and SR (i.e., lower FAR) for larger events
237 compared to smaller ones at all thresholds and all three forecast ranges in Fig. 4, with only a few exceptions after close
238 inspection. This indicates that the high-resolution CReSS not only produces larger rainfall for large rainfall events,
239 which leads to a higher TSs, but also produces larger rainfall for small events, which leads to lower SR values. In
240 summary, as for typhoon rainfall (W15), the 2.5-km CReSS is the most skillful in predicting the largest events in the
241 mei-yu season in Taiwan..

242 Next, the BS values are examined for over/under-prediction (i.e., above/below the diagonal line) in Fig. 4, where the
243 threshold with O/N falling below 1% is marked to indicate values that might be potentially unstable and less
244 meaningful. For day-1 QPFs, the BSs for all segments suggest slight under-prediction for low thresholds ≤ 10 mm
245 (per 24 h), but some over-prediction ($BS \approx 1.25-1.5$) across 50-350 mm (Fig. 4a). On the contrary, the model shows
246 slight over-prediction over 0.05-75 mm for the largest events of group A+ (with BSs up to 1.15), but under-prediction
247 toward higher thresholds, with the lowest BS of 0.52 at 350 mm. Mostly between the two curves mentioned above,
248 the curve for group A stays closer to unity and is more ideal across nearly all thresholds (Fig. 4a). For B-D groups
249 (Fig. 4b), their characteristics are similar to the All group, with BSs of 0.8-1.0 at low thresholds but generally some
250 over-prediction across higher thresholds. However, their BS values rarely exceed 2.5, unless the O/N values drop to
251 below 1%. The situation for BSs between different groups remains similar on days 2 and 3 (Figs. 4c-f), and the over-
252 forecasting across the middle thresholds in group A (at all ranges) can be confirmed to come mainly from groups B-
253 D, as groups A+ and A exhibit little or a much less tendency for over-prediction there (Fig. 4).

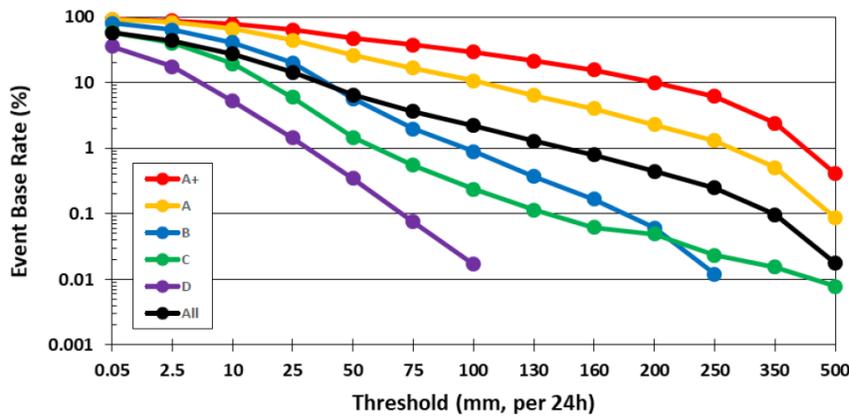
254 Toward the longer ranges of days 2 and 3, the BS values in general become smaller, particularly for the larger groups
255 (Fig. 4, left column). Thus, the over-prediction in group A is reduced and the under-prediction in A+, which is the
256 most important group, becomes more evident, especially toward the high thresholds (Figs. 4c,e). For example, the BS
257 of day-2 QPFs for A+ is ideal and ≥ 0.8 up to 200 mm but declines to about 0.35 at 500 mm, but it is already below
258 0.4 at 130 mm on day 3. This indicates that for larger events, the error growth with lead time in the model tends to
259 become less rainy, as reflected in the decrease in BS. Thus, the probability to under-forecast peak rainfall rises with
260 lead time. For smaller events that do not produce much rainfall (i.e., B-D and X), a similar tendency does not exist or
261 is weaker, and BS tends to be greater than unity. So, to say the least, one needs to practice caution in the interpretation
262 of BS, which can also become unstable when O/N approaches zero (which inevitably happens at certain thresholds).



264 **Figure 4: Performance diagrams of 24-h QPFs for (a),(b) day 1 (0-24 h), (c),(d) day 2 (24-48 h), and (e),(f) day 3 (48-72 h)**
 265 **by the 2.5-km CReSS, at 13 rainfall thresholds (inserts) from 0.05 to 500 mm, during three mei-yu seasons (May-Jun) in**
 266 **2012-2014 in Taiwan. Results for groups A+, A, and All (All, B, C, and D) are plotted in left (right) column with different**
 267 **colors. TS values are labelled at fixed thresholds of 0.05, 50, 160, and 500 mm (open symbols) or selected endpoints (smaller**
 268 **fonts), and data points with TS = 0 at high thresholds are omitted. For each group, the threshold where the observed rain-**
 269 **area size (O/N) falls below 1% is labeled in insert, and also marked by an arrow in (a),(b).**

270 **3.2 Improvement in heavy-rainfall QPFs**

271 To assess the improvement in heavy-rainfall QPFs in mei-yu season, our results in Fig. 4 are compared to Fig. 2 and
 272 those reviewed in Section 1. However, the differences in model resolution should be noted. Overall, the “all” curves
 273 in Fig. 4 indicate that the 2.5-km CReSS exhibits better skill than those reviewed in Section 1 (with Δx as fine as 5
 274 km at most), especially at thresholds above 100 mm (e.g., TS = 0.15 at 250 mm and 0.09 at 500 mm for day 1). With
 275 even higher TSs for larger and more hazardous events (groups A and A+), the improvement of heavy-rainfall QPFs
 276 in the present study from earlier results is therefore quite clear and dramatic. The physical explanation will be further
 277 elaborated and discussed later in Section 4.



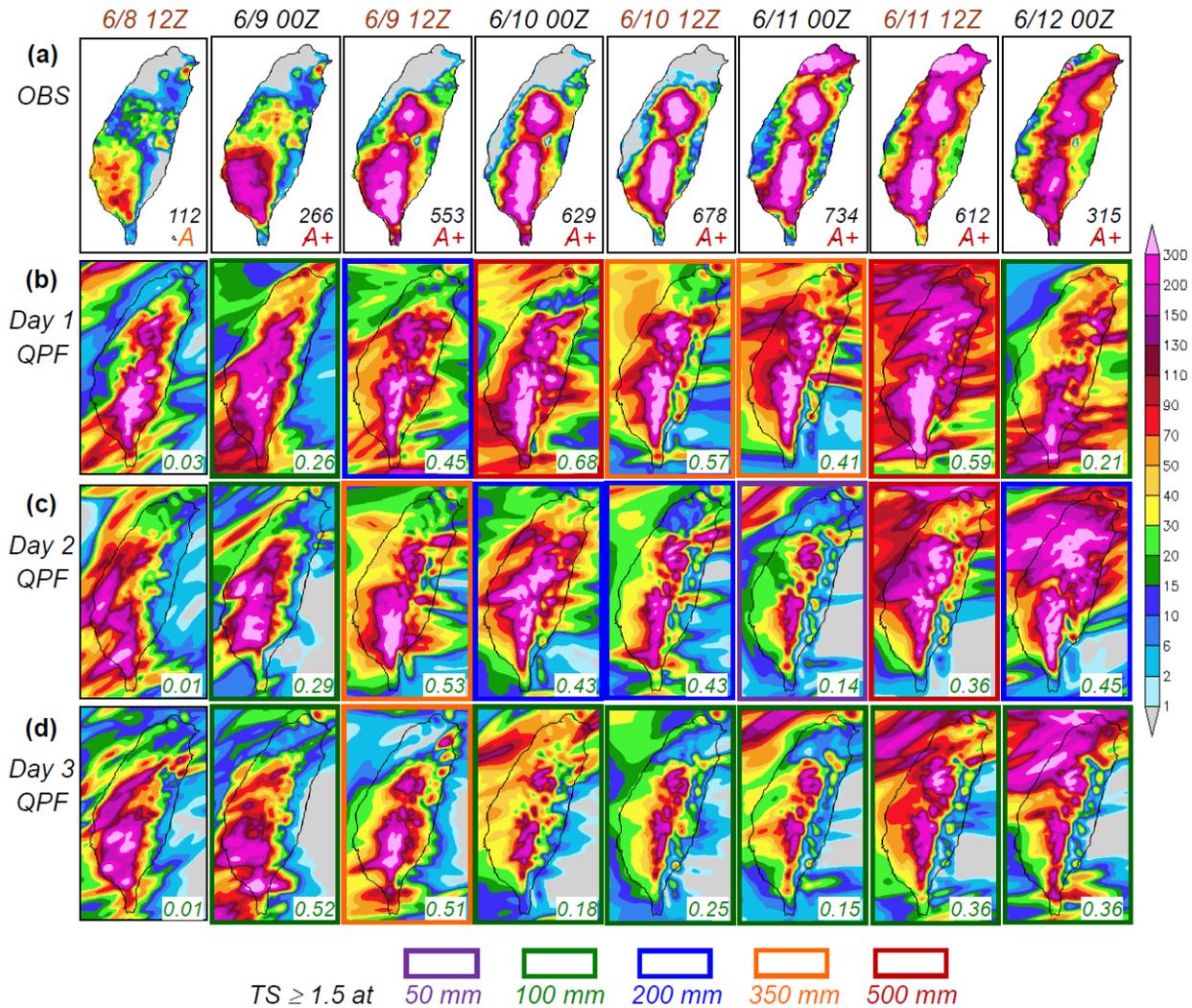
278 **Figure 5: Observed rain-area size or base rate (O/N , %) of 24-h rainfall (same for days 1-3) in logarithmic scale used to**
 279 **compute the scores in Fig. 4**
 280

281 **4 Examples of Model QPFs**

282 Given the success of the CRM in its overall performance, some examples of CReSS forecasts are selected and
 283 presented in this section. The main goal here is two-fold: 1) To illustrate how the model behaves and captures the
 284 rainfall in individual forecasts, and thus 2) to identify where such a CRM has a better capability in QPFs and where it
 285 has limitations in Taiwan. Since our focus is on heavy rainfall, the event during 9-12 June 2012, the largest during our
 286 study period, is chosen for illustration.

