

Article

Development of a Hydrological Ensemble Prediction System to Assist with Decision-Making for Floods during Typhoons

Sheng-Chi Yang ¹, Tsun-Hua Yang ²,*¹, Ya-Chi Chang ³,*, Cheng-Hsin Chen ⁴, Mei-Ying Lin ⁵, Jui-Yi Ho ⁵ and Kwan Tun Lee ^{6,7}

- ¹ National Applied Research Laboratories, Taipei 106, Taiwan; shirky@narlabs.org.tw
- ² Department of Civil Engineering, National Chiao Tung University, Hsinchu 300, Taiwan
- ³ Green Energy and Environment Research Laboratories, Industrial Technology Research Institute, Hsinchu 310, Taiwan
- ⁴ Department of Civil Engineering, National Chung Hsing University, Taichung 402, Taiwan; mas@dragon.nchu.edu.tw
- ⁵ National Science and Technology Center for Disaster Reduction, Taipei 231, Taiwan; mylin@ncdr.nat.gov.tw (M.-Y.L.); juiyiho@ncdr.nat.gov.tw (J.-Y.H.)
- ⁶ Department of River and Harbor Engineering, National Taiwan Ocean University, Keelung 202, Taiwan; ktlee@ntou.edu.tw
- ⁷ Center of Excellence for Ocean Engineering, National Taiwan Ocean University, Keelung 202, Taiwan
- * Correspondence: tshyang@nctu.edu.tw (T.-H.Y.); Rachel.Chang@itri.org.tw (Y.-C.C.)

Received: 18 March 2020; Accepted: 19 May 2020; Published: 22 May 2020



Abstract: Hydrological ensemble prediction systems (HEPSs) can provide decision makers with early warning information, such as peak stage and peak time, with enough lead time to take the necessary measures to mitigate disasters. This study develops a HEPS that integrates meteorological, hydrological, storm surge, and global tidal models. It is established to understand information about the uncertainty of numerical weather predictions and then to provide probabilistic flood forecasts instead of commonly adopted deterministic forecasts. The accuracy of flood forecasting is increased. However, the spatiotemporal uncertainty associated with these numerical models in the HEPS and the difficulty in interpreting the model results hinder effective decision-making during emergency response situations. As a result, the efficiency of decision-making is not always increased. Thus, this study also presents a visualization method to interpret the ensemble results to enhance the understanding of probabilistic runoff forecasts for operational purposes. A small watershed with area of 100 km² and four historical typhoon events were selected to evaluate the performance of the method. The results showed that the proposed HEPS along with the visualization approach improved the intelligibility of forecasts of the peak stages and peak times compared to that of approaches previously described in the literature. The capture rate is greater than 50%, which is considered practical for decision makers. The proposed HEPS with the visualization method has potential for both decreasing the uncertainty of numerical rainfall forecasts and improving the efficiency of decision-making for flood forecasts.

Keywords: hydrological ensemble prediction system; numerical weather model; flood forecast; peak flow; visualization

1. Introduction

Ensemble prediction systems (EPSs), which consist of an adequate number of equiprobable numerical weather prediction (NWP) models, have been established to provide probabilistic



precipitation forecasts instead of a single deterministic forecast [1]. An EPS provides predictions with greater skill than those obtained from individual runs of NWP models or deterministic model runs, especially for longer lead times [2]. A hydrological ensemble prediction system (HEPS) is an integrated system that couples an EPS with catchment-scale hydrological models to provide flood forecasts. Many studies have used HEPS to conduct flood forecasting, with each study applying a different meteorological EPS input coupled with a different hydrological model based on the properties of the catchment of interest [3–5]. Balint et al. [4] analyzed more than 5000 model runs and tested the feasibility of HEPS for flood forecasting of the Danube. Based on the European Center for Medium-Range Weather Forecasts (ECMWF) ensemble predictions, the HEPS is implemented by the National Hydrological Forecasting Service (NHFS) modeling system of VITUKI (the Hungarian hydrological service). Davolio et al. [5] proposed a HEPS based on multi-model ensemble QPF (quantitative precipitation forecast) along with the rainfall-runoff model TOPKAPI (TOPographic Kinematic APproximation and Integration) [6] for the Reno River basin in northern Italy. Researchers are in general agreement that HEPSs for flood warning purposes have added forecast value compared to the previously adopted deterministic forecasts [1,7].

Uncertainties stemming from factors including boundary conditions, initial conditions, and model parameter values or different models applied in HEPSs for flood forecasting affect the forecast results. Abebe et al. [8] suggested that the consideration of the degree of uncertainty under which the forecasts are produced is essential for emergency response. Effective communication of ensemble forecasts means that clearly express the uncertainties associated with HEPSs is important so that end-users can easily respond to the provided information [9–11]. Pappenberger et al. [9] argued that the uncertainty information provided by HEPSs sometimes results in resistance on the part of the public if experts or nonexperts cannot easily understand the provided information. Pagano et al. [11] noted that defining effective methods for the communication of ensemble forecasts is a challenge for operational river forecasting and represents a research challenge. Uncertainties or differences within HEPSs still rely on conventional visualization techniques, such as "spaghetti diagrams" or box plots, to display the distributions of forecast results. Pappenberger et al. [9] focused on expert users of HEPSs and the communication among these experts and identified key information for the public, such as discharge, lead time, warning levels, return periods, and worst/best scenario. Zappa et al. [10] proposed a visualization approach to support the interpretation and verification of HEPS results. Giordani et al. [7] modified the approach from Zappa et al. [10] and developed another method for multiple peak flow detections. Olsson et al. [12] used this approach to process the ensemble runoff forecasts for joint probabilistic visualization of peak flow runoff and timing. This approach has been used to obtain quantitative and qualitative insights, such as the timing, water level, and discharge associated with peak flow.

An average of 3.7 typhoons hit Taiwan annually, resulting in enormous loss of life and properties [13]. Hsiao et al. [14] used an ensemble meteorological modeling system to predict typhoon rainfall and coupled it with a hydrological model that considers only rainfall runoff and river routing simulations to predict flood in the Lanyang watershed in Taiwan. Instead of providing a single-valued forecast from a deterministic system, they deliver a probabilistic forecast to the users.

