Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Development of a hydrological ensemble prediction system and a visualization approach for improved interpretation during typhoon events

Sheng-Chi Yang[1], Tsun-Hua Yang[1], Ya-Chi Chang[1,*], Cheng-Hsin Chen[1], Mei-Ying Lin[1], Jui-Yi Ho[1] and Kwan-Tun Lee[1,2]

[1]*Taiwan Typhoon and Flood Research Institute (TTFRI), National Applied Research Laboratories (NARLabs), Taipei, Taiwan*

[2]*Department of River and Harbor Engineering, National Taiwan Ocean University, Keelung, Taiwan*

[*]*Correspondence to: 11 F, No. 97, Sec. 1, Roosevelt Rd., Zhongzheng Dist., Taipei City 10093, Taiwan (R.O.C.)*

E-mail: rachel.ev91@gmail.com

ABSTRACT

Typhoons are accompanied by heavy rainfall and cause loss of life and property. Hydrological ensemble prediction systems can provide decision makers with hydrological information, such as peak stage and peak time, with some lead time. This information assists decision makers in taking the necessary measures to prevent and mitigate disasters. This study proposes a hydrological ensemble prediction system that includes numerical weather models that perform rainfall forecasts and hydrologic models that produce assessments of surface runoff and the associated flooding. However, the spatiotemporal uncertainty associated with the numerical models and the difficulty in interpreting the model results hinder effective decision making during emergency response situations. Thus, this study also presents an extension of the 'Peak-

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Box' visualization methodology that assists in interpreting the forecast results for operational purposes. A small watershed with area of 100 km^2 and four typhoons that occurred from 2012 to 2015 were selected to evaluate the performance of these tools. The results showed that the modified 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 new approach includes all available forecasts to increase the sample size. The capture rate is greater than 50%, which is considered practical for decision makers. The proposed system and the modified visualization approach have demonstrated their potential for both decreasing the uncertainty of numerical rainfall forecasts and improving the performance of flood forecasts.

KEY WORDS Hydrological ensemble prediction system; peak flow; decision support; visualization.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

1. INTRODUCTION

Numerical weather prediction (NWP) models generate different precipitation forecasts for specified locations and times due to the incompleteness of the input observations, the approximate nature of the forecast models and their parameterizations, and the random errors that result from perturbing the initial atmospheric conditions (Palmer, 2001; Hostache et al., 2011). Ensemble prediction systems (EPSs), which consist of an adequate number of equiprobable NWP models, have been established to provide probabilistic precipitation forecasts instead of a single deterministic forecast (Cloke and Pappenberger, 2009). 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 (Demeritt et al., 2007; Cuo et al., 2011).

A hydrological ensemble prediction system (HEPS) is an integrated system that couples an EPS with catchment-scale hydrological models to provide flood forecasts with sufficient lead time. The importance of such systems in disaster mitigation, water resource management, and hydropower dam and lake operation is growing (Pappenberger et al., 2005; Cloke and Pappenberger, 2009; Zappa et al., 2010, 2013; Yang and Yang, 2014). However, uncertainties stemming from factors including boundary conditions, initial conditions, and model parameter values affect the forecast accuracy of these systems. The precipitation forecasts of NWP models dominate the overall uncertainty of these systems (Zappa et al., 2011; Rossa et al., 2011). It is necessary to develop guidelines and tools for communicating the uncertainties associated with complex HEPSs (e.g., Jaun et al., 2008; Thielen et al., 2009; Bartholmes et al., 2009; Todini, 2009; Bruen et al., 2010; Renard et al., 2010; Thirel et al., 2010; Zappa et al., 2010, 2013; Frick and Hegg, 2011; Pappenberger et al., 2011a, 2011b; Fundel and Zappa, 2011; Pappenberger et al., 2013).

Effective communication of ensemble forecasts means that clear expression of the uncertainties associated with HEPS is important so that end-users can easily respond to the information provided during operations (Demeritt et al., 2010; Ramos et al., 2010; Pappenberger et al., 2013; Zappa et al., 2013; Pagano et al., 2014). Pagano et al. (2014) noted that defining effective methods for the communication of ensemble forecasts is a challenge for future operational river forecasting and represents a future research opportunity. Pappenberger et al. (2013) 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 information provided. At present, HEPSs still rely on conventional visualization techniques, such as ‘spaghetti diagrams’ or box plots, to display the distributions of forecast results. Pappenberger et al. (2013) 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, worst/best scenario, etc. Zappa et al. (2013) proposed the 'Peak-Box' visualization approach to support the interpretation and verification of HEPS results. This approach has been used to obtain quantitative and qualitative insights, such as the timing, water level, and discharge associated with peak flow. This information is crucial for end-users and decision makers. Zappa et al. (2013) applied an operational HEPS, namely, the IFKIS-HYDRO hydrological nowcasting system, to five different basins in Switzerland to evaluate the performance of the 'Peak-Box' methodology. The sizes of the basins ranged from 186 km^2 to 1696 km^2. The study found that, of 485 operational forecasts performed from June 2007 through November 2008, 30% to 55% of the observed peaks fell outside the 'Peak-Box'.