287 The event of 9-12 June 2012 spanned four days and contributes more than half the segments in group A+ (7 in 13, cf.
 288 Table 3). In Fig. 6a, the observed 24-h rainfall distributions over Taiwan are shown every 12 h, from 1200-1200 UTC
 289 8 June to 0000-2400 UTC 12 June 2012. Except for the first forecast period, all seven segments are qualified as A+
 290 and five have a 24-h peak rainfall over 500 mm (those since 1200 UTC 9 June). Three rainfall maxima from this
 291 lengthy event exist: over southern CMR, near the intersection of CMR and SMR in central Taiwan, and over northern
 292 Taiwan (Fig. 6a). The rain at the two mountain centers (cf. Fig. 1b) is much more persistent than that in northern

293 Taiwan, which concentrated mainly over a 10-h period beginning 1400 UTC 11 June (Wang et al., 2016b). The
 294 southwestern plains also received considerable rainfall, especially around 9 June (Fig. 6a).



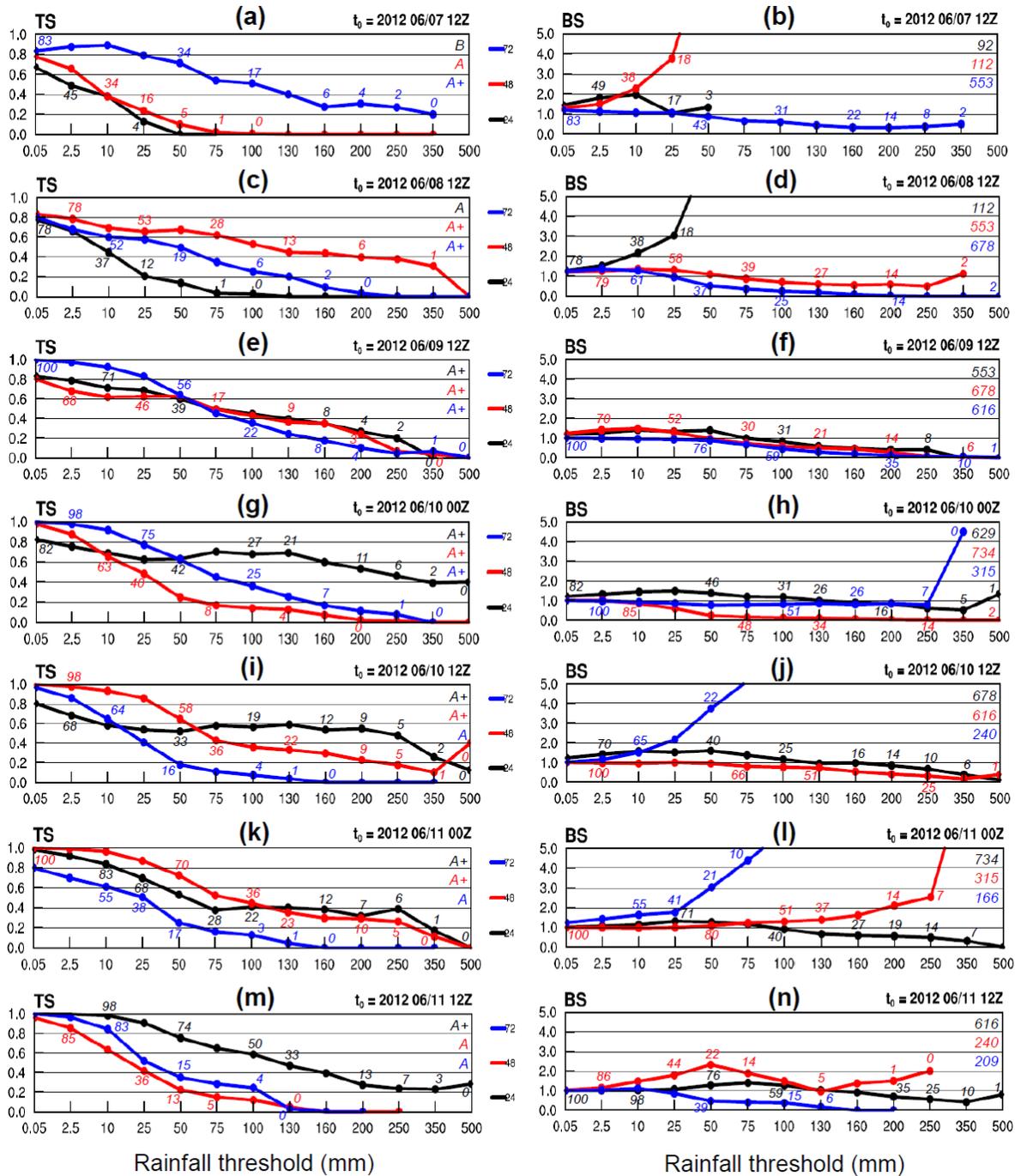
296 **Figure 6:** (a) The observed 24-h accumulated rainfall (mm, scale on the right) over Taiwan from 1200 UTC 8
 297 Jun to 0000 UTC 13 Jun 2012, given every 12 h (from left to right), with the beginning time of accumulation
 298 (UTC) labeled on top (black for 0000-2400 UTC and brown for 1200-1200 UTC). (b) Day-1 (0-24 h), (c) day-2
 299 (24-48 h), and (d) day-3 (48-72 h) QPFs valid for the same 24-h periods as shown in (a) by the 2.5-km CReSS
 300 (starting at 0000/1200 UTC under black/brown headings). In (a), peak 24-h rainfall (mm) and classification
 301 group are labeled. In (b)-(d), thick boxes in purple, green, blue, orange, and scarlet denote forecasts having a
 302 TS ≥ 0.15 at the threshold of 50, 100, 200, 350, and 500 mm (per 24 h), respectively, and the TS at 100 mm is
 303 also given (lower right corner).

304 The 24-h QPFs produced by the 2.5-km CReSS (at 0000 or 1200 UTC) in real time targeting the same periods as in
 305 Fig. 6a, at the ranges of days 1-3 are presented in Figs. 6b-d, with the general quality expressed by the TS at 100 mm
 306 (lower right corner inside panels) and thickened outline for TS ≥ 0.15 at the threshold of 50, 100, 200, 350, or 500
 307 mm. The day-1 QPFs (Fig. 6b) are made from the forecasts starting (with initial time t_0) at the time of the heading,
 308 while day-2 (Fig. 6c) and day-3 QPFs (Fig. 6d) are those made 24 and 48 h earlier (for the same target period),
 309 respectively. In Fig. 6, this extreme and long-lasting event was generally well captured by the model, especially on

310 day 1 where the overall rainfall pattern and TS both tend to be better, as expected. The best day-1 QPF is for 0000-
 311 2400 UTC 10 June (TS = 0.68 at 100 mm and 0.40 at 500 mm), followed by the one for 1200 UTC 11-12 June (TS =
 312 0.59 at 100 mm and 0.29 at 500 mm, columns 4 and 7, Fig. 6b). At longer ranges on days 2 and 3, the rainfall
 313 magnitudes produced over the mountains and southwestern plains are also comparable to observations, but the event
 314 starts somewhat earlier and becomes less rainy during 10-11 June showing under-forecast (Figs. 6c,d). As a result, the
 315 TSs for the segments starting at 1200 UTC 8 June and during 10-11 June (columns 1 and 4-7) mostly increase from
 316 longer to shorter ranges, i.e., with better QPFs at later times. This relationship with range does not hold true for the
 317 other segments, among which the day-3 and day-2 QPFs for the period of 1200 UTC 9-10 June (TS \geq 0.51-0.53 at
 318 100 mm and 0.20-0.31 at 350 mm) and the day-2 QPF for 1200 UTC 11-12 June (TS = 0.40 at 500 mm) are particularly
 319 impressive (columns 3 and 7). Compared to the rain over the terrain, the maximum across Taipei in northern Taiwan
 320 during 11-12 June was largely over lower and flatter regions (cf. Figs. 1b and 2) and more challenging for the model
 321 to predict at the right location (Fig. 6), an aspect that will be further elaborated on later. Note, nevertheless, that since
 322 the mountain regions are the only places where rainfall amounts reach 300 mm in both the observations and the model
 323 (Fig. 6), any hits at and above this threshold occur in the mountains.