It is well known in the scientific community around the world that ensemble or probabilistic forecasts contain more information than a single-valued forecast. This study therefore integrates the EPS from the Taiwan Cooperative Precipitation Ensemble Forecast Experiment (TAPEX) with 5 km spatial resolution and hydrological models to establish HEPSs. Since Taiwan is an island, the water level downstream are related to storm surge and tides during a typhoon. The hydrological models in the proposed HEPSs therefore include not only rainfall-runoff and river-routing models but also a storm-surge model, which have rarely been used. In practice, however, the users might prefer deterministic hydro-meteorological forecasts because they commonly misused the information of probabilistic forecasts. To solve the conflicts between probabilistic and deterministic forecasts for end-users, this study also proposes a visualization approach to simplify the HEPS forecast information

and therefore decrease the complexity of the decision-making process for common users. Together with the presented visualization approach, the forecast accuracy and emergency response efficiency will be improved. The same idea can be applied to areas or countries that are usually influenced by typhoon events. The remainder of this paper is organized as follows. Section 2 briefly describes the study area and typhoon events analyzed in the study. Section 3 details the developed HEPS and the proposed visualization approach. Finally, Sections 4 and 5 present the results, discussion, and conclusions.

2. Study Area and Typhoon Events

2.1. Study Area

The Yilan River in northeastern Taiwan was selected as the study area (Figure 1). The river flows through the city of Yilan, the main river length is approximately 24.4 km, and the catchment area is 149.06 km². It has four main tributaries: the Wushi River, the Dahu River, the Dajiao River, and the Xiaojiao River. There are 11 rainfall gauging stations, 16 water-stage gauging stations, five water surface-velocity gauging stations, and 36 inundation-depth gauging stations installed in the Yilan River Basin. Figure 1 also shows the locations of the water-stage and rainfall gauging stations that collected the data used in this study. The monitoring data have been carefully collected and processed. In addition, the flow discharge data used in this study were measured by moving-boat Acoustic Doppler Current Profiler (ADCP) and velocity-extrapolation methods during typhoon events.



Figure 1. Study area and locations of streamflow gauges. Black dots and triangles indicate the locations of water-stage gauging stations and rain gauge stations, respectively.

2.2. Typhoon Events

Typhoon tracks are usually determined by the center of a typhoon. The Central Weather Bureau (CWB) categorized the typhoon tracks that invaded Taiwan into nine specific categories (Category 1 through 9) and one special category using historical typhoon data from 1911 to 2018 as shown in Figure 2 [15]. Of the 10 categories, Track-2 and Track-3 typhoons account for approximately 28% of all typhoons and bring heavy rainfall to the Yilan River Basin. For instance, a rainfall of 158 mm in 4 h was observed at the rainfall gauging station C1U610 (shown in Figure 1) during Typhoon Soulik. Table 1 shows all the typhoons with a warning issued by CWB from 2012 through 2015. Five of these events are Track-2 and Track-3 typhoons, which have the biggest impact on the Yilan River Basin. Therefore, this study selected the typhoons Saola (2012), Soulik (2013), Soudelor (2015), and Dujuan (2015) to

calibrate and validate the Yilan River HEPS. Typhoon Matmo, a Track-3 typhoon that occurred in 2014, was not included due to its weak intensity.



Special Tracks 3.95%

Figure 2. Schematic diagram showing the tracks of typhoons invading Taiwan. The percentages shown in the figure are the statistical results from 1911 through 2018 obtained from the Central Weather Bureau (CWB) in Taiwan. The dark gray polygon located in northern Taiwan indicates the Yilan River catchment. Track-2 and Track-3 typhoons bring heavy rainfall to the Yilan River catchment.

Table 1. All typhoons with warning issued by CWB in Taiwan during 2012 through 2015. A total of four typhoons of Track-2 and Track-3, namely, Saola in 2012, Soulik in 2013, Soudelor in 2015, and Dujuan in 2015, were selected to calibrate the system and test the performance. Typhoon Matmo, a Track-3 typhoon that occurred in 2014, was not selected due to its weak intensity.

Typhoon	Track	Intensity Warning Period				
DUJUAN	2	3	27–29 September 2015			
GONI	-	-	20–23 August 2015			
SOUDELOR	3	3	6-9 August 2015			
LINFA	-	-	6–9 July 2015			
CHAN-HOM	-	2	9–11 July 2015			
NOUL	-	-	10–11 May 2015			
FUNG-WONG	Special	-	19–22 September 2014			
MATMO	3	-	21–23 July 2014			
HAGIBIS	-	3	14–15 June 2014			
FITOW	1	-	4–7 October 2014			
USAGI	5	3	19–22 September 2013			
KONG-REY	6	-	27–29 August 2013			
TRAMI	1	-	20–22 August 2013			
CIMARON	-	-	17–18 July 2013			
SOULIK	2	1	11–13 July 2013			
JELAWAT	-		27-28 September 2012			
TEMBIN	Special	-	21–25 August 2012			
	Special	-	26–28 August 2012			
KAI-TAK	-	1	14–15 August 2012			
HAIKUI	-	-	6–7 August 2012			
SAOLA	2	4	30 July–3 August 2012			
DOKSURI	-	-	28–29 June 2012			
TALIM	9	-	19–21 June 2012			

3. Hydrological Ensemble Prediction System

Figure 3 shows the flowchart of data processing in the HEPS. It integrates NWP models that provide 72-h ensemble precipitation forecasts, a rainfall-runoff model that calculates the surface runoff and estimates flow discharge at Hsincheng and Yuanshan Bridges as the upstream boundary conditions, a storm-surge model coupled with a global tide model that generate the water stage at Kemalan Bridge as the downstream boundary conditions, and a flood-routing model that simulates Yilan River flows. The abovementioned geographical locations can be found in Figure 1. The Yilan River HEPS produces ensemble flood forecasts with a 72-h lead time four times a day. The details of the models are described in the following subsections.



Figure 3. Flowchart describing the flow of data processing within the Yilan River hydrological ensemble prediction system (HEPS).

3.1. Ensemble Precipitation Forecasts

TAPEX began in 2010. It is a collective effort among academic institutes and government agencies, such as the National Taiwan University, the National Central University, the National Taiwan Normal University, the Chinese Culture University, the CWB, the National Center for High-Performance Computing, the Taiwan Typhoon and Flood Research Institute (TTFRI), and the National Science and Technology Center for Disaster Reduction. TAPEX is the first attempt to design a high-spatial-resolution (5 km) numerical ensemble model in Taiwan. This effort applies various NWP models, such as the Weather Research and Forecasting Model (WRF), the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), the Cloud-Resolving Storm Simulator (CReSS), and the Hurricane Weather Research and Forecasting Model physics. TAPEX generates four runs a day and provides ensemble predictions of the wind and pressure fields and quantitative estimates of precipitation with a lead time of 72 h. Further information can be found in reference [13]. This study focuses on a one-way coupling in which TAPEX provides rainfall forecast to the rainfall-runoff model; feedbacks from the rainfall-runoff model to TAPEX are not considered.

