
Chung-Chieh Wang 1, Shin-Hau Chen 1,*, Kazuhisa Tsuboki 2, Shin-Yi Huang 1 and Chih-Sheng Chang 1

1 Department of Earth Sciences, National Taiwan Normal University, Taipei 11677, Taiwan; cwang@ntnu.edu.tw (C.-C.W.); hshinyi0828@gmail.com (S.-Y.H.); shurin.chang@gmail.com (C.-S.C.)
2 Institute for Space-Earth Environmental Research, Nagoya University, Nagoya 464-8601, Japan; tsuboki@nagoya-u.jp
* Correspondence: b96209006@ntu.edu.tw

Abstract: Typhoon Morakot (2009) struck Taiwan during 7–9 August and brought extreme rainfall to the southern mountainous regions of the island, with a peak of 2635 mm in 48 h (and most heavy rainfall observed on 8 August), subsequently causing widespread damage. In this study, the time-lagged ensemble method is applied to this unique case to examine its effectiveness and usefulness for the prediction of heavy rainfall events. A series of hindcasts experiments was carried out using a cloud-resolving model with a grid size of 2.5 km at 6-h intervals from 0000 UTC 1 August to 0000 UTC 9 August 2009, with real-time global model products as initial and boundary conditions. It was found that most of the hindcast members initialized at and after 0600 UTC 6 August in the short range (i.e., within 72 h) could capture the magnitude of the rainfall to a reasonable extent, with a peak 48-h amount near or over 2500 mm, under the condition that the track errors were adequately small. With longer lead times, however, the predictability was limited due to larger track errors, and reliable rainfall predictions were impossible for runs with an initial time before or on 5 August. Thus, the probabilities derived from lagged members for extreme rainfall increased dramatically starting from 6 August, when a >80–90% likelihood was indicated that the southern mountainous regions of Taiwan would receive a 48-h rainfall in excess of 1000 mm, and ≥80% in parts of the area to receive over 1500 mm. Therefore, the time evolution of the probability may be useful in decision making and hazard mitigation. The limitations of such a time-lagged system and the potential to make further improvements are also discussed.

Keywords: Typhoon Morakot (2009); Quantitative Precipitation Forecast (QPF); tropical cyclone; time-lagged ensemble; cloud-resolving model; Taiwan

1. Introduction

Quantitative Precipitation Forecasts (QPFs) of heavy-rainfall thresholds are a challenging area in numerical weather predictions [1–4], and are in particular demand when associated with high-impact and hazardous weather events such as tropical cyclones (TCs). Situated over the western North Pacific (WNP), roughly 3–4 TCs (or typhoons) strike Taiwan each year on average [5]; they can bring extreme rainfall to the island, leading to hazards like flash floods, urban inundation, landslides, and debris flows. Such hazards are typical because Taiwan has steep and complex topography (Figure 1a) which enhances rainfall production under the influence of TCs [6–9] and, at times, in combination with the southwesterly monsoon flow [10–13]. Therefore, while other aspects of TC predictions such as track and intensity are not to be overlooked [14–18], the performance and accuracy of model QPFs for TCs are especially important in Taiwan, as well as in many other regions around the world with similar conditions [19–21].
Due to the chaotic behavior of nonlinear systems like the atmosphere [22,23], it is well known that NWP is not an exact science, and each model forecast has errors and therefore uncertainty. To gauge forecast errors and better cover actual scenarios, nowadays, it is common practice at many operational centers to run ensemble forecasts using multiple members (and sometimes even a large number) simultaneously at each initial time [24–28]. However, for agencies that are not as resourceful, such as those in many developing countries (which are also often less able to cope with natural hazards), there is an inevitable tradeoff between ensemble (i.e., probability) information and higher model resolution. Moreover, different parameters require different grid resolutions to be effectively captured. Among those related to TCs, for example, their tracks are mainly controlled by large-scale environmental flow, and can thus often be predicted reasonably well with relatively coarse grids using global models [29,30]. On the other hand, the rapid intensification process and QPFs of TCs, linked to detailed cloud microphysics, tend to require high resolution, sometimes cloud-resolving grid sizes [31,32]. Indeed, real-time forecast experiments for a large number of typhoons in Taiwan over the past decade [33–36] have demonstrated that QPFs are improved dramatically using the Cloud-Resolving Storm Simulator [37,38] at a grid size of 2.5 km to better resolve both deep convection and island topography, compared to 5-km models. The same is true for QPFs during the Mei-yu regime (May–June) [39]. Therefore, when there is a tradeoff, it is important to identify the parameters which are responsible for local hazards and ensure an adequate resolution for them, e.g., [40–42]. Without the ability to properly simulate the key parameters, the ensemble system will not be able to provide the right estimates for probability, regardless of the total number of members used.