Typhoons are common natural events that cause severe damage in countries at the edge of the northwestern Pacific Ocean, such as Japan, the Philippines, and Taiwan. For example, based on records covering 1958 to 2010, an average of 3.4 typhoons affect Taiwan annually, and these events cause an annual average loss of more than 500 million U.S. dollars (Li et al., 2004). Typhoon-related flood events cause these losses. If they provide early warnings with sufficient lead time, flood forecasts from a HEPS can help authorities prepare disaster prevention and mitigation measures. A customized visualization method for typhoons is also necessary to make the ensemble flood forecasts generated by HEPS meaningful for emergency responders. Therefore, this

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

study presents a HEPS that can provide ensemble flood forecasts during typhoon events and proposes a customized visualization approach especially for typhoons to simplify the forecast information. This approach is an extension of the one presented by Zappa et al. (2013); it has been modified to increase the percentage of observed peaks that fall within the predicted range during typhoon events. The remainder of this paper is organized as follows. Section 2 includes the details of the proposed HEPS. Section 3 briefly describes the study area and typhoon events used in the study. Section 4 compares the original 'Peak-Box' approach with the proposed extended version. Finally, Sect. 5 and 6 present the results, discussion, and conclusions.

2. SETUP OF THE HYDROLOGICAL ENSEMBLE PREDICTION SYSTEM

This study proposes a HEPS that integrates various models. These models include NWP models that provide ensemble precipitation forecasts, a rainfall-runoff model that generates upstream boundary conditions, a storm surge model that generates downstream boundary conditions, and a flood routing model that simulates river flows. The data processing is shown in Figure 1. The HEPS produces ensemble flood forecasts with a 72-hour lead time four times a day. The models used in the HEPS are described in the following subsections.

2.1 Ensemble precipitation forecasts

The Taiwan Cooperative Precipitation Ensemble Forecast Experiment (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 Central Weather Bureau (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-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 (HWRF). It also considers different setups in terms of the model initial conditions, data assimilation processes and 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 hours. Further information can be found in Hsiao et al. (2013). A typhoon's average impact duration is 73.68 hours (Huang et al., 2012) and the average lag between observed peak precipitation and flooding in Taiwan is between 2 and 10 hours (Jang et al., 2012). This study focuses on a one-way

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

coupling in which TAPEX provides rainfall forecast to the rainfall-runoff model; feedbacks from the rainfall-runoff model to TAPEX are not considered.

2.2 Rainfall-runoff model

The HEPS uses the surface runoff forecast generated by a kinematic-wave-based geomorphologic instantaneous unit hydrograph model (the KW-GIUH model) as its upstream boundary condition. The KW-GIUH model, which was developed by Lee and Yen (1997), 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 (Lee et al., 2001; 2006).

2.3 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 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 (1987), is a three-dimensional, nonlinear, primitive equation finite difference ocean model. It has been applied to simulate a wide range of ocean problems, including coastal storm surge in Taiwan (Ou et al., 2008; Chiou, 2010). In this study, the TAPEX

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

model provides ensemble pressure field and wind field forecasts to POM and the TPXO6.2 model and obtains tidal level predictions. As with TAPEX, it generates four runs a day, and each run has a 72-hour lead time.

2.4 Flood 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) was developed by Yeh et al. (1998) to simulate one-dimensional channel networks, two-dimensional overland flow, and three-dimensional variably saturated subsurface flow. 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 (e.g., Yeh et al., 2006; Yeh and Shih., 2011; Shih et al., 2012; Hsiao et al., 2013). The HEPS uses the one-dimensional channel model of WASH123D as its flood routing model to simulate water stages in rivers.