3.2. Hydrological Methods

3.2.1. Rainfall-Runoff Model

The Yilan River HEPS uses the surface runoff forecast generated by a kinematic-wave-based geomorphologic instantaneous unit hydrograph model (the KW-GIUH model) as the upstream boundary condition of the river-routing model. The KW-GIUH model, which was developed by Lee and Yen [16], can reflect the effects of watershed geomorphology, land cover conditions, soil characteristics and rainfall intensity on runoff. It has been successfully applied to many Taiwanese catchments [17–19].

Two parameters in the proposed HEPS KW-GIUH model have been calibrated using in situ observations made during typhoon events. These parameters are the roughness coefficient for overland flow (n_0) and the roughness coefficient for channel flow (n_c) and depend on the type of land use and bed material of the river in the analyzed basin. Five rainfall gauges, including LTGX, YSGZ, C1U610, C0U520, and C1U630 (see Figure 2 for locations), were used, and the Thiessen polygon method [20] was used to estimate the hourly spatial-average rainfall intensities used as rainfall input data to the KW-GIUH model. The observed flow data at the Hsincheng and Yuanshan Bridges during Typhoons Saola, Soulik, and Soudelor were used to calibrate and validate the parameters. As shown in Figure 4, the percent errors in the peak discharges between observed and simulated flow of the selected typhoons were 4.59% (Saola), 2.07% (Soulik), and -5.89% (Soudelor) at the Hsincheng Bridge and 14.88% (Saola), 5.28% (Soulik), and -3.05% (Soudelor) at the Yuanshan Bridge. All of the errors in the peak times were less than one hour. The results show that the proposed HEPS KW-GIUH model is capable of providing accurate simulations for peak time as well as peak discharge.



Figure 4. Comparison of simulated discharges (red circles) and recorded discharges (solid lines) for model calibration (Typhoons Saola and Soulik) and validation (Typhoon Soudelor) experiments at Hsinsheng (left) and Yuanshen (right). The blue bars are the hourly spatial-average rainfall intensities measured in the watershed upstream of Hsinsheng and Yuanshen.

3.2.2. Storm-Surge Model

Storm surges are abnormal increases in water levels above those expected from astronomical tides. They are generated by strong winds and atmospheric pressure changes and affect water levels downstream (near estuaries) during typhoons. The Yilan River HEPS uses the storm surge and tide forecasts generated by the Princeton Ocean Model (POM) and the TOPEX-POSEIDON global tidal model (TPXO6.2) as downstream boundary conditions. The POM model, which was developed by Blumberg and Mellor [21], is a three-dimensional, nonlinear, primitive equation finite difference ocean model. It has been applied to simulate a wide range of ocean scenarios, including coastal storm surge in Taiwan [22,23]. In this study, the TAPEX model provides ensemble pressure field and wind field forecasts to POM and the TPXO6.2 model for obtaining tidal level predictions. As with TAPEX, it generates four runs a day, and each run has a 72-h lead time. Two typhoons (Saola and Soulik) were used to evaluate the storm-surge model due to the severe lack of observations for Typhoon Soudelor at Suao tide station. As shown in Figure 5, the percent errors in the peak storm surge were -9.4% for Saola and 19.4% for Soulik.



Figure 5. Comparison of simulated (red circles) and observed (solid lines) storm surges for model calibration (Typhoon Saola) and validation (Typhoon Soulik) experiments at Suao tide station.

3.2.3. River-Routing Model

The Numerical Model Simulating Water Flow and Contaminant and Sediment Transport in WAterSHed Systems of 1D Stream/River Networks, 2D Overland Regimes, and 3D Subsurface Media (WASH123D) developed by Yeh et al. [24] was used to simulate one-dimensional channel networks, two-dimensional overland flow, and three-dimensional variably saturated subsurface flow. The flow in river/stream networks in WASH123D is based on the one-dimensional St. Venant equation and the simulation in this model is conducted with finite-element approaches of various spatial-temporal scale computations [25]. It has been applied successfully in Taiwan and around the world, and it was chosen by the US Army Corps of Engineers as the core computational code used in modeling the Lower East Coast (LEC) Wetland Watershed [26–28]. The Yilan River HEPS uses the one-dimensional channel model of WASH123D as its flood-routing model to simulate water stages in rivers.

The Yilan River HEPS adopted the most recent available cross-sectional bathymetry of the Yilan River, which was measured in 2010, as its input topography data. The upstream boundary of the model is set at the Hsincheng and Yuanshan Bridges, and the downstream boundary of the model is set at the Kemalan Bridge. Figure 6 shows that the percent errors in comparison with the observed peak stages for Typhoons Saola, Soulik, and Soudelor were 2.1%, 5.7%, and 10.6% at Zhongshan, 12.9% and 2.2% at Leawood, and 7.4%, 6.0%, and 2.1% at Jhuangwei, respectively. There was one data gap at Leawood due to incomplete observed data collection during Typhoon Soudelor. Nevertheless, all of the errors in the peak times were less than one hour. The results show that the Yilan River HEPS is capable of providing confident predictions of peak times, as well as peak stages.



Figure 6. Comparison of simulated (red circles) and observed (solid lines) water levels for model calibration (Typhoons Saola and Soulik) and validation (Typhoon Soudelor) experiments at Zhongshan (left), Leawood (central) and Jhungwei (right).

3.3. A Visualization Approach for Supporting the Interpretation of Operational Ensemble Peak Flow Forecasts during Typhoon Events

Forecasts of HEPSs are usually characterized by a large ensemble spread. Consequently, their evaluation may be complex for decision-making by common users, especially during emergency response. Therefore, Zappa et al. [10] first proposed a visualization approach, called "Peak-Box", with the purpose of helping decision-making when forecasting flood events. Giordani et al. [7] modified it to detect multiple peak flows. Figure 7 Left shows the "Peak-Box" visualization approach and the concept is described as follows. First, a rectangle (the "Peak-Box"), which is enveloped by the coordinates of four values, i.e., the earliest time of peak flow (t_0) , the latest time of peak flow (t_{100}) , the lowest peak discharge (p_0) and the highest peak discharge (p_{100}) of all ensemble members. Second, an inner rectangle (the "IQR-Box"), which is enveloped by the coordinates of four values, i.e., the 25% quartile of the peak timing (t_{25}), the 75% quartile of the peak timing (t_{75}), the 25% quartile of the peak discharge (p_{25}) and the 75% quartile of the peak discharge (p_{75}) of all ensemble members. A horizontal line is drawn from t0 to t100, which represents the median of the peak discharge (p_{50}) of all ensemble forecasts. Finally, a vertical line is plotted from p0 to p100, which represents the median of the peak timing (t_{50}) of all ensemble forecasts. Based on the above-mentioned approach, this study proposed a modified visualization method specifically for the proposed HEPS and for typhoon events. Figure 7 Right shows the proposed visualization approach and the modifications are described below.