To reach high resolution as discussed above, time-lagged ensemble is a viable choice [43,44], as the method uses computational resources very efficiently. In Wang et al. [45], this strategy was applied to perform routine typhoon QPFs for Taiwan from 2012 with a range of eight days, with information that is particularly useful for hazard preparation (see also [46]). Using the example of Typhoon (TY) Kong-Rey (2013), which produced an atypical rainfall pattern for its track due to the presence of a southwesterly monsoon, catching many by surprise at the time [47], the study showed that the lagged ensemble provided a variety of rainfall scenarios associated with different storm tracks at longer
ranges (roughly at least 4–5 days ahead), precisely because the uncertainty was high at this stage [45]. Based on these scenarios, the authorities can start to prepare for the worst-case scenario, because however low the probability, there is no guarantee that such a scenario of heavy rainfall will not happen at this stage. Later, as the TC moved closer (to within 3 days), its tracks from successive predictions gradually converged toward the best track (not fully known at the time), and the probability of that particular scenario increased. Adjustments could then be made accordingly. Such a strategy fits the traditional wisdom of “hope for the best, but prepare for the worst.” For TY Kong-Rey, where the worst-case scenario turned out to occur, one can also see that the above strategy allowed for the longest preparation time in that situation, and therefore, provided the best chance to minimize the loss or impact from the event [45,46]. At the short range (≤72 h), the QPFs for several TCs in 2012 examined in [45] were comparable to the long-term results shown in [33,36].

In early August of 2009, TY Morakot hit Taiwan, causing the worst weather disaster in the nation in half a century, with rainfall peaking at 2855 mm over a period of four days [48–51] (Figure 1b). Fatalities were very high (699), mainly due to a massive mudslide that buried Shiao-Lin Village in southern CMR at around 2210 UTC on 8 August [48,50]. Due to a change in its environmental steering flow linked to a monsoon gyre at a quasi-biweekly (or 10–30 day) timescale, Morakot slowed down and changed its course from west to north-northwest as it passed over Taiwan [52–55] (Figure 2). Moreover, TY Morakot (2009) interacted strongly with the southwesterly flow and developed an asymmetric rainfall structure, with strongest convection near/over Taiwan’s terrain and to the rear of its motion. This localized latent heat release caused a further slowdown during its departure on 8 August [50,56] (cf. Figure 2), which was the period of heaviest rainfall in the mountains of southern Taiwan. Largely due to the above reasons, as well as other factors examined in some past studies, including the topography of Taiwan, e.g., [53,57–60] and the storm-scale interaction between updrafts inside the rainband and its TC environment [61], extreme rainfall occurred.

![Figure 2. Central Weather Bureau (CWB) best track (every 6 h) of TY Morakot (2009), overlaid with a schematic showing the approximate location and propagation of the background cyclonic monsoon gyre (at frequencies of 10–30 day) from 5 to 8 August (dashed gray) based on Hong et al. [52]. Different colors represent typhoon intensities of tropical depression (TD, <17.2 m s$^{-1}$, purple), tropical storm or mild typhoon (TS, 17.2–32.6 m s$^{-1}$, blue), moderate typhoon (MTY, 32.7–50.9 m s$^{-1}$, green), and severe typhoon (STY, ≥51.0 m s$^{-1}$, red). The month and date at 0000 UTC are labeled in “mm/dd” format. The original plot was produced and provided by the CWB.](image-url)
Due to the complex interactions of features or processes across a wide range of scales as described above, real-time forecasting for TY Morakot (2009) was very challenging, as a good track and QPF depended on the ability of the models to properly capture such interactions. For example, Hendricks et al. [62] shows that an ideal QPF for this event could not be obtained without a good TC track in the forecast. Wang et al. [51] suggested that a quality deterministic QPF could be made on 6 August, i.e., roughly 48 h prior to the onset of the heaviest rainfall in southern Taiwan, while simulations using analyses as initial and boundary conditions (IC/BCs) could capture the event reasonably well (including mid- and upper-level features) at an initial time \( t_0 \) as early as 3–4 August [63]. Chou et al. [64] found that Morakot (2009) was among the most difficult TCs, in the sense that dropsonde data acquired on 6 August did not improve its track forecast. As the track forecast had higher uncertainty before 6 August, an ensemble approach was recommended in several studies [42,65,66]. However, these studies performed ensemble forecasts using a large number of members simultaneously at only one \( t_0 \), most likely due to the heavy demand on computational resources, as mentioned. Even at short range (within 3 days), most forecast studies showed deterministic results with a \( t_0 \) at or after 0000 UTC 6 August.