324 Figure 7 shows the TS and BS of day-1 to day-3 QPFs from the runs made at a series of initial times, including 1200
 325 UTC of 7-9 June and the next four from 0000 UTC 10 to 1200 UTC 11 June (top to bottom), and our focus is mainly
 326 over the thresholds \geq 100 mm. Inside the panels, the observed event base rate (O/N , i.e., rain-area size, identical at the
 327 same threshold for the same target period) and the hit probability (H/N , note that $H/N \leq O/N$) are given at selected
 328 points. Figures 7a-7f provide some examples on how the model did in predicting the beginning of the event (cf. Fig.
 329 6, columns 1-3). As mentioned, the day-3 QPF made from 1200 UTC 7 June (Figs. 7a,b, blue curves) and day-2 QPF
 330 made one day later (Figs. 7c,d, red curves), both targeting 1200 UTC 9-10 June, are of fairly high quality. With rain
 331 areas (O/N) occupying 31%, 14%, and just 2% of Taiwan at 100, 200, and 350 mm, the day-2 QPF in Fig. 7c, with
 332 BS \approx 0.6-1.1 (Fig. 7d), yields TSs of 0.53-0.31 at these thresholds. The day-3 QPF with t_0 at 1200 UTC 7 June, with
 333 less predicted rain and BS \approx 0.3-0.6 (cf. Fig. 6d, column 3), the TSs are 0.51-0.2 (Figs. 7a,b). With TS at least 0.2 at
 334 350 mm (an amount predicted only in southern CMR), both QPFs (for 1200 UTC 9-10 June) are quite good. Valid for
 335 periods with varying magnitude (B, A, and A+), the forecasts in Figs. 7a,b are also good examples to illustrate the
 336 dependency property (Fig. 4 and W15), where the rainfall amount apparently exhibits a larger influence on the scores
 337 than the forecast range. In Figs. 7e,f, the TS curves at the three ranges (all for A+ events) are closer.

338 In Fig. 8, the actual forecast near Taiwan between 42 and 69 h, from the run made at 1200 UTC 7 June, is compared
 339 with radar observations every 6 h to examine general rainfall locations. While a wind-shift line existed off eastern
 340 Taiwan, the surface mei-yu front was well to the north with prefrontal low-level southwesterly flow impinging on the
 341 island during this period (also Wang et al. 2016b). Active convection constantly developed over the mountains in
 342 central and southern Taiwan and moved from the upstream ocean into the southwestern plains, and this scenario was
 343 well captured by the 2.5-km CReSS (Fig. 8), yielding a high-quality QPF on day 3 despite some under-forecast at
 344 thresholds \geq 75 mm (cf. Figs. 7a,b).



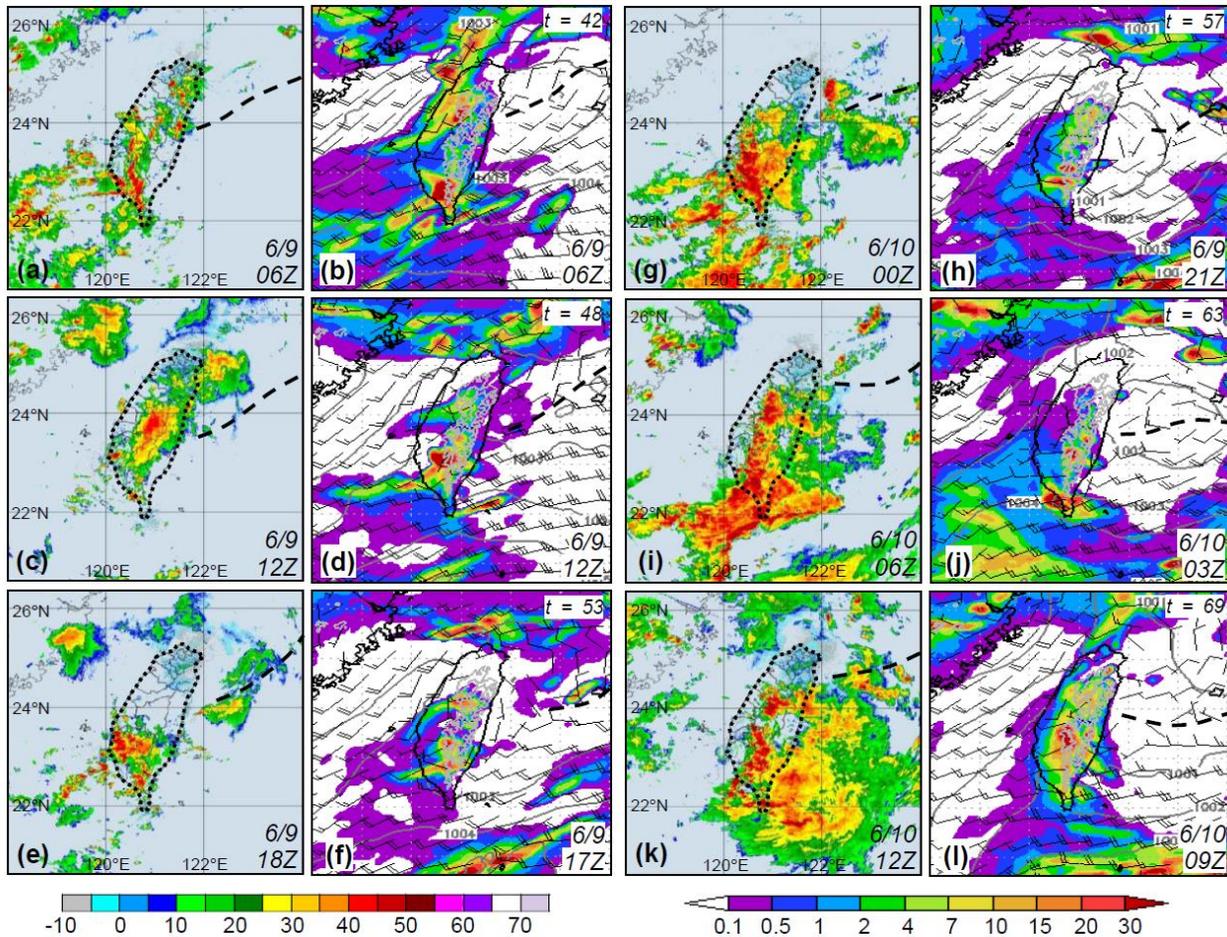
345

346 **Figure 7: (a) TS and (b) BS of day-1 (black), day-2 (red), and day-3 (blue) QPFs made at 1200 UTC 7 Jun 2012 as a function**
 347 **of rainfall threshold (mm, per 24 h). The hit rate (H/N , %, rounded to integer) at selected points and the classification group**
 348 **for each day are labeled in the left panel (for TS). The observed base rate (O/N , %) and peak 24-h rainfall (mm) are labeled**
 349 **in the right panel (for BS). (c),(d) to (m),(n) As in (a),(b), except for the QPFs made at (c),(d) 1200 UTC 8 Jun, (e),(f) 1200**
 350 **UTC 9 Jun, (g),(h) 0000 and (i),(j) 1200 UTC 10 Jun, and (k),(l) 0000 and (m),(n) 1200 UTC 11 Jun, 2012, respectively.**

351

352 In the four following forecasts made on 10-11 June (Figs. 7g-n), the dependency on event magnitude exists, but the
 353 QPFs made for A+ periods tend to have higher TSs above 75-100 mm at the shorter ranges (Figs. 7g,i,k). The TSs of
 354 these day-1 QPFs can be as high as 0.48 at 250 mm and 0.40 at 500 mm. At 350-500 mm, such high TS occurs with
 O/N below < 10% (or even only 1%), and thus indicates remarkable model accuracy in predicting the peak rainfall at

355 the correct location in the mountains in this event. Over thresholds ≥ 200 mm, BS values in Fig. 7 indicate that under-
 356 prediction for this extreme event occurs much more often than over-prediction, while they also tend to be closer to
 357 unity (with less under-forecast) for QPFs achieving higher TSs. Consistent with Fig. 4, an over-prediction is more
 358 likely to happen for smaller events (A or below), across low thresholds below 50 mm, and/or when the rain area
 359 becomes small. In Fig. 7, for example, $BS \geq 2$ at high thresholds for A+ group occurs only when O/N approaches zero
 360 (Figs. 7h,n), with the only exception in Fig. 7l on day 2. Overall, the model does not have a tendency to over-predict
 361 such a large event (cf. Fig. 6).