- a. **Simplify and avoid misused information.** Pappenberger et al. [9] noted a considerable desire on the part of end-users to reduce probabilistic forecasts to deterministic actions. The proposed visualization approach removes the horizontal and vertical lines because end-users commonly misused them for decision-making. The two lines may lead end-users to believe that the information provided represents a single deterministic forecast, rather than a probabilistic one. The outer rectangle "Peak-Box" is also removed when the availability of too much data may obscure critical information during typhoons. Therefore, only one rectangle is shown in the proposed approach. This rectangle only indicates where the observed peak stage and its occurring time are likely to occur.
- b. **Rescale the rectangle.** This study defines an "SD-Box" that uses the mean (μ) and the standard deviation (σ), instead of the 25% and 75% quartiles, to define the enveloping rectangle. As shown

in the right panel of Figure 7, the lower left coordinate of the "SD-Box" is defined as the mean peak time minus one standard deviation ($\mu_t - \sigma_t$), and the mean peak stage minus one standard deviation $(\mu_h - \sigma_h)$ produced by all of the ensemble members. The upper right coordinate is defined as the mean peak time plus one standard deviation ($\mu_t + \sigma_t$) and the mean peak stage plus one standard deviation $(\mu_h + \sigma_h)$ of all of the ensemble members. The standard deviation takes into account the spread of ensemble forecasts. In addition, using the mean and the standard deviation instead of the quartile deviation to determine the second rectangle allows the inclusion of a higher number of ensemble forecasts and have more opportunities to cover observed peaks. In principle, the "IQR-Box" should contain 25% (50% of the peak discharge multiplied by 50% of the peak times) of all forecasts. Zappa et al. [10] showed that this box only contained between 12.5% and 37.5% due to the distribution of ensemble members. The "SD-Box" includes the mean and the standard deviation and results in a larger area. It includes 46.60% of the ensemble forecasts (68.27% of peak water level multiplied by 68.27% of the peak times); therefore, it should have a greater chance of including the observed peaks. A greater rectangle may generate overestimation. Because there is only one rectangle remaining, the simplified information can make for efficient decision-making.

c. Include all forecasts with different lead times in the rectangle. Descriptive statistics, such as the quartile deviation and the standard deviation, are susceptible to outliers when calculated using insufficient sample sizes. Since adding extra ensemble members to produce more forecasts and thus increase sample size consumes computer resources, this study includes present (*t*) and previous forecasts (t - 1, t - 2, t - 3... t - n, where *n* is the number of available forecasts when the system is initiated) to expand the sample size and provide a better interpretation of results. As shown in the right panel of Figure 7, the green area illustrates the "SD-Box". The black and gray solid dots represent the current and previous peak flow forecasts, respectively. The "SD-Box" is designed exclusively for typhoons for operational purposes, and it is initiated when the CWB issues a sea warning and ends when the next ensemble forecast is six hours less than the left edge of the "SD-Box".



Figure 7. The left panel shows a graphical explanation of the "Peak-Box" approach from Zappa et al. [10]. The outer rectangle is the "Peak-Box", and the internal rectangle (the yellow area) is the inner rectangle box (IQR-Box). The solid dots represent all of the ensemble forecasts, and the red cross is the observed peak flow. The right panel shows a graphic comparison of the "IQR-Box" and "SD-Box" in the proposed modification of the "Peak-Box" approach. The enveloping rectangle is the "SD-Box" (the green area), and the yellow rectangle is "IQR-Box". The solid black and gray dots represent current and previous peak flow forecasts, respectively.

4. Results and Discussion

4.1. Performance Evaluation of the Yilan River HEPS

This study focused exclusively on peak flow in terms of its magnitude during the peak stage and peak timing. Therefore, this study applied two widely used performance measures to evaluate the performance of the HEPS: the root-mean-square error (RMSE) and the spread-skill score. The RMSE and the spread-skill score are defined as follows:

$$RMSE = \sqrt{\left(O_{peak} - \mu\right)^2} \tag{1}$$

$$Spread - skill\ score = \frac{RMSE}{\sigma}$$
(2)

$$\mu = \frac{1}{m} \sum_{i=1}^{m} P_{peak,i} \tag{3}$$

$$\sigma = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(P_{peak,i} - \mu \right)^2} \tag{4}$$

where μ is the mean of the peak flow forecast ensemble; σ is the standard deviation of peak flow forecast ensemble; O_{peak} is the observed peak flow; $P_{peak,i}$ is the peak flow prediction of the *i*th member; and *m* is the number of ensemble members.

The RMSE, which is commonly referred to as skill, measures the difference between the observations and the ensemble mean without considering the direction. The closer the RMSE is to zero, the better the ensemble mean is as a forecast. The spread-skill score is one of many metrics used to evaluate an ensemble system, which ranges from zero to infinity, is the ratio of the standard deviation of the ensemble peak flow forecasts (spread) to the RMSE [29]. Ensemble prediction should provide spread between the members that is large enough to cover the uncertainty in the prediction. For a well-calibrated ensemble forecast system, the spread should be of the same magnitude as the RMSE [10]. In that regard, the spread-skill score might be appropriate to evaluate the ensemble system proposed in this study and should be close to one to meet the requirements of a qualified ensemble system.

The forecasts used in this study were initiated when the CWB issues a sea warning and end when the next ensemble forecast is six hours less than the left edge of the "SD-Box". In that regard, 93 forecasts are available for the four selected typhoons. Table 2 shows the spread-skill score in peak stages and peak times considering only the current forecasts for the Dujuan, Soudelor, Soulik and Saola events at the Zhongshan, Leawood, and Zhuangwei Bridges. We evaluated 7, 9, 8 and 10 forecasts from 27 September 08:00 to 28 September 28 20:00 with 6-h intervals in Dujuan, from 6 August 08:00 to 8 August 08:00 with 6-h intervals in Soudelor, from 11 July 08:00 to 13 July 02:00 with 6-h intervals in Soulik, and 30 July 20:00 to 2 August 02:00 with 6-h intervals in Saola, respectively. The spread-skill scores were not calculated for the Leawood Bridge during Typhoon Soudelor due to the lack of complete observations. The spread-skill scores that are less than one (indicating that the spread of the forecasts exceeds the prediction uncertainty) in the table are highlighted. Most of the spread-skill scores of Dujuan were greater than 2. The system performed the worst for this event due to the large forecast track error from the NWP model used in the Yilan River HEPS, resulting in poor performance in rainfall forecasts [30]. Because most of the uncertainty comes from the NWP model, the accuracy of rainfall forecasts exhibited a significant impact on the performance of the HEPS. This also explained why Saola and Soulik had the best performance among all events. Overall, the "SD-Box" method yielded average spread-skill scores of 1.13 for the peak stages and 1.00 for the peak times.