In this study, therefore, the method of time-lagged ensemble recommended by [45] was applied to the unique case of TY Morakot (2009), as such a strategy has the benefits of both deterministic and ensemble approaches, with cloud-resolving capability and the capacity to provide probability information. Because of its high efficiency, the method is more affordable and thus more feasible than the multi-member approach. With high resolution, realistic QPFs for Taiwan could be produced not only at short range, but potentially also at longer ranges, i.e., up to about a week. To find out whether this is attainable is our first focus of the present study. Also, in the present work, as our second focus, the time evolution of probability information derived from the lagged ensemble is emphasized, as such an aspect was not explored for Morakot previously. Compared to [51], the present study also includes model runs at 0600 and 1800 UTC and at longer lead times. To understand what can be achieved in real time, only data or information available at each \( t_0 \) are used.

The remainder of this paper is arranged as follows. In Section 2, the CReSS model, hindcast experiment, and the methodology for QPF verification are described. In Section 3, our results of the time-lagged CReSS ensemble for TY Morakot (2009) are presented with a discussion, including the tracks, QPFs, and probability information. Finally, the conclusion and summary are given in Section 4.

2. The CReSS Model, Hindcast Experiments, and QPF Verification

2.1. The CReSS Model and Hindcast Experiments

In this study, the same version and CReSS configuration as used in [45,46] are applied, so only a brief description is given here, with the basic setup and major physical packages listed in Table 1. The CReSS model [37,38] is a cloud-resolving model that employs a terrain-following vertical coordinate. All clouds in CReSS are simulated explicitly using a bulk cold-rain scheme, with a total of six species: vapor, cloud water, cloud ice, rain, snow, and graupel [67–71], without any cumulus parameterization. Sub-grid scale parameterization includes turbulence in the planetary boundary layer (PBL) [37,72,73] and surface radiation and momentum/energy fluxes with a substrate model [73–75]. For further details, readers are referred to [45] and [37,38].

To assess the practical predictability [65] of TY Morakot (2009) in our hindcasts, the gross analysis and forecasts in real time by the Global Forecast System (GFS) of the National Center for Environmental Prediction (NCEP), which are freely available, are used as IC/BCs for all our runs, as in [33,36,45]. Thus, the results presented here are reforecasts using only data available before each run in 2009. Note that the track errors from the global forecasts (especially during the exit period) tend to be large prior to about 0000 UTC 6 August 2009 [51]. A few studies since then have shown that the use of a more advanced global
model or a more sophisticated DA system [76], perhaps with a multi-member ensemble approach, is likely to improve the results, e.g., [42, 65, 66]. A different data source such as the analyses from the European Center for Medium-range Weather Forecasts (ECMWF) might give somewhat different results as well.

Table 1. The basic domain configuration and physical packages of the 2.5-km CReSS in this study.

<table>
<thead>
<tr>
<th>Map Projection</th>
<th>Lambert Conformal (Center at 120° E, Secant at 10° N and 40° N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid spacing (km)</td>
<td>2.5 × 2.5 × 0.2 − 0.663 (0.5) *</td>
</tr>
<tr>
<td>Grid dimension (x, y, z)</td>
<td>744 × 544 × 40</td>
</tr>
<tr>
<td>Domain size (km)</td>
<td>1860 × 1360 × 20</td>
</tr>
<tr>
<td>Forecast frequency</td>
<td>Every 6 h, from 0000 UTC 1 to 0000 UTC 9 August 2009</td>
</tr>
<tr>
<td>Forecast length</td>
<td>8 days (192 h)</td>
</tr>
<tr>
<td>IC/BCs</td>
<td>NCEP GFS 1.0° × 1.0° analyses and forecasts (26 levels)</td>
</tr>
<tr>
<td>Cloud microphysics</td>
<td>Bulk cold-rain scheme</td>
</tr>
<tr>
<td>PBL parameterization</td>
<td>1.5-order closure with turbulent kinetic energy prediction</td>
</tr>
<tr>
<td>Surface processes</td>
<td>Energy/momentum fluxes and shortwave/longwave radiation</td>
</tr>
<tr>
<td>Soil model</td>
<td>41 levels, every 5 cm to 2 m in depth</td>
</tr>
</tbody>
</table>

* The vertical grid spacing of CReSS is uneven and stretched (smallest at bottom); the parentheses give the averaged value.