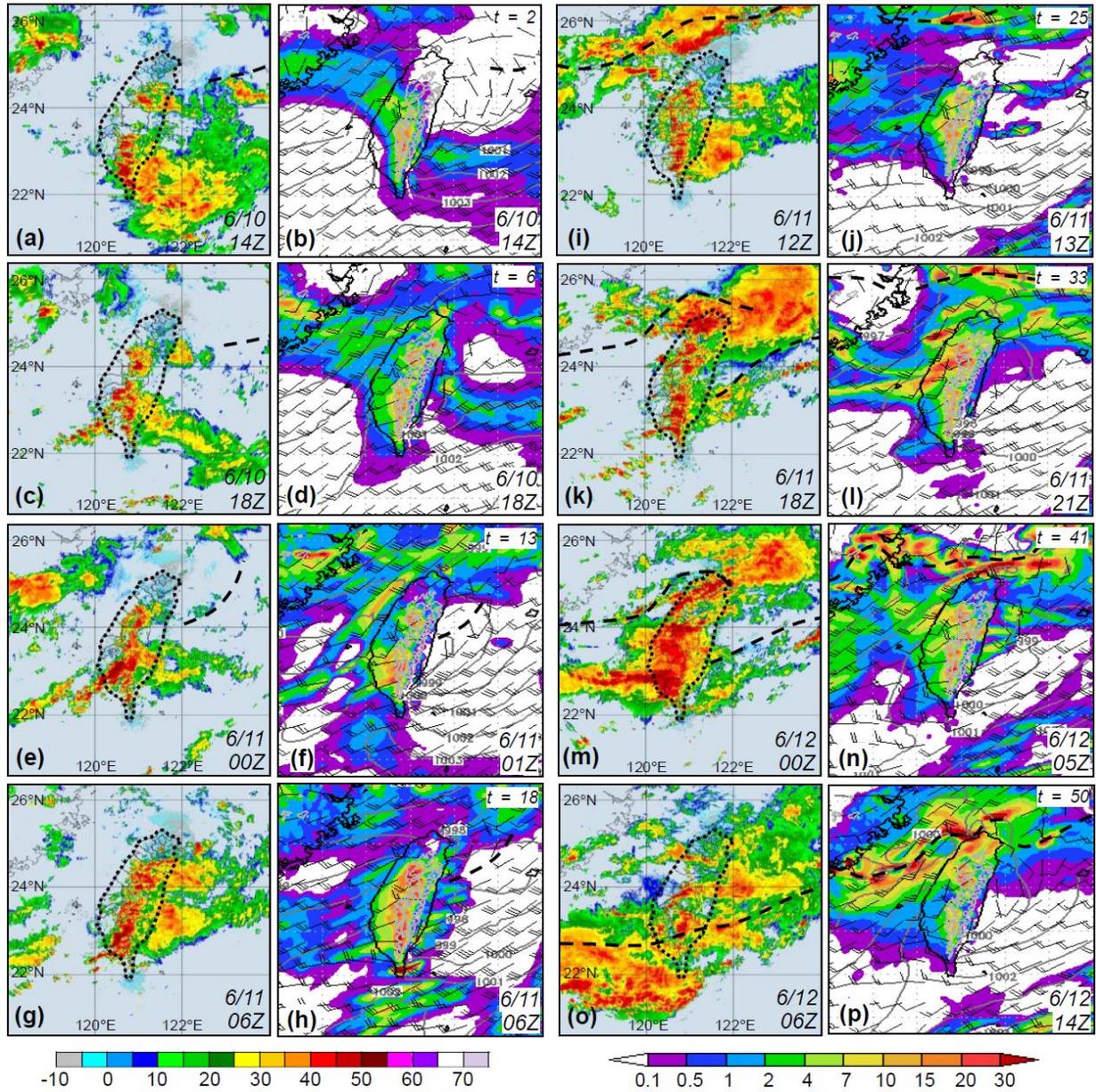


362
 363 **Figure 8: (First and third column) Radar reflectivity composite (dBZ, scale at bottom left) in the Taiwan area (width**
 364 **roughly 600 km) every 6 h from (a) 0600 UTC 9 Jun to (k) 1200 UTC 10 Jun, 2012 (original plots provided by the CWB).**
 365 **(Second and fourth column) The CReSS forecast, starting from 1200 UTC 7 Jun 2012, of sea-level pressure (hPa, every 1**
 366 **hPa, over ocean only), surface wind (kts, barbs, at 10 m), terrain height at 1 and 2 km (gray contours), and hourly rainfall**
 367 **(mm, color, scale at bottom right) valid at the time or within 3 h of the radar composite as labeled [in UTC (forecast time**
 368 **in h) at lower (upper) right corner] over the same area. The thick dashed lines mark the position of surface frontal or wind-**
 369 **shift line, based on NCEP gridded analyses for the observation (outline of Taiwan also highlighted).**

370 The forecasts on days 1-2 produced by the run starting at 1200 UTC 10 June are compared with radar observations in
 371 Fig. 9. Together with Fig. 8, the radar panels cover the wettest 72 h (0600 UTC 9-12 June) of the entire event. During
 372 day 1 (Figs. 9a-h), the scenario remains similar to Fig. 8, and the model again was able to capture the mountain rainfall.
 373 The convection moving in from the Taiwan Strait, however, was too active and the rain along the western coast on
 374 day 1 was over-predicted with BSs ≈ 1.2 -1.6 from 0.05 up to 100 mm (cf. Fig. 7j). Note that in Fig. 7, some over-

375 prediction across low thresholds can also exist for group A+ and lowers the TS, which otherwise can often exceed 0.8
376 at and below 25 mm. In any case, the model's performance over the low thresholds is of secondary importance.
377 Since 1200 UTC 11 June, the mei-yu front gradually approached northern Taiwan, and its western section moved
378 rapidly across the island after about 0000 UTC 12 June (Figs. 9i-p). Studied by Wang et al. (2016b), the heavy rainfall
379 in northern Taiwan (during 1400-2400 UTC) was caused by quasi-linear convection that developed south of the front
380 (Figs. 9i,k,m), along a convergence zone between the low-level flow blocked and deflected by Taiwan's topography,
381 and unblocked flow further to the northwest (but still prefrontal) in the environment (also e.g., Li and Chen, 1998;
382 Yeh and Chen, 2002; Chen et al., 2005; Wang et al., 2005). In the model forecast, with apparent errors in the position
383 and moving speed of the front (Figs. 9i-p), it is highly challenging to produce a similar system at the correct location
384 and time even when the overall scenario surrounding northern Taiwan are reasonably predicted. In the simulation of
385 Wang et al. (2016b), the rainbands cannot be fully captured even with a finer grid of $\Delta x = 1.5$ km and the NCEP final
386 analyses as IC/BCs. Likely mainly linked to the IC/BCs, the position error of the front in this case is still a major error
387 source for the rainfall associated with the front. Thus, although the model did indicate a real possibility of heavy
388 rainfall in northern Taiwan in Fig. 9, the high TS of 0.4 at 500 mm on day 2 (Fig. 7i) came from the mountains, where
389 the rainfall prediction is clearly more accurate (cf. Figs. 6a,c, column 7), consistent with Walser and Schär (2004). Of
390 course, the day-1 QPF with $t_0 = 1200$ UTC 11 June performed better in northern Taiwan than our example, but the
391 goal here is to illustrate the relatively high accuracy to predict heavy rainfall phase-locked to the topography versus
392 the low accuracy for rainfall produced by transient systems over low-lying plains.

393 The above example, together with other cases including those on 20 May 2013 and 20-21 May 2014, (cf. Table 3, not
394 shown), suggests a lower accuracy and a more challenging task for model QPFs to capture the heavy rainfall produced
395 by transient systems often in close association with the mei-yu front, compared to topographic rainfall in Taiwan.
396 Even though the overall scenario is reasonably and realistically predicted (cf. Figs. 8 and 9), some position errors on
397 the mei-yu front are almost inevitable and the intrinsic predictability can limit the accuracy of the QPF (e.g., Hochman
398 et al., 2021). Also, for such rainfall caused by transient systems, categorical statistics are known to be less effective in
399 verifying model QPFs (e.g., Davis et al., 2006; Wernli et al., 2008; Gilleland et al., 2010). However, for the quasi-
400 stationary, phase-locked rainfall over the topography in the majority of large events (in both mei-yu and typhoon
401 seasons, e.g., Chang et al., 1993; Cheung et al., 2008) in Taiwan, they are still valid and useful as shown herein. As
402 model resolution increases, both the topography and deep convection can be better resolved, leading to improved
403 QPFs over the terrain.



404

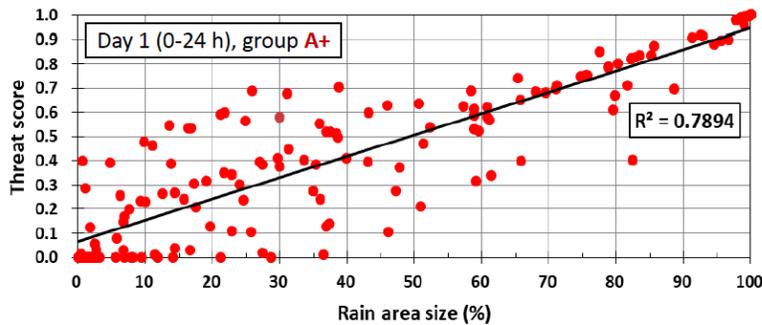
405 **Figure 9:** As in Fig. 8, but showing (columns 1 and 3) radar reflectivity composite (dBZ) at (a) 1400 UTC 10 Jun and every
 406 6 h from (c) 1800 UTC 10 Jun to (o) 0600 UTC 12 Jun, 2012, and (columns 2 and 4) the CReSS forecast, starting from 1200
 407 UTC 10 Jun 2012, of sea-level pressure (hPa), surface wind (kts), and hourly rainfall (mm) valid at the time or within up to
 408 8 h (towards the end) of the radar composite (as labeled).