Table 2. Comparisons of the spread-skill scores in the peak stage and peak time for the Yilan River HEPS. Scores less than one (highlighted) indicate that the enveloping box did contain the observed peak.(a) The spread-skill scores in peak stage forecasts. (b) The spread-skill scores in peak timing forecasts.

(a)												
Leastion/Tunkeen	Forecast											
Location/ Typhoon	1	2	3	4	5	6	7	8	9	10		
Zhongshan Bridge												
Dujuan (2015)	2.54	2.59	2.64	2.09	4.79	2.62	0.57	-	-	-		
Soudelor (2015)	0.41	0.60	1.88	0.93	2.76	2.82	2.27	4.59	1.78	-		
Soulik (2013)	1.07	1.27	1.39	0.76	0.64	0.38	0.15	0.40	-	-		
Saola (2012)	0.20	0.07	0.71	0.56	0.55	0.55	1.36	1.23	2.18	0.54		
Leawood Bridge												
Dujuan (2015)	1.21	1.27	1.75	1.24	3.48	1.48	1.67	-	-	-		
Soudelor (2015)	-	-	-	-	-	-	-	-	-	-		
Soulik (2013)	0.79	0.95	1.06	0.36	0.20	0.10	0.27	0.54	-	-		
Saola (2012)	0.93	1.25	1.66	1.32	1.41	0.16	0.29	0.22	0.04	1.36		
Zhuangwei Bridge												
Dujuan (2015)	1.97	2.13	0.60	0.21	0.46	1.51	2.94	-	-	-		
Soudelor (2015)	1.19	0.17	0.45	0.10	1.01	1.24	0.55	1.81	2.64	-		
Soulik (2013)	0.62	0.71	0.79	0.17	0.03	0.32	0.47	0.90	-	-		
Saola (2012)	0.82	1.08	1.40	1.14	1.26	0.09	0.70	0.09	0.22	1.29		
				(b)								
	Forecast											
Location/Typhoon	1	2	3	4	5	6	7	8	9	10		
Zhongshan Bridge												
Dujuan (2015)	1.34	1.38	4.33	1.83	2.83	1.86	0.68	-	-	-		
Soudelor (2015)	0.68	0.70	1.74	0.97	3.49	1.75	1.08	1.08	0.66	-		
Soulik (2013)	1.48	1.60	2.64	0.59	1.37	0.23	0.36	1.29	-	-		
Saola (2012)	0.07	0.26	0.28	0.02	0.37	0.58	0.30	0.01	0.79	0.48		
Leawood Bridge												
Dujuan (2015)	0.46	0.11	1.69	0.32	2.24	0.58	0.71	-	-	-		
Soudelor (2015)	-	-	-	-	-	-	-	-	-	-		
Soulik (2013)	0.40	1.17	1.96	0.39	0.71	0.11	0.09	0.96	-	-		
Saola (2012)	0.04	0.09	0.34	0.17	0.04	0.11	0.46	0.07	0.67	0.53		
			Zhu	uangwei	Bridge							
Dujuan (2015)	2.90	3.54	3.06	4.17	2.57	3.91	0.86	-	-	-		
Soudelor (2015)	0.40	0.48	1.32	0.72	3.20	1.42	1.04	1.08	0.28	-		
Soulik (2013)	0.42	0.59	1.08	0.81	0.16	0.68	0.08	0.70	-	-		
Saola (2012)	0.25	0.07	0.17	0.28	0.54	1.05	0.79	0.33	0.09	0.68		

4.2. Comparison of Enveloping Rectangles Defined Using the "SD-Box" and the "IQR-Box" Methods for Supporting the Interpretation of Ensemble Peak Flow Results

Both "SD-Box" and "IQR-Box" were applied to provide peak flow forecasts. There is only one box left for both approaches because "Peak-Box" was removed to avoid any confusion. Figures 8 and 9 show all forecast results. The horizontal axis is the number of forecasts. The vertical axes represent the length in water depth in Figure 8 and amount of time in Figure 9 to address the size of the rectangle, which is the size of the "Box" in Figure 7. In this analysis, only one forecast ensemble cycle at a time was compared, as indicated by the solid line shown as "xxx-Box single". If the observed peak fell inside the box, it was labeled as a "hit" (shown as solid circles and squares in Figures 8 and 9) at that

number of forecasts. The average hit rates using the "IQR-Box" method were 33.3% (31/93) and 52.6% (49/93) for peak flow stage and timing, respectively. Using the "SD-Box" method improved the average rate to 51.6% (48/93) and 64.5% (60/93) for stage and timing, respectively. It is noted that the individual performance of Typhoons Saola and Soulik are better than other events. One of the reasons is that the parameters in the hydrological models were calibrated using observations from Typhoons Saola and Soulik. The purpose of the "Box" is not to generate a deterministic forecast but to simplify the information out of the ensemble forecasts. The sizes of the box help users to understand the event range in terms of peak stage and peak timing. The decision maker can act accordingly based on a range as long as the observations are contained in the box. The results showed that the enveloping rectangles defined using the "SD-Box" method were more reliable for decision-making during typhoon events (i.e., both have average hit rates exceeding 50%).

Figures 8 and 9 shows that "SD-Box" has comparatively larger values of length in water stage and in time than "IQR-Box". The results show that the length in water stage of the downstream location (i.e., Zhuanwei Bridge) is smaller than the ones of upstream (i.e., Zhongshan and Leawood Bridge). However, the results of the length in time are consistent along the stream. Because the method of defining the size of the box differed, the "SD-Box" method resulted in a comparatively larger region than the "IQR-Box" method for most cases in the figures. It is why blue makers are usually larger than black ones in the figures. It is arguable that a larger box should have a better hit rate. However, the performance (hits which are solid ones in figures) still highly depends on the performance of rainfall forecasts. For example, the length in water stage for "SD-Box All" is larger than that for "SD-Box Single" during of Dujuan Typhoon, however, it shows no hits for "SD-Box All". A larger box does not always mean a higher hit rate if none of spread results can catch the observations. The rainfall forecast capability of the system is important and reliable so the box can be described based on the forecast results. As a comparison with previous study, the "SD-Box" method produced a smaller box than the "Peak-Box" method. For decision-making during typhoon events, it is considered more efficient to apply the "SD-Box" method instead of a combination of the "Peak-Box" and "IQR-Box" methods.



Figure 8. Changes in the magnitude of the peak stage using the "SD-Box" (blue) and "IQR-Box" (black) methods considering only the current forecast (solid lines) and inclusion of all the available forecasts at the present time (dotted lines). The solid circles and squares represent that the observations are contained in the "Box", which is a hit, at that number of forecasts.