2.2. Verification of Model QPFs

For easy comparison with earlier results [33, 36, 45, 51], the widely-used categorical measures based on the 2 × 2 contingency table, e.g., [77, 78] are adopted here, in particular the threat score (TS) and bias score (BS). At any given point among a set of N verification points over any accumulation period, the outcome of a model prediction to reach a specified rainfall threshold (i.e., an event) can be one of the following four: hit (H, event observed and predicted), miss (M, event observed but not predicted), false alarm (FA, event not observed but predicted), and correct negative (CN, event neither observed nor predicted). Thus, \( N = H + M + FA + CN \). Through counting the number of points in each category, TS and BS are:

\[
TS = \frac{H}{H + M + FA}, \\
BS = \frac{H + FA}{H + M}.
\]

Thus, TS is the fraction of points with correct predictions of events among all event points in either observations or predictions (or both). Hence, 0 ≤ TS ≤ 1, and the higher the better. As one can see from Equation (1), “hits” \( H \) are required for TS to exceed zero, and the model’s ability to produce rainfall at the correct location is rewarded, a desirable property owning to the types of hazards in Taiwan [79], especially at high rainfall thresholds. BS, on the other hand, is the ratio of event points in the model prediction \( H + FA \) to those observed \( H + M \), and thus measures over- or under-prediction by the model (if BS > 1 or <1) regardless of whether hits are produced or not. The most ideal BS value is therefore unity. In this study, mainly the 48-h QPFs are verified due to the prolonged heavy rainfall from TY Morakot, at thresholds from 10 to the extreme value of 1500 mm (per 48 h). In addition to the categorical measurements of TS and BS, the Fractions Skill Score (FSS) [80] gauges the overall similarity between two patterns, in this case, the predicted and observed rainfall patterns in Taiwan, as:

\[
FSS = 1 - \frac{1}{N} \frac{\sum_{i=1}^{N} (f_i - o_i)^2}{\left(\sum_{i=1}^{N} f_i^2 + \sum_{i=1}^{N} o_i^2\right)},
\]

where \( f_i \) and \( o_i \) are the forecast and observed rainfall at the \( i \)th point among \( N \), respectively. Thus, 0 ≤ FSS ≤ 1 and a higher score means a smaller overall difference between observa-
tions and predictions. As shown in Equation (3), all \( N \) points are involved in the calculation, and correct predictions of no rainfall also contribute to a better FSS, in contrast to no effect for \( CN \) in the calculation of both TS and BS. As in [33,36,45,79], all scores are computed at the rain-gauge sites where accurate observations were made (Figure 1a), through the use of bilinear interpolation of model QPFs onto these locations. The total number of gauge sites during TY Morakot were around 400.


3.1. Tracks and the Associated QPFs

Using the data and methodology described above, hindcasts (or reforecasts) are performed every 6 h from 0000 UTC 1 to 0000 UTC 9 August 2009 (49 runs), and the tracks produced prior to 0000 UTC 8 August are presented in Figure 3. Consistent with earlier studies, those made during 1–5 August contain larger track errors, and none of the storm centers makes landfall in Taiwan (Figure 3a), although the one with a \( t_0 \) at 1800 UTC 5 August comes quite close. On the other hand, all TCs predicted during 6–7 August make landfall (Figure 3b), even though those on 6 August still tend to move slightly too far west or north during 8–9 August, i.e., the slow translation speed of the storm upon its exit is not fully captured.

![Figure 3](image-url) Figure 3. The best track (black) of TY Morakot (2009) from the Joint Typhoon Warning Center (JTWC) and predicted tracks produced from the time-lagged ensemble every 6 h during (a) 0000 UTC 1 to 1800 UTC 5 August and (b) 0000 UTC 6 to 1800 UTC 7 August, respectively. Different colors and symbols (at 0000 UTC of each day) are used for the tracks (see legend).