409 **5 Dependency of Categorical Scores to Event Size**

410 In Section 3, a positive dependency in categorical measures by CReSS, including TS, POD, and FAR, on rainfall
 411 amount is shown for the mei-yu regime in Taiwan, as predicted. Also discussed in W15, this property arises mainly
 412 due to the positive correlation between the scores and rain-area sizes, as illustrated in Fig. 10 with a correlation
 413 coefficient $r = 0.89$ for the mei-yu regime. However, to explore whether the model is indeed more skillful in predicting

414 larger rainfall events, further analysis with the factor of rain-area size removed is needed. Different from W15, our
415 approach here is described below.

416 For each segment, the statistics (H , M , FA , and CN) at 13 fixed thresholds of 0.05-500 mm, each occurring at a certain
417 O/N (if the threshold \leq the observed peak amount), are known. The observed base rate (0-100%) is divided into bins
418 every 5% except at 0-5%, where it is subdivided into 0-0.5, 0.5-2, and 2-5% to give more comparable sample size.
419 For each group (A+ or A-D), the statistics are then summed for each bin regardless of their rainfall threshold. Thus,
420 those in the same bin come from rain areas with similar sizes. In Fig. 11a, the distribution of total counts of thresholds
421 across O/N is plotted, and the larger events toward A+ are more capable to produce rain areas larger in size (say, $\geq 60\%$
422 of Taiwan). Also, the counts remain mostly around 50 for $O/N \geq 40\%$, then rise to 200-300 with $O/N \leq 10\%$.

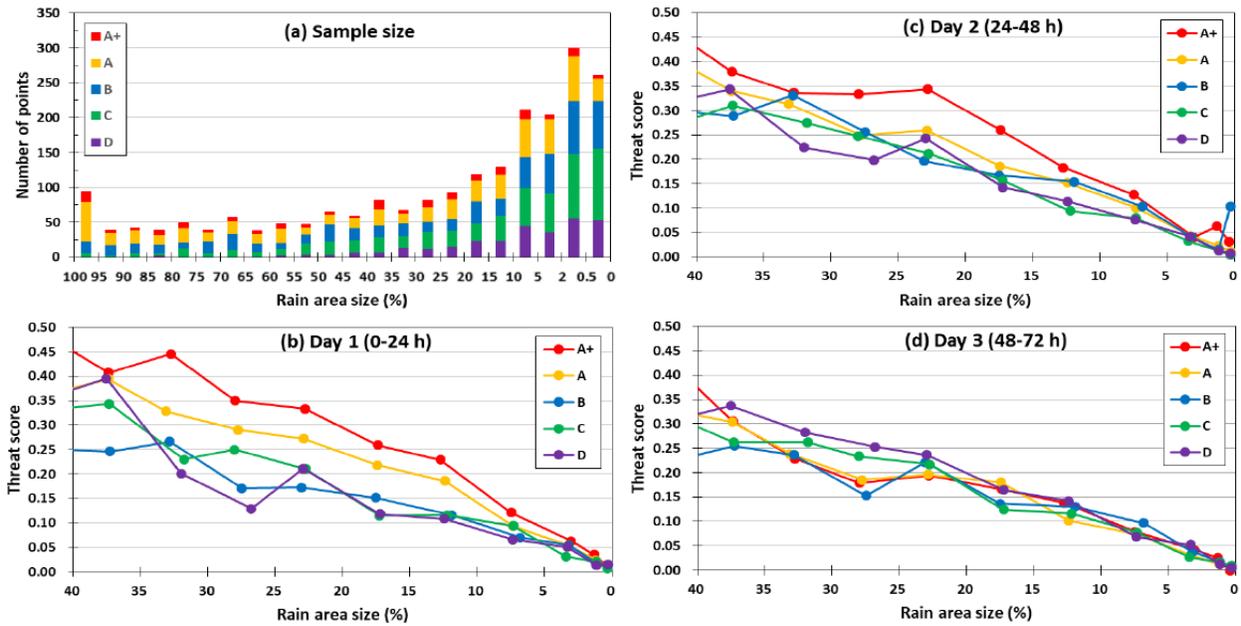


423
424 **Figure 10: Scatter plot of TS versus observed rain-area size (%) from day-1 QPFs for group A+ from 0% to 100% (from**
425 **high to low rainfall threshold). The square of correlation coefficient (R^2) is given.**

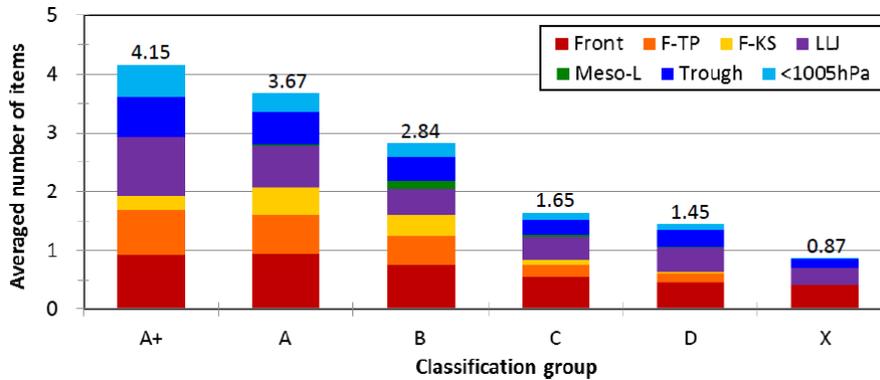
426 Due to fewer samples at larger O/N values, the TSs for different groups (from a single 2×2 table for each bin) are
427 presented only for $O/N \leq 40\%$ in Figs. 11b-d. While the scores for B-D are roughly the same, the TSs for A are clearly
428 higher compared to them on day 1, and those for A+ are again higher compared to A on days 1 and 2 over most part
429 of this range, sometimes by 0.05-0.1, when the factor of rain-area size is removed (Figs. 11b,c). On day 3 (Fig. 11d),
430 however, the TSs for larger events (A+ and A) show no particular advantage. Therefore, similar to typhoons in W15,
431 the 2.5-km CReSS is more skillful in predicting the larger mei-yu events in Taiwan within 2 days, over the heavy-
432 rainfall area (again, mainly over the mountains).

433 The higher TSs and better skill for large events at O/N within 40% (Fig. 11) are most likely linked to the more favorable
434 conditions at synoptic to meso- α scale, which the model is capable to capture with higher accuracy (e.g., Walser and
435 Schär, 2004). To briefly elaborate on this aspect, seven items on the checklist used by CWB forecasters in the mei-yu
436 season as a guidance to issue heavy-rainfall warning (e.g., Wang et al., 2012a) are selected, and their occurrence
437 frequency, judged using surface weather maps and NCEP gridded analyses at the starting time of each 24-h periods
438 are compiled for different groups. These items include: 1) presence of surface mei-yu front inside 20° - 28° N, 118° -
439 124° E; 2) Taipei (cf. Fig. 3) within 200 km south and 100 km north of the front; 3) Kaoshiung (cf. Fig. 3) within 200
440 km south of the front; 4) presence of low-level jet (LLJ) inside 18° - 26° N, 115° - 125° E at 850 or 700 hPa; 5) presence
441 of mesolow near Taiwan; 6) Taiwan inside a low pressure zone; and 7) the mean sea-level pressure in Taiwan is below
442 1005 hPa. The results (Fig. 12) indicate that among the seven items, an average of 4.15 items are met in group A+,
443 and this figure gradually declines toward smaller groups, from 3.67 in A, 2.84 in B, and finally to only 0.87 in X.

444 Thus, as expected, the synoptic and meso- α -scale conditions tend to be more favorable in larger events, which in
 445 general also correspond to higher TS values (Figs. 3 and 11) in combination with the orographic forcing in Taiwan.



446
 447 **Figure 11:** (a) The distribution of data points in the bins of observed rain-area size (%), every 5% from 100% to 5%, then
 448 2-5%, 0.5-2%, and < 0.5%; same for days 1-3) among groups A+ and A to D. (b)-(d) The TS of 24-h QPFs for (b) day 1 to
 449 (d) day 3, respectively, as a function of observed rain-area size between 40% and 0%.



450
 451 **Figure 12:** The average number of items met, among the seven items on the checklist, at the starting time of 24-h segments
 452 for different classification groups (from A+, A to D, and X), with the value labeled on top. Following the order (bottom to
 453 top), the seven items are: presence of surface mei-yu front (front), front near Taipei (F-TP), front near Kaoshiung (F-KS),
 454 presence of LLJ, mesolow (meso-L), Taiwan inside a low pressure zone (trough), and the mean sea-level pressure lower
 455 than 1005 hPa (<1005 hPa), respectively. The items are plotted in different colors (see insert) to show their proportion.