Figure 9. Changes in the duration of the peak stage using the "SD-Box" (blue) and "IQR-Box" (black) methods considering only the current forecast (solid lines) and inclusion of all the available forecasts at the present time (dotted lines). The solid circles and squares represent that the observations are contained in the "Box", which is a hit, at that number of forecasts.

4.3. Inclusion of All Forecasts with Different Lead Times during an Event to Expand the Sample Size

The sample size has a strong effect in terms of determining whether a result is statistically significant. In other words, the number of available ensemble members is important for both the "SD-Box" and "IQR-Box" methods. For example, the number of available ensemble members for each forecast ranged from 11 to 14 during operations. Thus, the descriptive statistics were calculated using insufficient sample sizes (less than 30). The same issue exists in other studies that employ HEPSs [10,31]. It is difficult to increase the number of ensemble members used in HEPSs, due to the limited computational resources that are available.

As mentioned in Yang et al. [32], there is no evidence to show the forecast performance will be significantly improved with lead time during a typhoon event. In order to discuss the relationship between accuracy of forecast and lead time, this study calculated the error in the maximum 4-h rainfall between the average forecasts and the average observations at the watershed upstream of the Zhongshan Bridge (4-h used herein indicates the time of concentration of the peak flow at the Zhongshan Bridge is approximately 4 h). Figure 10 shows that there is no obvious trend in the errors in stage and timing, regardless of the length of the lead time. The correlation coefficients were -0.09 and 0.11, respectively, and these values indicate that no significant correlations exist between errors in stage or timing on the one hand and lead time on the other. For example, the best and worst forecasts during Typhoon Dujuan in terms of stage error were the first and fifth forecasts, respectively. However, the 6th forecast was better than the 5th, which implies that there is no trend in the cascading forecasting process.

It is worth noting that the conclusion that there is no significant trend in forecast accuracy and lean time is made based on limited ensemble members. Bulk ensemble members of forecast might improve the accuracy of metrological variables when the forecast is close to the event. However, as mentioned above, it is difficult to increase ensemble members for each run or the number of tropical cyclone events analyzed. As a trade-off, this study proposes a method for including present and previous forecasts in order to expand the sample size during the estimation process.





Figure 10. The box-and-whisker plot of (**a**) the magnitude error in the maximum 4-h rainfall and (**b**) the timing error of maximum 4-h rainfall at the watershed upstream of the Zhongshan Bridge during the four selected typhoon events. The blue dots indicate the ensemble means. The inverted triangles indicate the time of occurrence of the maximum 4-h rainfall. The results show that there is no obvious trend in lead time for the errors in either the rainfall or timing. The error is defined as the predicted rainfall (timing) minus observed rainfall (timing).

Figure 11 illustrates comparisons between using one forecast and using all available forecasts with the "SD-Box" method (hereinafter indicated as "SD-Box Single" and "SD-Box All") at Zhongshan Bridge. The performance of "SD-Box All" was more consistent than that of "SD-Box Single" in terms of both stage and timing. For example, the spread-skill scores for stage during Typhoon Soudelor ranged from 0 to 5 when the "SD-Box Single" method was used, but they were more consistent and closer to 1 with "SD-Box All". The results showed that the inclusion of all available forecasts ("SD-Box All") in the calculation process could expand the sample size and obtain more robust descriptive statistics than "SD-Box Single" does. The spread-skill score of "SD-Box Single" ranges between 2 and 5 implies that the spread between the members is too concentrated to cover the observed peaks. In addition, Figures 8 and 9 show the features in terms of the size of the box using "SD-Box All" and "SD-Box Single" approaches. The performance of "SD-Box Single" is comparatively similar to the one of "SD-Box All". The biggest differences appear especially in the performance of upstream location. For example, the "SD-Box Single" cannot catch the peak stage at Zhongshan Bridge during Soudelor Typhoon but "SD-Box All" can. Although the box sizes of "SD-Box All" is larger than that of "SD-Box Single" (in most of forecast runs), the variations in the water depth and length of time were consistent in each update for "SD-Box All". In other words, the variation in box size of "SD-Box All" is getting stable from update to update and the stable variation of the box might be useful for decision-making. This result implies that the forecasts including present and previous forecasts improve the performance of the HEPS and help decision-making because the results of "SD-Box All" is more consistent and stable. Finally, Figure 12 illustrates the scores of all of the forecasts for the different typhoon events. The "SD-Box Single" contained 51.7% of the observed peaks in terms of stage (40.9% + 10.8%), whereas "SD-Box All" contained 69.9% (62.4% + 7.5%) of the observed peaks. Furthermore, the "SD-Box Single" contained 64.6% (40.9% + 23.7%) of the observed peaks in terms of timing, whereas "SD-Box All" contained 77.5% (62.4% + 15.1%). Since the scores of "SD-Box Single" are widespread, it implies that the performance may experience wild swings in each run. In terms of stage and timing 62.4% of the scores of "SD-Box All" in comparison with 40.9% of "SD-Box Single" are less than one. It indicates that due to larger sample size the spread of "SD-Box All" is more appropriate to cover the uncertainty in the prediction. This just confirms the stability of the performance of "SD-Box All" and steady performance provides decision makers a basis to respond to a potential disaster more consistently.





Figure 11. The spread-skill scores for (**a**) peak stage and (**b**) peak timing of the single ("SD-Box Single") and accumulating ("SD-Box All") methods at the Zhongshan Bridge during the four selected typhoon events. The inverted triangles indicate the time of occurrence of the observed peak stage. The results show that the performance of the "SD-Box All" method (solid circles) was more stable than that of the "SD-Box Single" method (open circles) in terms of both stage and timing.



Figure 12. Comparison of the spread-skill scores obtained for "SD-Box Single" and "SD-Box All". The results show that the "SD-Box All" approach provides comparatively consistent forecasts in terms of variation of the score and it serves as a good basis for decision makers to respond the disasters.

5. Conclusions

This study proposed a HEPS that employs NWP models to perform rainfall forecasts and employs hydrologic models, including rainfall-runoff, river-routing, and combined storm surge/global tide models to produce ensemble flood forecasts during typhoon events. The communication of ensemble forecasts is critical for helping end-users respond. A visualization approach called "SD-Box" was also developed to support the interpretation of ensemble forecast results for operational purposes. A total of 93 forecast runs from four typhoon events during the period 2012–2015 and observations collected in the Yilan Experimental Watershed were used to evaluate the performance of these techniques. The results showed that the proposed HEPS is able to provide flood forecasts during the selected typhoon events. The peak flow information, such as peak stage and timing, are usually the most important for decision-making. For the purposes of simplicity and efficiency of operation, the "SD-Box" visualization approach considers the mean and the standard deviation to define the size of the box and focuses on the peak flow, capturing more of the observed peaks during typhoon events.