The accumulated rainfall patterns produced by the time-lagged ensemble and targeted for the 48 h from 1200 UTC 7 to 1200 UTC 9 August 2009, i.e., the period with the heaviest rainfall in southern CMR [50,51], are shown in Figure 4. With one member added every 6 h, these hindcast results can be compared with the observed distribution, where the peak amount is 2635 mm (in 48 h), containing the bulk of the total rainfall from the entire event. Since the TC track is poorly-predicted before 6 August, only the 18 members starting since 0600 UTC 3 August are presented. As expected, not much rainfall occurs in the hindcasts (for the targeted 48-h period) before the run at 1200 UTC 5 August, where the predicted rainfall in southern Taiwan starts to increase (Figure 4). Starting from the hindcast at 0600 UTC 6 August, most of the lagged members can capture the magnitude of the extreme rainfall in southern CMR, with a peak 48-h amount near or over 2500 mm (Figure 4). Thus, the event of TY Morakot (2009) in Taiwan can be predicted beforehand, provided that the lead time is short enough and the track error under the interactions across the scales can be reduced to a reasonable degree, as discussed in Section 1 and consistent with previous studies [51]. However, as also confirmed in Figures 3 and 4, a rainy scenario cannot be obtained prior to 1200 UTC 5 August for TY Morakot (2009), and the system cannot provide a worst-case scenario like in the case of Kong-Rey (2013) and several other TCs [45] due to the larger track errors which occur at this earlier stage.
3.2. Skill Scores of 48-h QPFs

The quality of the QPFs by the 2.5-km CReSS in Figure 4 can be quantified through categorical statistics (Section 2.2). The TS values at a series of thresholds from 10 to 1500 mm (per 48 h) are presented in Figure 5. As shown in Equation (1), TS is the fraction of points (i.e., rain gauge sites, cf. Figure 1a) that meet a specified rainfall threshold in both the observation and prediction (i.e., hits or $H$) among all points that are either observed or predicted, or both (i.e., $H + M + FA$). As shown, the QPFs (valid for 1200 UTC 7 to 1200 UTC 9 August) made on 4–5 August have low TS values in association with large track errors (Figure 3a), except at thresholds below about 100–200 mm (Figure 5a,b). Among them, the hindcast made from 1200 UTC 5 August is slightly better, with TSs above zero up to 800 mm (per 48 h). The above situation changes on 6 August, when both the 0600- and 1800-UTC runs produce much-improved QPFs (Figure 5c), with TS values near 0.5 at 800 mm and about 0.15 at 1500 mm, in agreement with Figure 4. The scores for the three runs on 7 August (up to 1200 UTC) are, in general, even slightly higher, especially the one made at 0000 UTC, with TS above 0.6 up to 900 mm and 0.47 at 1200 mm (Figure 5d). Note that at this time, the Shiao-Lin tragedy would still be about 46 h away.

Compared to the TS values obtained in previous studies, due to the 48-h accumulation period and thus a higher total amount [33,34,36], the TSs obtained here (with a real-time setup) are at least comparable to, if not better than, earlier results from simulations or forecasts at short range. For example, the run made at 0000 UTC 7 August produced a TS up to 0.52 at 1000 mm (for 48-h QPF at the range of 12–60 h), compared to 0.57 at the same threshold but for a 3-day total ($t_0 = 1200$ UTC 6 August) in the control simulation of [81], to 0.3 also at 1000 mm on day 3 (24-h QPF) in the 3-km CReSS simulation (their E6A experiment, $t_0 = 0000$ UTC 6 August) of [51], or to 0.4 on day 2 (24-h QPF) by the 4-km CReSS forecast ($t_0$ also at 0000 UTC 7 August) in [79]. On the other hand, at slightly longer range, the simulation would perform better mainly due to a smaller track error, as mentioned earlier. For instance, starting from 0000 UTC 6 August, the hindcast here can only reach a TS of 0.3 at thresholds ≤ 200 mm, compared to 1000 mm in E6A of [51] targeted at 24 h on 8 August. In our hindcast runs, as the TC still moves too fast, a significant portion...
Atmosphere 2022, 13, 585

of rainfall occurred before 1200 UTC 7 August and thus was outside the 48-h target period to help raise the TS values.

Figure 5. Threat scores (TSs) of 48-h QPFs over Taiwan, targeted for the 48-h period from 1200 UTC 7 to 1200 UTC 9 August during TY Morakot (2009), from six-hourly time-lagged ensemble hindcasts made during (a) 4 August, (b) 5 August, (c) 6 August, and (d) 7 August (to 1200 UTC only), respectively.