456 6 Summary and Concluding Remarks

457 In this study, the QPFs at the ranges of 1-3 days by the 2.5-km CReSS during three mei-yu seasons (May-June) in
 458 Taiwan of 2012-2014 are evaluated using categorical statistics, with an emphasis on heavy to extreme rainfall events
 459 (100-500 mm per 24 h). Overall, the TSs of day-1 QPFs for all events (no classification) at thresholds of 100, 250,
 460 and 500 mm are 0.18, 0.15 and 0.09, respectively. Compared to previous and contemporary results from models at

461 lower resolutions for mei-yu season in Taiwan, where the TSs are no higher than 0.1 at 100 mm and 0.02 at 250 mm
462 and beyond (Section 1, e.g., Hsu et al., 2014; Li and Hong, 2014; Su et al., 2016; Huang et al., 2016), the results herein
463 show considerable improvements by the 2.5-km CReSS, especially over the heavy-rainfall thresholds.
464 Moreover, the ability to represent the extreme and top events (group A+) in terms of the TS are much higher when a
465 proper classification based on observed rain area size (i.e., event magnitude) is used. For the top 4% and most
466 hazardous mei-yu events, the day-1 QPFs have TSs of 0.34, 0.24, and 0.16, respectively. The QPFs for larger events
467 also exhibit higher POD, lower FAR, and higher TS than smaller ones, across nearly all thresholds at all ranges of days
468 1-3. Thus, the positive dependency in categorical scores on the overall rainfall amount also exists in mei-yu regime in
469 Taiwan, as predicted by W15 (and Wang, 2016).
470 For a selected case study, the improvement by the 2.5-km CReSS in Taiwan is shown to lie in an improved ability to
471 capture the phase-locked topographic rainfall at its correct location in larger events at heavy-rainfall thresholds. For
472 rainfall in the mountains, the QPFs tend to be more accurate as the CRM can better resolve both the terrain and
473 convection (e.g., also Walser and Schär, 2004; Roberts and Lean, 2007). In contrast, the accuracy of QPFs for
474 concentrated rainfall caused by transient systems (such as frontal squall lines) could not be demonstrated probably
475 due to the difficulty to predict at the correct time and location owing to nonlinearity, even though a realistic scenario
476 is produced. Overall, the high-resolution models showed a higher QPF accuracy in categorical statistics for extreme
477 events than coarser resolution models over the geographic region of Taiwan as demonstrated here (and in W15). Such
478 QPFs can be helpful to hazard preparation and mitigation.

479 **Code and data availability**

480 The model used in this study is called Cloud-Resolving Storm Simulator and its website for downloading model and
481 user's guide is at http://www.rain.hyarc.nagoya-u.ac.jp/~tsuboki/cress_html/index_cress_eng.html. And the rainfall
482 figures of Model is at the website <http://cressfcst.es.ntnu.edu.tw/>.

483 **Author contribution**

484 Chung-Chieh Wang designed the experiments and Pi-Yu Chuang carried them out. Chih-Sheng Chang operated the
485 real-time model forecasting and Kazuhisa Tsuboki created the model code. Shin-Yi Huang helped with some figures
486 and Guo-Chen Leu provided some CWB results. Chung-Chieh Wang prepared the manuscript with contributions from
487 all co-authors.

488 **Competing interests**

489 The authors declare that they have no conflict of interest.

490 **Acknowledgments**

491 Constructive comments and suggestions from two anonymous reviewers and Prof. George Tai-Jen Chen of the
492 National Taiwan University are appreciated. The first author, CCW, wishes to thank assistants Ms. Y.-W. Wang, Mr.
493 T.-C. Lin, and Mr. K.-Y. Chen for their help on this study. The CWB is acknowledged for providing the observational
494 data, the radar plots, and the QPF verification results in Fig. 2. The National Center for High-performance Computing
495 (NCHC) and the Taiwan Typhoon and Flood Research Institute (TTFRI) provided the computational resources. This
496 study is jointly supported by the Ministry of Science and Technology of Taiwan under Grants MOST-103-2625-M-
497 003-001-MY2, MOST-105-2111-M-003-003-MY3, MOST-108-2111-M-003-005-MY2, MOST-110-2111-M-003-
498 004, and MOST-110-2625-M-003-001.

499 **References**

- 500 Barnes, L. R., Schultz, D. M., Grunfest, E. C., Hayden, M. H., and Benight, C. C.: Corrigendum: False alarm rate or
501 false alarm ratio?, *Weather Forecast.*, 24, 1452–1454, <https://doi.org/10.1175/2009WAF2222300.1>, 2009.
- 502 Bryan, G. H., Wyngaard, J. C., and Fritsch, J. M.: Resolution requirements for the simulation of deep moist convection,
503 *Mon. Weather Rev.*, 131, 2394–2416, [https://doi.org/10.1175/1520-0493\(2003\)131%3C2394:RRFTSO%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131%3C2394:RRFTSO%3E2.0.CO;2),
504 2003.
- 505 Chang, C.-P., Yeh, T.-C., and Chen, J.-M.: Effects of terrain on the surface structure of typhoons over Taiwan, *Mon.*
506 *Weather Rev.*, 121, 734–752, [https://doi.org/10.1175/1520-0493\(1993\)121%3C0734:EOTOTS%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121%3C0734:EOTOTS%3E2.0.CO;2), 1993.
- 507 Chang, C.-P., Yang, Y.-T., and Kuo, H.-C.: Large increasing trend of tropical cyclone rainfall in Taiwan and the roles
508 of terrain, *J. Clim.*, 26, 4138–4147, <https://doi.org/10.1175/JCLI-D-12-00463.1>, 2013.
- 509 Chen, C.-S., and Chen, Y.-L.: The rainfall characteristics of Taiwan. *Mon. Weather Rev.*, 131, 1324–1341, 2003.
- 510 Chen, G. T.-J., Wang, C.-C., and Lin, D. T.-W.: Characteristics of low-level jets over northern Taiwan in Mei-yu
511 season and their relationship to heavy rain events, *Mon. Weather Rev.*, 133, 20–43, [https://doi.org/10.1175/MWR-](https://doi.org/10.1175/MWR-2813.1)
512 2813.1, 2005.
- 513 Chen, T.-C., Yen, M.-C., Hsieh, J.-C., and Arritt, R. W.: Diurnal and seasonal variations of the rainfall measured by
514 the Automatic Rainfall and Meteorological Telemetry System in Taiwan. *Bull. Amer. Meteor. Soc.*, 80, 2299–2312,
515 doi:10.1175/1520-0477(1999)080,2299: DASVOT.2.0.CO;2, 1999.
- 516 Cheung, K. K. W., Huang, L.-R., and Lee, C.-S.: Characteristics of rainfall during tropical cyclone periods in Taiwan,
517 *Nat. Hazards Earth Syst. Sci.*, 8, 1463–1474, <https://doi.org/10.5194/nhess-8-1463-2008>, 2008.
- 518 Chi, S.-S.: The Mei-Yu in Taiwan, SFRDEST E-06-MT-03-4, Chung-Shin Engineering Technology Research and
519 Development Foundation, Taipei, Taiwan, 65 pp, 2006. (in Chinese)
- 520 Chien, F.-C., and Jou, B. J.-D.: MM5 ensemble mean precipitation in the Taiwan area for three early summer
521 convective (Mei-Yu) seasons, *Weather Forecast.*, 19, 735–750, 2004.
- 522 Chien, F.-C., Kuo, Y.-H., and Yang, M.-J.: Precipitation forecast of MM5 in the Taiwan area during the 1998 Mei-yu
523 season, *Weather Forecast.*, 17, 739–754, [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0434(2002)017%3C0739:PFOMIT%3E2.0.CO;2)
524 0434(2002)017%3C0739:PFOMIT%3E2.0.CO;2, 2002.

525 Chien, F.-C., Liu, Y.-C., and Jou, B. J.-D.: MM5 ensemble mean forecasts in the Taiwan area for the 2003 Mei-yu
526 season, *Weather Forecast.*, 21, 1006–1023, <https://doi.org/10.1175/WAF960.1>, 2006.

527 Clark, A. J., Gallus, Jr., W. A., and Chen, T.-C.: Comparison of the diurnal precipitation cycle in convection-resolving
528 and non-convection-resolving mesoscale models, *Mon. Weather Rev.*, 135, 3456–3473,
529 <https://doi.org/10.1175/MWR3467.1>, 2007.

530 Clark, A. J., Kain, J. S., Stensrud, D. J., Xue, M., Kong, F., Coniglio, M. C., Thomas, K. W., Wang, Y., Brewster, K.,
531 Gao, J., Wang, X., Weiss, S. J., and Du, J.: Probabilistic precipitation forecast skill as a function of ensemble size and
532 spatial scale in a convection-allowing ensemble, *Mon. Weather Rev.*, 139, 1410–1418,
533 <https://doi.org/10.1175/2010MWR3624.1>, 2011.

534 Cotton, W. R., Tripoli, G. J., Rauber, R. M., and Mulvihill, E. A.: Numerical simulation of the effects of varying ice
535 crystal nucleation rates and aggregation processes on orographic snowfall, *J. Appl. Meteorol. Climatol.*, 25, 1658–
536 1680, [https://doi.org/10.1175/1520-0450\(1986\)025%3C1658:NSOTEO%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025%3C1658:NSOTEO%3E2.0.CO;2), 1986.

537 Cuo, L., Pagano, T. C., and Wang, Q. J.: A review of quantitative precipitation forecasts and their use in short- to
538 medium-range streamflow forecasting, *J. Hydrometeorol.*, 12, 713–728, <https://doi.org/10.1175/2011JHM1347.1>,
539 2011.

540 Davis, C., Brown, B., and Bullock, R.: Object-based verification of precipitation forecasts. Part I: Methodology and
541 application to mesoscale rain areas, *Mon. Weather Rev.*, 134, 1772–1784, <https://doi.org/10.1175/MWR3145.1>, 2006.

542 Deardorff, J. W.: Stratocumulus-capped mixed layers derived from a three-dimensional model, *Bound.-Layer*
543 *Meteorol.*, 18, 495–527, 1980.