This study used a spread-skill score based on peak stage and peak timing accuracy to evaluate the similarity of the ensemble averages and observations and to verify the peak stage forecasts. Further evaluation measures related to flow volumes can be found in the literature (e.g., series-distance methods [33,34]). The average spread-skill scores of the proposed HEPS for the selected events were 1.13 and 1.00 in terms of peak stage and peak timing, respectively. Scores of close to one indicated that the spread of the ensemble forecasts was large enough to contain the prediction uncertainty. As an

input to the HEPS, the rainfall forecasts from NWP models exhibited a significant impact on the performance. The worst performance was found during Typhoon Dujuan because of its large track forecast errors. As a result, the spread-skill scores were above 2. However, details of the performance of the NWP models are beyond the scope of this study. Overall, the average score was close to one and satisfied the statement "One of the main objectives of ensemble flood forecasts is the representation of the full spectrum of forecast uncertainty or predictability in form of different hydrological responses to the input of the various members obtained from an atmospheric EPS" made by Zappa et al. [10]. The supporting visualization approach can therefore decrease the complexity of the HEPS forecast information and improve the efficiency of end-user's decision-making.

This study also proposed a method that enables increasing the sample size, leading to statistically significant results. This method involves including current and previously available forecasts in the calculation process. For example, the proposed HEPS generated 11 available ensemble members at each forecast during Typhoon Dujuan. By including all of the present and previously available forecasts (the "SD-Box All" method), the sample size increased to 22 for the second forecast, 33 for the third forecast, and so on. The results showed that the "SD-Box All" made more consistent predictions by including all available forecasts in the calculation process. The descriptive statistics, such as the quartile deviation and the standard deviation, can be more meaningful after expanding sample size. As a result, the rectangles defined by the "SD-Box All" method contained 57.8% of the observed peaks in stage and timing. Coughlan de Perez et al. [35] suggested that a HEPS that produces a false alarm rate below 50% is tolerable for decision makers in terms of the economic and practical consequences of taking action. However, this study assumed that the forecast performance of the proposed HEPS is independent of the length of the lead time and conducted an experiment to prove it. Other studies, such as that of Zappa et al. [10], have claimed that the most accurate forecasts were obtained for lead times of two or more days. Such statements imply that the performance of HEPSs do not improve with shorter lead times or are independent of lead time, and Yang et al. [32] found that the best performance is obtained before a typhoon makes landfall. This assumption is still susceptible to the topography of the applied area and the type of extreme event being considered. Further investigation of various conditions must be performed before firm conclusions can be drawn. Regardless, the proposed HEPS and the modified visualization approach have been shown to produce convincing peak stage and peak timing forecasts for operational purposes during a typhoon. In addition, the modified visualization approach is based on the well calibrated HEPS and is applicable only for typhoon events. Therefore, the users should properly apply it under these prerequisites.

Author Contributions: Data curation, C.-H.C.; formal analysis, S.-C.Y. and Y.-C.C.; funding acquisition, T.-H.Y.; methodology, Y.-C.C., M.-Y.L., J.-Y.H., and K.T.L.; validation, Y.-C.C.; visualization, S.-C.Y.; writing—original draft, S.-C.Y. and Y.-C.C.; writing—review and editing, T.-H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Ministry of Science and Technology, R.O.C., under grant MOST 105-3011-F-492-009.

Acknowledgments: The authors thank the Water Resources Agency of Taiwan for providing the hydrological observations from the rainfall gauges and water-level stations in the Yilan River Basin. Thanks are also due to the Taiwan Typhoon and Flood Research Institute and the National Applied Research Laboratories for providing the results from the Taiwan Cooperative Precipitation Ensemble Forecast Experiment and historical records from the Yilan Experimental Watershed.

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

References

- 1. Cloke, H.L.; Pappenberger, F. Ensemble flood forecasting: A review. J. Hydrol. 2009, 375, 613–626. [CrossRef]
- Cuo, L.; Pagano, T.; Wang, Q. A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. *J. Hydrometeorol.* 2011, *12*, 713–728. [CrossRef]