As expected, the BS values of the 48-h QPFs (targeted for the period from 1200 UTC 7 to 1200 UTC 9 August) made during 4–5 August show an increasing under-prediction from low to the high thresholds (Figure 6a,b). In most of these runs, the BSs are ≥0.7 at the lowest threshold of 10 mm, but decrease to below 0.1 beyond the thresholds of 200–500 mm. The run with a $t_0$ at 1200 UTC 4 August, interestingly, produced a peak rainfall amount reaching 1200 mm but at the wrong location (cf. Figure 4, first row, sixth panel), while the hindcast one day later (from 1200 UTC 5 August) also produced at least 800 mm (cf. Figure 4, second row, second panel). Thus, a few of these earlier members indeed provided some indication of the potential of extreme rainfall in Taiwan, although the location was not always correct. Again, the conditions improved from 0600 UTC 6 August (Figure 6c,d), when most members yielded BS values between 0.6 and 1.15 up to about 800 mm or higher. In terms of BS, the best member may be the one from 0000 UTC 7 August, which has slight over-prediction (1.0 ≤ BS ≤ 1.15) over the threshold range of 10–1000 mm (Figure 6d). Even at 1200 mm, the BS remains at about 0.9. This particular run also performs the best in TSs (cf. Figure 5d). Overall, for the extreme case of TY Morakot (2009), the over-predictions tend to be relatively mild (BS ≤ 1.15) in the hindcasts, even at short range, and under-prediction occurs more frequently than over-prediction (Figure 6).

While the TS and BS provide objective measurements in categorical statistics on the degree of overlapping between predicted and observed rain areas (TS) and their relative size (BS) at the selected thresholds, the FSS in Equation (3) gives the overall similarity between two rainfall patterns (observed and model) in one measure [80], as shown in Figure 7. Consistent with Figures 4 and 5, the FSS increases over the course of 5–7 August,
from about 0.15 to 0.85 and above. On 6–7 August, high-quality QPFs (say, with FSS reaching 0.8) are produced by the run at 0600 UTC and all those after 1800 UTC 6 August (Figure 7). Here, it is perhaps worth noting that in the two runs at 0000 and 1200 UTC 6 August, the model TCs still move too fast on 8 August, so the strongest rainfall occurs early in the southern CMR. Had the 48 h of 7–8 August been selected for verification instead, their results would be improved in terms of skill scores.

Figure 6. As in Figure 5, except for Bias scores (BSs) from the time-lagged ensemble hindcasts (a) 4 August 2009. (b) 5 August 2009. (c) 6 August 2009. (d) 7 August 2009.

Figure 7. Fractions skill score (FSS) of 48-h QPFs over Taiwan, targeted for the period from 1200 UTC 7 to 1200 UTC 9 August during TY Morakot (2009), from six-hourly time-lagged ensemble hindcasts made from 0000 UTC 1 until 1200 UTC 7 August (labelled along abscissa in the format of “ddhh” for date and hour in UTC), respectively.
3.3. Time Evolution of Probability

As reviewed and shown in [45], the probability information of QPFs can be derived from successive time-lagged members, and updated regularly in real time when new members become available. With one member added every 6 h, seven members (or runs) are used here for this purpose, and the results at five selected thresholds, up to the extreme value of 1500 mm, are presented in Figure 8. To emphasize the temporal evolution of heavy-rainfall probability, three sets are selected: the seven runs with \( t_0 \) from 0000 UTC 3 to 1200 UTC 4 August (top row); those with \( t_0 \) from 1200 UTC 4 to 0000 UTC 6 August (middle row); and those with \( t_0 \) from 0000 UTC 6 to 1200 UTC 7 August (bottom row), respectively. Also, while equal weights could be used [82–84], here a weight of \( k/\Sigma n \) is assigned to the \( k \)th member in time among a total of \( n \) members involved (\( n = 7 \)), while \( \Sigma n \) is the summation of 1 to \( n \) (i.e., \( \Sigma n = 28 \)). Thus, a weight of 1/28, 2/28, \( \ldots \), to 7/28 is applied, and more weight is placed on the later, newer members since their QPFs are more accurate on average.

Figure 8. Probabilities of 48-h QPFs (% color) from seven successive time-lagged members shown in Figure 7, with \( t_0 \) between (top row) 0000 UTC 3 and 1200 UTC 4, (middle row) 1200 UTC 4 and 0000 UTC 6, and (bottom row) 0000 UTC 6 and 1200 UTC 7 August (as labelled on left), to reach the thresholds of 200, 350, 500, 1000, and 1500 mm (left to right), respectively, for the target period of 1200 UTC 7 to 1200 UTC 9 August during TY Morakot (2009). The observed areas (over land) at each respective threshold are depicted by the thick black contours. For each group, the weights assigned to each member are 1/28, 2/28, \( \ldots \), 7/28, from the earliest to the latest, with a total weight of 1.