544 Done, J., Davis, C. A., and Weisman, M.: The next generation of NWP: explicit forecasts of convection using the
545 weather research and forecasting (WRF) model, *Atmos. Sci. Lett.*, 5, 110–117, 2004.

546 Ebert, E. E.: Ability of a poor man’s ensemble to predict the probability and distribution of precipitation, *Mon.*
547 *Weather Rev.*, 129, 2461–2480, [https://doi.org/10.1175/1520-0493\(2001\)129%3C2461:AOAPMS%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129%3C2461:AOAPMS%3E2.0.CO;2),
548 2001.

549 Ebert, E. E., and McBride, J. L.: Verification of precipitation in weather systems: Determination of systematic errors,
550 *J. Hydrol.*, 239, 179–202, [https://doi.org/10.1016/S0022-1694\(00\)00343-7](https://doi.org/10.1016/S0022-1694(00)00343-7), 2000.

551 Ebert, E. E., Damrath, U., Wergen, W., and Baldwin, M. E.: The WGNE assessment of short-term quantitative
552 precipitation forecasts (QPFs) from operational numerical weather prediction models, *Bull. Am. Meteorol. Soc.*, 84,
553 481–492, <https://doi.org/10.1175/BAMS-84-4-481>, 2003.

554 Fang, X., and Kuo, Y.-H.: Improving ensemble-based quantitative precipitation forecasts for topography-enhanced
555 typhoon heavy rainfall over Taiwan with a modified probability-matching technique, *Mon. Weather Rev.*, 141, 3908–
556 3932, <https://doi.org/10.1175/MWR-D-13-00012.1>, 2013.

557 Fritsch, J. M., and Carbone, R. E.: Improving quantitative precipitation forecasts in the warm season. A USWRP
558 research and development strategy, *Bull. Am. Meteorol. Soc.*, 85, 955–965, <https://doi.org/10.1175/BAMS-85-7-955>,
559 2004.

560 Gilleland, E., Ahijevych, D. A., Brown, B. G., and Ebert, E. E.: Verifying forecasts spatially, *Bull. Am. Meteorol.*
561 *Soc.*, 91, 1365–1373, <https://doi.org/10.1175/2010BAMS2819.1>, 2010.

562 Golding, B. W.: Quantitative precipitation forecasting in the UK. *J. Hydrol.*, 239, 286–305,
563 [https://doi.org/10.1016/S0022-1694\(00\)00354-1](https://doi.org/10.1016/S0022-1694(00)00354-1), 2000.

564 Hochman, A., Scher, S., Quinting, J., Pinto, J. G., and Messori, G.: A new view of heat wave dynamics and
565 predictability over the eastern Mediterranean, *Earth Syst. Dynam.*, 12, 133–149, [https://doi.org/10.5194/esd-12-133-](https://doi.org/10.5194/esd-12-133-2021)
566 2021, 2021.

567 Hong, J.-S.: Evaluation of the high-resolution model forecasts over the Taiwan area during GIMEX, *Weather and*
568 *Forecast.*, 18, 836–846, [https://doi.org/10.1175/1520-0434\(2003\)018%3C0836:EOTHMF%3E2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018%3C0836:EOTHMF%3E2.0.CO;2), 2003.

569 Hong, J.-S., Fong, C.-T., Hsiao, L.-F., Yu, Y.-C., and Tseng, C.-Y.: Ensemble typhoon quantitative precipitation
570 forecasts model in Taiwan, *Weather Forecast.*, 30, 217–237, <https://doi.org/10.1175/WAF-D-14-00037.1>, 2015.

571 Hsu, J.: ARMTS up and running in Taiwan, *Väisälä News*, 146, 24–26, 1998.

572 Hsu, J. C.-S., Wang, C.-J., Chen, P.-Y., Chang, T.-H., and Fong, C.-T. (2014), Verification of quantitative
573 precipitation forecasts by the CWB WRF and ECMWF on 0.125° grid, in: *Proceedings of 2014 Conference on*
574 *Weather Analysis and Forecasting*, Central Weather Bureau, Taipei, Taiwan, 16-18 September 2014, A2-24, 2014.
575 (in Chinese)

576 Huang, T.-S., Yeh, S.-H., Leu, G.-C., and Hong, J.-S.: A synthesis and comparison of QPF verifications at the CWB
577 and major NWP guidance, in: *Proceedings of 2015 Conference on Weather Analysis and Forecasting*, Central Weather
578 Bureau, Taipei, Taiwan, 15-17 September 2015, A7-11, 2015. (in Chinese)

579 Huang, T.-S., Yeh, S.-H., Leu, G.-C., and Hong, J.-S.: Postprocessing of ensemble rainfall forecasts---Ensemble mean,
580 probability matched mean and exceeding probability, *Atmospheric Sciences*, 44, 173–196, 2016. (in Chinese with
581 English abstract)

582 Ikawa, M., and Saito, K: Description of a non-hydrostatic model developed at the Forecast Research Department of
583 the MRI, Technical Report, 28, Meteorological Research Institute, Tsukuba, Ibaraki, Japan, 245 pp, 1991.

584 Jou, B. J.-D., Lee, W.-C., and Johnson, R. H.: An overview of SoWMEX/TiMREX, in: *The Global Monsoon System:*
585 *Research and Forecast*, 2nd Edition, edited by: Chang, C.-P., Ding, Y., Lau, N.-C., Johnson, R. H., Wang, B., Yasunari,
586 T., World Scientific, Toh Tuck Link, Singapore, 303–318, https://doi.org/10.1142/9789814343411_0018, 2011.

587 Kalnay, E., Kanamitsu, M., and Baker, W. E.: Global numerical weather prediction at the National Meteorological
588 Center, *Bull. Am. Meteorol. Soc.*, 71, 1410–1428, 1990.

589 Kanamitsu, M.: Description of the NMC global data assimilation and forecast system, *Weather Forecast.*, 4, 335–342,
590 [https://doi.org/10.1175/1520-0434\(1989\)004%3C0335:DOTNGD%3E2.0.CO;2](https://doi.org/10.1175/1520-0434(1989)004%3C0335:DOTNGD%3E2.0.CO;2), 1989.

591 Kleist, D. T., Parrish, D. F., Derber, J. C., Treadon, R., Wu, W. S., and Lord, S.: Introduction of the GSI into the
592 NCEP global data assimilation system, *Weather Forecast.*, 24, 1691–1705,
593 <https://doi.org/10.1175/2009WAF2222201.1>, 2009.

594 Kondo, J.: Heat balance of the China Sea during the air mass transformation experiment, *J. Meteor. Soc. Japan*, 54,
595 382–398, https://doi.org/10.2151/jmsj1965.54.6_382, 1976.

596 Kuo, Y.-H., and Chen, G. T.-J.: The Taiwan Area Mesoscale Experiment (TAMEX): An overview, *Bull. Am.*
597 *Meteorol. Soc.*, 71, 488–503, [http://dx.doi.org/10.1175/1520-0477\(1990\)071%3C0488:TTAMEA%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1990)071%3C0488:TTAMEA%3E2.0.CO;2),
598 1990.

599 Li, C.-H., and Hong, J.-S.: Study on the application and analysis of regional ensemble quantitative precipitation
600 forecasts, in: Proceedings of 2014 Conference on Weather Analysis and Forecasting, Central Weather Bureau, Taipei,
601 Taiwan, 16-18 September 2014, A2-19, 2014. (in Chinese)

602 Li, J. and Chen, Y.-L.: Barrier jets during TAMEX, *Mon. Weather Rev.*, 126, 959–971, [https://doi.org/10.1175/1520-0493\(1998\)126%3C0959:BJDT%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126%3C0959:BJDT%3E2.0.CO;2), 1998.

604 Lin, Y.-L., Farley, R. D., and Orville, H. D.: Bulk parameterization of the snow field in a cloud model, *J. Appl. Meteorol. Climatol.*, 22, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022%3C1065:BPOTSF%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022%3C1065:BPOTSF%3E2.0.CO;2),
605 1983.

607 Louis, J. F., Tiedtke, M., and Geleyn, J. F.: A short history of the operational PBL parameterization at ECMWF, in:
608 Proceedings of Workshop on Planetary Boundary Layer Parameterization, Shinfield Park, Reading, UK, 25-27
609 November 1981, 59–79, 1982.

610 Moorthi, S., Pan, H. L., and Caplan, P.: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast
611 system, NWS Technical Procedures Bulletin, 484, Office of Meteorology, National Weather Service, Silver Spring,
612 Maryland, USA, 2001.

613 Murakami, M.: Numerical modeling of dynamical and microphysical evolution of an isolated convective cloud---The
614 19 July 1981 CCOPE cloud, *J. Meteor. Soc. Japan*, 68, 107–128, https://doi.org/10.2151/jmsj1965.68.2_107, 1990.

615 Murakami, M., Clark, T. L., and Hall, W. D.: Numerical simulations of convective snow clouds over the Sea of Japan:
616 Two-dimensional simulation of mixed layer development and convective snow cloud formation, *J. Meteor. Soc. Japan*,
617 72, 43–62, https://doi.org/10.2151/jmsj1965.72.1_43, 1994.

618 Paul, S., Wang, C.-C., Chien, F.-C., and Lee, D.-I.: An evaluation of the WRF Mei-yu rainfall forecasts in Taiwan,
619 2008-2010: differences in elevation and sub-regions. *Meteorol. Appl.*, 25, 269-282, doi: 10.1002/met.1689, 2018.