- 3. Pappenberger, F.; Beven, K.J.; Hunter, N.M.; Bates, P.D.; Gouweleeuw, B.T.; Thielen, J.; De Roo, A.P.J. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrol. Earth Syst. Sci.* **2005**, *9*, 381–393. [CrossRef]
- 4. Balint, G.; Csik, A.; Bartha, P.; Gauzer, B.; Bonta, I. Application of Meterological Ensembles for Danube Flood Forecasting and Warning. In *Gheorghe Stancalie and Gabor Balint, Transboundary Floods: Reducing Risks through Flood Management*; Marsalek, J., Ed.; Springer: Berlin/Heidelberg, Germany, 2006.
- Davolio, S.; Miglietta, M.M.; Diomede, T.; Marsigli, C.; Morgillo, A.; Moscatello, A. A meteo-hydrological prediction system based on a multi-model approach for precipitation forecasting. *Nat. Hazards Earth Syst. Sci.* 2008, *8*, 143–159. [CrossRef]
- Todini, E. Predictive Uncertainty Assessment in Real Time Flood Forecasting. In Uncertainties in Environmental Modelling and Consequences for Policy Making; Baveye, P.C., Laba, M., Mysiak, J., Eds.; Springer: Berlin/ Heidelberg, Germany, 2009; pp. 205–228.
- 7. Giordani, A.; Zappa, M.; Rotach, M. Estimating Ensemble Flood Forecasts' Uncertainty: A Novel "Peak-Box" Approach for Detecting Multiple Peak-Flow Events. *Atmosphere* **2019**, *11*, 2. [CrossRef]
- 8. Abebe, A.J.; Price, R.K. Decision support system for urban flood management. *J. Hydroinformatics* **2005**, *7*, 3–15. [CrossRef]
- 9. Pappenberger, F.; Stephens, L.; Thielen, J.; Salamon, P.; DeMeritt, D.; Van Andel, S.J.; Wetterhall, F.; Alfieri, L. Visualizing probabilistic flood forecast information: Expert preferences and perceptions of best practice in uncertainty communication. *Hydrol. Process.* **2012**, *27*, 132–146. [CrossRef]
- 10. Zappa, M.; Fundel, F.; Jaun, S. A 'Peak-Box' approach for supporting interpretation and verification of operational ensemble peak-flow forecasts. *Hydrol. Process.* **2012**, *27*, 117–131. [CrossRef]
- Pagano, T.C.; Wood, A.W.; Ramos, M.-H.; Cloke, H.L.; Pappenberger, F.; Clark, M.; Cranston, M.; Kavetski, D.; Mathevet, T.; Sorooshian, S.; et al. Challenges of Operational River Forecasting. *J. Hydrometeorol.* 2014, 15, 1692–1707. [CrossRef]
- 12. Olsson, J.; Pers, C.; Bengtsson, L.; Pechlivanidis, I.; Berg, P.; Körnich, H. Distance-dependent depth-duration analysis in high-resolution hydro-meteorological ensemble forecasting: A case study in Malmö City, Sweden. *Environ. Model. Softw.* **2017**, *93*, 381–397. [CrossRef]
- 13. Wu, C.-C.; Kuo, Y.-H. Typhoons Affecting Taiwan: Current Understanding and Future Challenges. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 67–80. [CrossRef]
- Hsiao, L.-F.; Yang, M.-J.; Lee, C.-S.; Kuo, H.-C.; Shih, D.-S.; Tsai, C.-C.; Wang, C.-J.; Chang, L.-Y.; Chen, D.Y.-C.; Feng, L.; et al. Ensemble forecasting of typhoon rainfall and floods over a mountainous watershed in Taiwan. *J. Hydrol.* 2013, 506, 55–68. [CrossRef]
- 15. Huang, J.-C.; Yu, C.-K.; Lee, J.-Y.; Cheng, L.-W.; Lee, T.-Y.; Kao, S.-J. Linking typhoon tracks and spatial rainfall patterns for improving flood lead time predictions over a mesoscale mountainous watershed. *Water Resour. Res.* **2012**, *48*. [CrossRef]
- 16. Lee, K.T.; Yen, B.C. Geomorphology and Kinematic-Wave–Based Hydrograph Derivation. *J. Hydraul. Eng.* **1997**, 123, 73–80. [CrossRef]
- 17. Lee, T.K.; Chang, C.H.; Yang, M.S.; Yu, W.S. Reservoir Attenuation of Floods from Ungauged Watersheds. *Hydrol. Sci. J.* **2001**, *46*, 349–362. [CrossRef]
- 18. Lee, T.K.; Chung, Y.R.; Lau, C.C.; Meng, C.C.; Chiang, S. A Windows-Based Inquiry System for Design Discharge Based on Geomorphic Runoff Modeling. *Comput. Geosci.* **2006**, *32*, 203–211. [CrossRef]
- 19. Lee, K.T.; Chen, N.-C.; Gartsman, B.I. Impact of stream network structure on the transition break of peak flows. *J. Hydrol.* **2009**, *367*, 283–292. [CrossRef]
- 20. Thiessen, A.H. Precipitation averages for large areas. Mon. Weather Rev. 1911, 39, 1082–1084. [CrossRef]
- 21. Blumberg, F.A.; Mellor, G.L. A Description of a Three-Dimensional Coastal Ocean Circulation Model, in Three-Dimensional Coastal Ocean Models; American Geophysical Union: Washington, DC, USA, 1987.
- 22. Ou, S.; Liu, J.; Tsai, C.; Hsu, T. Numerical Studies of Typhoon-Induced Storm Surge Using Pom and Finite Element Depth-Averaged Model in Taiwan. In Proceedings of the Chinese-German Joint Symposium on Hydraulic and Ocean Engineering, Darmstadt, Germany, 24–30 August 2008.
- 23. Chiou, M.D. Characteristic and Numerical Simulation of Astronomic Tide and Storm Surge in Taiwan Water. Ph.D. Thesis, National Cheng Kung University, Tainan, Taiwan, 2010.

- 24. Yeh, T.G.; Cheng, H.P.; Cheng, J.R.; Lin, J.H. A Numerical Model to Simulate Flow and Contaminant and Sediment 1 Transport in Watershed Systems (Wash12d). In *Technical Rep. Chl-98-15, Waterways Experiment Station*; U.S. Army Corps of Engineers: Vicksburg, MS, USA, 1998.
- 25. Shih, D.-S.; Liau, J.-M.; Yeh, G.-T. (George) Model Assessments of Precipitation with a Unified Regional Circulation Rainfall and Hydrological Watershed Model. *J. Hydrol. Eng.* **2012**, *17*, 43–54. [CrossRef]
- Yeh, T.G.; Huang, G.B.; Zhang, F.; Cheng, H.P.; Lin, H.C. Wash123d: A Numerical Model of Flow, Thermal Transport, and Salinity, Sediment, and Water Quality Transport in Watershed Systems of 1-D Stream-River Network, 2-D Overland Regime, and 3-D Subsurface Media; Technical Report Submitted to US EPA; Department of Civil and Environmental Engineering, University of Central Florida: Orlando, FL, USA, 2005.
- 27. Yeh, G.-T.; Shih, D.-S.; Cheng, J.-R.C. An integrated media, integrated processes watershed model. *Comput. Fluids* **2011**, *45*, 2–13. [CrossRef]
- 28. Shih, D.-S.; Chen, C.-H.; Yeh, G.-T. Improving our understanding of flood forecasting using earlier hydro-meteorological intelligence. *J. Hydrol.* **2014**, *512*, 470–481. [CrossRef]
- 29. Levine, R.A.; Wilks, D.S. Statistical Methods in the Atmospheric Sciences. *J. Am. Stat. Assoc.* 2000, *95*, 344. [CrossRef]
- 30. Yang, T.-H.; Tsai, C.-C. Using numerical weather model outputs to forecast wind gusts during typhoons. *J. Wind. Eng. Ind. Aerodyn.* **2019**, *188*, 247–259. [CrossRef]
- 31. Yang, S.-C.; Yang, T.-H. Uncertainty Assessment: Reservoir Inflow Forecasting with Ensemble Precipitation Forecasts and HEC-HMS. *Adv. Meteorol.* **2014**, 2014, 1–11. [CrossRef]
- 32. Yang, T.-H.; Hwang, G.-D.; Tsai, C.-C.; Ho, J.-Y. Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 4731–4745. [CrossRef]
- 33. Ehret, U.; Zehe, E. Series distance–an intuitive metric to quantify hydrograph similarity in terms of occurrence, amplitude and timing of hydrological events. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 877–896. [CrossRef]
- 34. Seibert, S.P.; Ehret, U.; Zehe, E. Disentangling timing and amplitude errors in streamflow simulations. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 3745–3763. [CrossRef]
- 35. De Perez, E.C.; Hurk, B.V.D.; Van Aalst, M.K.; Amuron, I.; Bamanya, D.; Hauser, T.; Jongma, B.; Lopez, A.; Mason, S.; De Suarez, J.M.; et al. Action-based flood forecasting for triggering humanitarian action. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 3549–3560. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).