From Figure 8, one can easily see that the probability of heavy to extreme rainfall, up to 500–1500 mm in the southern CMR, increases dramatically for the runs after 0000 UTC 6 August, as the prediction increasingly anticipates a slower-moving storm crossing Taiwan.
For example, the forecasts made after 0000 UTC 6 August suggest high probabilities, i.e., at least 80–90%, of 48-h rainfall in excess of 1000 mm over a large portion of the southern CMR, and a ≥80% likelihood that some area will receive 1500 mm (Figure 8, bottom row). Thus, the probability information and, in particular, its temporal evolution obtained from such a high-resolution time-lagged ensemble can be highly informative at the short range (as the TC approaches Taiwan), even for an extreme and challenging event such as TY Morakot (2009). In some other TC cases where the tracks are more predictable with smaller errors, high-quality QPFs can become available at longer lead times [45], i.e., sometimes more than 4–5 days before landfall [85]. As emphasized in Section 1, however, the strength of the lagged ensemble lies in different areas at such longer ranges, mainly in the various scenarios produced (to cover the actual one) due to a larger spread (i.e., low predictability), rather than in the accuracy and consistency of QPFs.

While it is acknowledged that the later members made their QPFs at a shorter range (and likely with smaller track errors), the above probabilities of extreme rainfall are quite high compared to those obtained in some earlier studies. For example, at least 50% (30%) from a total of 60 ensemble forecast members (4.5-km) predicted 3-day total rainfall over 1000 (1500) mm in a study by Zhang et al. [65] which used perturbations produced by an ensemble data assimilation system, at a relatively early initial time of 0000 UTC 5 August. In [58], the 32-member, 4-km ensemble simulations (i.e., using analyses for both IC and BCs) with $t_0$ at 0000 UTC 6 August yielded probabilities of 1000 and 1500 mm (in 48 h), i.e., comparable to ours, but the computational cost was estimated to be about 10 times as much applying the former approach. Thus, efficiency, cost-effectiveness, and improved feasibility are also strengths of the time-lagged method adopted here, as reviewed in Section 1.

### 3.4. 24-h QPFs and the Probability

Similar QPFs and probabilities can be produced for the 24-h rainfall on 8 August for TY Morakot (2009) using seven runs, as shown in Figures 9 and 10. For this period, two more runs are available with $t_0$ at 1800 UTC 7 and 0000 UTC 8 August, and the derived probabilities at the short range are highly accurate (Figure 10, bottom row), i.e., up to 80–90% probability that regions in southern CMR will receive up to 500–1000 mm (per 24 h) of rainfall, and typically less than about 20% for most regions that turned out to not receive these prescribed rainfall amounts. Even before 1200 UTC 6 August, the potential in southern CMR to receive at least 700, or even 1000 mm, is indicated (Figure 10, top row). Overall, due to the strong topographic control, the QPFs made since 0600 UTC 6 August (64 h before the Shiao-Lin tragedy) are mostly of high quality and resemble the observations (Figures 4 and 9), so the rainfall variability in the hindcasts (or forecasts if in real time) may not be as large as some might think once the track can be accurately predicted (cf. Figure 3). In any case, the timely and rapid evolution of rainfall probabilities, as shown in Figures 8 and 10, may provide invaluable information for decision making and hazard mitigation in real time in an operational setting. As no additional cost in computation is required, the lagged strategy can also be useful for risk assessments of other weather or climate phenomena besides QPFs, such as TC intensity at landfall, for example [86].

Having only four members added per day, the only major shortcoming in the single-model time-lagged ensemble [45,46] presented here, one may argue, is the relatively few runs from which to derive the probability information. When computational resources allow, this deficiency can be remedied by adding a few more members, each with a similar resolution and capability, and the entire system will still be more efficient than running a large number of ensemble members simultaneously. With more reliable probability information from a greater number of members (and perhaps also using ensemble analyses/forecasts as IC/BCs), the system can be further improved. Such a configuration is consistent with the recommendations of Fang and Kuo [42] in their high-resolution (4-km) ensemble, but with a combined use of information from lagged members, and thus, improved efficiency.
Figure 9. As in Figure 4, but for the observed 24-h rainfall (mm, enlarged panel), from 0000 UTC 8 to 0000 UTC 9 August, during TY Morakot (2009) and the predicted rainfall (mm) targeted for the same period by the time-lagged ensemble every 6 h (smaller panels), at initial times from 1800 UTC 3 to 0000 UTC 8 August as labelled. The same color scale (lower right) is used for all panels.