620 Roberts, N. M., and Lean, H. W.: Scale-selective verification of rainfall accumulations from high-resolution forecasts
621 of convective events, *Mon. Weather Rev.*, 136, 78–97, <https://doi.org/10.1175/2007MWR2123.1>, 2007.

622 Roebber, P. J.: Visualizing multiple measures of forecast quality, *Weather Forecast.*, 24, 601–608,
623 <https://doi.org/10.1175/2008WAF2222159.1>, 2009.

624 Schaefer, J. T.: The critical success index as an indicator of warning skill, *Weather Forecast.*, 5, 570–575,
625 [https://doi.org/10.1175/1520-0434\(1990\)005%3C0570:TCSIAA%3E2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005%3C0570:TCSIAA%3E2.0.CO;2), 1990.

626 Segami, A., Kurihara, K., Nakamura, H., Ueno, M., Takano, I., and Tatsumi, Y.: Operational mesoscale weather
627 prediction with Japan Spectral Model, *J. Meteor. Soc. Japan*, 67, 907–924, https://doi.org/10.2151/jmsj1965.67.5_907,
628 1989.

629 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description
630 of the advanced research WRF version 2, National Center for Atmospheric Research, Boulder, Colorado, USA, 88
631 pp, <http://dx.doi.org/10.5065/D6DZ069T>, 2005.

632 Su, Y.-J., Hong, J.-S., and Li, C.-H.: The characteristics of the probability matched mean QPF for 2014 Meiyu season,
633 *Atmospheric Sciences*, 44, 113-134, 2016. (in Chinese with English abstract)

634 Tsuboki, K., and Sakakibara, A.: Large-scale parallel computing of cloud resolving storm simulator, in: High
635 Performance Computing, edited by: Zima H. P., Joe K., Sato M., Seo Y., Shimasaki M., Springer, Berlin, Heidelberg,
636 Germany, 243–259, https://doi.org/10.1007/3-540-47847-7_21, 2002.

637 Tsuboki, K., and Sakakibara, A.: Numerical Prediction of High-Impact Weather Systems: The Textbook for the
638 Seventeenth IHP Training Course in 2007, Hydrospheric Atmospheric Research Center, Nagoya University, Nagoya,
639 Japan, and UNESCO, Paris, France, 273 pp, 2007.

640 Walser, A., and Schär, C.: Convection-resolving precipitation forecasting and its predictability in Alpine river
641 catchments, *J. Hydrol.*, 288, 57–73, <https://doi.org/10.1016/j.jhydrol.2003.11.035>, 2004.

642 Wang, C.-C.: On the calculation and correction of equitable threat score for model quantitative precipitation forecasts
643 for small verification areas: The example of Taiwan, *Weather Forecast.*, 29, 788–798, [https://doi.org/10.1175/WAF-](https://doi.org/10.1175/WAF-D-13-00087.1)
644 [D-13-00087.1](https://doi.org/10.1175/WAF-D-13-00087.1), 2014.

645 Wang, C.-C.: The more rain, the better the model performs—The dependency of quantitative precipitation forecast
646 skill on rainfall amount for typhoons in Taiwan, *Mon. Weather Rev.*, 143, 1723–1748, [https://doi.org/10.1175/MWR-](https://doi.org/10.1175/MWR-D-14-00137.1)
647 [D-14-00137.1](https://doi.org/10.1175/MWR-D-14-00137.1), 2015.

648 Wang, C.-C.: News and notes, Paper of notes: The more rain from typhoons, the better the models perform, *Bull. Am.*
649 *Meteorol. Soc.*, 97, 16–17, https://doi.org/10.1175/BAMS_971_11-18_Nowcast, 2016.

650 Wang, C.-C., Chen, G. T.-J., Chen, T.-C., and Tsuboki, K.: A numerical study on the effects of Taiwan topography
651 on a convective line during the mei-yu season, *Mon. Weather Rev.*, 133, 3217–3242,
652 <https://doi.org/10.1175/MWR3028.1>, 2005.

653 Wang, C.-C., Chen, G. T.-J., and Huang, S.-Y.: Remote trigger of deep convection by cold outflow over the Taiwan
654 Strait in the Mei-yu season: A modeling study of the 8 June 2007 Case, *Mon. Weather Rev.*, 139, 2854–2875,
655 <https://doi.org/10.1175/2011MWR3613.1>, 2011.

656 Wang, C.-C., Kung, C.-Y., Lee, C.-S., and Chen, G. T.-J.: Development and evaluation of Mei-yu season quantitative
657 precipitation forecast in Taiwan river basins based on a conceptual climatology model, *Weather Forecast.*, 27, 586–
658 607, <https://doi.org/10.1175/WAF-D-11-00098.1>, 2012a.

659 Wang, C.-C., Kuo, H.-C., Chen, Y.-H., Huang, H.-L., Chung, C.-H., and Tsuboki, K.: Effects of asymmetric latent
660 heating on typhoon movement crossing Taiwan: The case of Morakot (2009) with extreme rainfall, *J. Atmos. Sci.*, 69,
661 3172–3196, <https://doi.org/10.1175/JAS-D-11-0346.1>, 2012b.

662 Wang, C.-C., Chen, Y.-H., Kuo, H.-C., and Huang, S.-Y.: Sensitivity of typhoon track to asymmetric latent
663 heating/rainfall induced by Taiwan topography: A numerical study of Typhoon Fanapi (2010), *J. Geophys. Res.*
664 *Atmos.*, 118, 3292–3308, <https://doi.org/10.1002/jgrd.50351>, 2013a.

665 Wang, C.-C., Kuo, H.-C., Yeh, T.-C., Chung, C.-H., Chen, Y.-H., Huang, S.-Y., Wang, Y.-W., and Liu, C.-H.: High-
666 resolution quantitative precipitation forecasts and simulations by the Cloud-Resolving Storm Simulator (CRSS) for
667 Typhoon Morakot (2009), *J. Hydrol.*, 506, 26–41. <https://doi.org/10.1016/j.jhydrol.2013.02.018>, 2013b.

668 Wang, C.-C., Huang, S.-Y., Chen, S.-H., Chang, C.-S., and Tsuboki, K.: Cloud-resolving typhoon rainfall ensemble
669 forecasts for Taiwan with large domain and extended range through time-lagged approach, *Weather Forecast.*, 31,
670 151–172, <https://doi.org/10.1175/WAF-D-15-0045.1>, 2016a.

671 Wang, C.-C., Chiou, B.-K., Chen, G. T.-J., Kuo, H.-C., and Liu, C.-H.: A numerical study of back-building process
672 in a quasistationary rainband with extreme rainfall over northern Taiwan during 11-12 June 2012, *Atmos. Chem. Phys.*,
673 16, 12359–12382, <https://doi.org/10.5194/acp-16-12359-2016>, 2016b.

674 Wang, C.-C., Paul, S., Chien, F.-C., Lee, D.-I., and Chuang, P.-Y.: An evaluation of WRF rainfall forecasts in Taiwan
675 during three mei-yu seasons of 2008-2010, *Weather Forecast.*, 32, 1329–1351, [https://doi.org/10.1175/WAF-D-16-](https://doi.org/10.1175/WAF-D-16-0190.1)
676 0190.1, 2017.

677 Wernli, H., Paulat, M., Hagen, M., and Frei, C.: SAL—A novel quality measure for the verification of quantitative
678 precipitation forecasts, *Mon. Weather Rev.*, 136, 4470–4487, <https://doi.org/10.1175/2008MWR2415.1>, 2008.

679 Wilks, D. S.: *Statistical methods in the atmospheric sciences*, 3rd Edition. Academic Press, San Diego, California,
680 USA, 2011.

681 Wu, C.-C., and Kuo, Y.-H.: Typhoons affecting Taiwan: Current understanding and future challenges, *Bull. Am.*
682 *Meteorol. Soc.*, 80, 67–80, [https://doi.org/10.1175/1520-0477\(1999\)080%3C0067:TATCUA%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080%3C0067:TATCUA%3E2.0.CO;2), 1999.

683 Yang, M.-J., Jou, B. J.-D., Wang, S.-C., Hong, J.-S., Lin, P.-L., Teng, J.-H., and Lin, H.-C.: Ensemble prediction of
684 rainfall during the 2000–2002 mei-yu seasons: Evaluation over the Taiwan area, *J. Geophys. Res.*, 109, D18203,
685 <https://doi.org/10.1029/2003JD004368>, 2004.

686 Yeh, H.-C., and Chen, Y.-L.: Characteristics of the rainfall distribution over Taiwan during TAMEX, *J. Appl.*
687 *Meteorol. Climatol.*, 37, 1457–1469, [https://doi.org/10.1175/1520-0450\(1998\)037%3C1457:CORDOT%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1998)037%3C1457:CORDOT%3E2.0.CO;2),
688 1998.

689 Yeh, H.-C., and Chen, Y.-L.: The role of off shore convergence on coastal rainfall during TAMEX IOP 3, *Mon.*
690 *Weather Rev.*, 130, 2709–2730, [https://doi.org/10.1175/1520-0493\(2002\)130%3C2709:TROOCO%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130%3C2709:TROOCO%3E2.0.CO;2),
691 2002.