Figure 10. As in Figure 8, but for the probabilities of 24-h QPFs (%, color) from the seven members in Figure 9 with $t_0$ between (top row) 0000 UTC 5 and 1200 UTC 6 and (bottom row) 1200 UTC 6 and 0000 UTC 8 August (as labelled), to reach thresholds of 200, 350, 500, 700, and 1000 mm (left to right), respectively, for the target date of 8 August during TY Morakot (2009). The observed areas (over land) at the respective threshold are depicted by the thick black contours. The weights of members are the same as those in Figure 8.

4. Conclusions and Summary

During 7–9 August in 2009, TY Morakot struck Taiwan and brought extreme rainfall to the southern part of the island. With a peak amount of 2635 mm in 48 h, the storm caused widespread flooding and numerous mudslides, especially over in mountainous interior of the CMR. In this study, the highly-efficient time-lagged ensemble strategy [43–45] is applied to this unique case, with emphases on the two following aspects: (1) To find out whether a
quality QPF for Morakot is possible beyond the short range, and (2) the time evolution of probability information derived from the time-lagged ensemble, as both aspects were not explored previously in the literature. Thus, a series of hindcast (or re-forecast) experiments was carried out using the 2.5-km CReSS model every 6 h at initial times from 0000 UTC 1 August to 0000 UTC 9 August 2009, with real-time NCEP GFS gross analyses and forecasts as IC/BCs. To assess the predictability of TY Morakot (2009) in our hindcasts, only data available at and before each run were used. The major findings can be summarized below.

(i) Starting from 0600 UTC 6 August, all but one of the time-lagged members in the ensemble captured the magnitude of the extreme rainfall in southern CMR reasonably well, with a peak 48-h amount near or over 2500 mm. Thus, the TY Morakot event (2009) could have been predicted beforehand, provided that the lead time was short enough to bring about adequately small track errors. In real time, the initial times of these runs would be at least 34 h, and up to 64 h, before the Shiao-Lin tragedy. The above results are consistent with earlier studies.

(ii) However, rain predictions close to the worst-case scenario for Taiwan could not be obtained prior to 6 August by the time-lagged ensemble, as in some other typhoons, mainly due to the larger track errors at longer lead times. This situation limits the practical predictability of the event in real time. When resources allowed, adding more members with a similar capability and/or using different analysis/forecast products, such as IC/BCs, would be potentially helpful to make further improvements to the ensemble information.

(iii) Even with a limited number of members, the time-lagged ensemble can provide useful information regarding the probability of heavy to extreme rainfall. Using more weight for later runs, ≥80–90% chance is obtained for a 48-h rainfall in excess of 1000 mm in southern CMR from runs starting at 6 August, and ≥80% predicted over 1500 mm in parts of the area. In real time situations, the evolution of probabilities can be effective to reflect the changing conditions in the forecasts as a TC approaches, thus providing invaluable information concerning decision making and hazard mitigation.

Author Contributions: Conceptualization, C.-C.W.; Formal analysis, C.-C.W. and S.-H.C.; Funding acquisition, C.-C.W.; Investigation, C.-C.W. and S.-H.C.; Data curation, S.-H.C. and C.-S.C.; Methodology, C.-C.W. and S.-H.C.; Project administration, C.-C.W.; Software, S.-H.C. and K.T.; Supervision, C.-C.W.; Visualization, S.-H.C. and S.-Y.H.; Writing—original draft, C.-C.W. and S.-H.C.; Writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Ministry of Science and Technology (MOST) of Taiwan, under grants MOST-108-2111-M-003-005-MY2, MOST-110-2111-M-003-004, and MOST-110-2625-M-003-001.


Acknowledgments: The authors thank the anonymous reviewers for their valuable comments and suggestions to improve the manuscript. The GFS analyses and forecasts used to drive the CReSS hindcasts are produced and made available by the NCEP. The CWB is acknowledged for providing Figure 2 and the rain-gauge data used for the verification of model QPFs in this study.

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

References
1. Golding, B. Quantitative precipitation forecasting in the UK. J. Hydrol. 2000, 239, 286–305. [CrossRef]


23. Epstein, E.S. Stochastic dynamic prediction. *Tellus* 1969, 21, 739–759. [CrossRef]


59. Liou, Y.-C.; Wang, T.-C.C.; Tsai, Y.-C.; Tang, Y.-S.; Lin, P.-L.; Lee, Y.-A. Structure of precipitating systems over Taiwan’s complex terrain during Typhoon Morakot (2009) as revealed by weather radar and rain gauge observations. *J. Hydrol.* **2013**, 506, 14–25. [CrossRef]


77. Schaefer, J.T. The critical success index as an indicator of warning skill. Weather Forecast. 1990, 5, 570–575. [CrossRef]