3. STUDY AREA AND TYPHOON EVENTS

3.1 Study area

This study selected the Yilan River in northeastern Taiwan as the study area (Figure 2). The river flows through the city of Yilan and has a main stream length of approximately 24.4 km and a watershed area of 149.06 km^2. It has four main tributaries,

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

which are the Wushi River, the Dahu River, the Dajiao River and the Xiaojiao River. The Water Resource Agency (WRA) and TTFRI have selected this river as one of two watersheds where long-term monitoring experiments are being carried out (the other is the Dianbao Creek basin in southwestern Taiwan). The purpose of the experimental watersheds is to generate long-term and high-density hydrological monitoring data that can be used for scientific studies, including the development of hydrological and hydraulic models and the study of environmental changes. In total, 11 rainfall gauging stations, 16 water-stage gauging stations, five river-velocity gauging stations, and 36 inundation-depth gauging stations have been installed in the Yilan River Basin. Figure 2 shows the locations of the water-stage and rainfall gauging stations that collected the data that we used in this study. The monitoring data have been carefully collected and processed. For full information and to download the available data, please refer to the official website (http://wraew.ttfri.narl.org.tw/index.php).

TAPEX provides 72-hour rainfall forecasts for five rainfall gauges in the upstream portion of the Yilan River Basin. The KW-GIUH model calculates the surface runoff and estimates river flow at the Hsincheng and Yuanshan Bridges. This study uses the POM and TPXO6.2 models to forecast the tides at Suao and to estimate the water stages at the Kemalan Bridge. WASH123D then generates ensemble flow forecasts using flows at the bridges mentioned above as the upstream boundary condition and the water stage at the Kemalan Bridge as the downstream boundary condition. The detailed locations of these places are shown in Figure 2.

3.2 Typhoon events

Figure 3 shows the tracks of the different typhoons that have affected Taiwan, according to historical records (Huang et al., 2012). Of the ten categories, Type-2 and Type-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 hours was observed at rainfall gauging station C1U610 (shown in Figure 2) during Typhoon Soulik. Table 1 shows all of the typhoons that invaded Taiwan from 2012 through 2015. Five of these events are Type-2 and Type-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 the HEPS and test its performance. Typhoon Matmo, a Type-3 typhoon that occurred in 2014, was not included due to its weak intensity. This study used historical observations of rainfall, river stage, and tide to validate the parameters in the proposed HEPS.

4. A VISUALIZATION APPROACH FOR SUPPORTING THE INTERPRETATION OF OPERATIONAL ENSEMBLE PEAK-

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

FLOW FORECASTS DURING TYPHOON EVENTS

This study modified the 'Peak-Box' approach originally proposed by Zappa et al. (2013) to provide better communication of HEPS forecasts during typhoon events. Figure 4 compares the two approaches, and the modifications are described in detail below. The purpose of the modifications is to develop a visualization approach that simplifies the ensemble flow forecast information for use in formulating emergency responses during typhoon events.

a. **Remove the horizontal and the vertical lines.** The horizontal and vertical lines that indicate the medians of ensemble forecasts in the original 'Peak-Box' approach are removed to prevent some information from being misused. Although uncertainties exist in the HEPS, Pappenberger et al. (2013) noted a considerable desire on the part of end-users to reduce probabilistic forecasts to deterministic actions. The two lines may lead end-users to believe that the information provided represents a single deterministic forecast, rather than a probabilistic one.

b. **Remove the outer rectangle.** In the original 'Peak-Box' approach, two rectangles are displayed. The outer rectangle is the 'Peak-Box,' which highlights all possibilities from the ensemble forecast, and the inner rectangle is the 'IQR-Box' that emphasizes the 25th and 75th percentiles of the peak times and peak discharges of the ensemble forecast. Zappa et al. (2013) argued that the outer rectangle

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

provides the forecaster with additional information. However, this argument does not hold during typhoons, when the availability of too much data may obscure critical information. Therefore, only one rectangle is shown in the study. This rectangle indicates where the observed peak stage is likely to occur.

c. **Use the mean and the standard deviation to define the rectangle.** This study defines an 'SD-Box' that uses the mean (μ) and the standard deviation (σ), instead of the first and third quartiles, to define the enveloping rectangle. As shown in the right panel of Figure 4, 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. In principle, the 'IQR-Box' should contain 25% (50% of the peak discharge times and 50% of the peak times) of all forecasts. In practice, it contained from 12.5% to 37.5%, due to the distribution of ensemble members (Zappa et al., 2013). Using the mean and the standard deviation (the 'SD-Box') results in a larger area, includes 46.60% of the ensemble forecasts (68.27% of peak water level times and 68.27% of the peak times) and has a greater chance of including the observed peaks.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

d. **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. Adding extra ensemble members to produce more forecasts consumes computer resources. Yang et al. (2016) showed that the performance of NWP models is independent of the length of the lead time during typhoon events. Therefore, in order to expand the sample size, 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 provide ensemble flow forecasts. As shown in the right panel of Figure 4, the green area illustrates the 'SD-Box'. The black and gray solid dots represent the current and previous peak-flow forecasts, respectively.

5. RESULTS AND DISCUSSION

5.1 Performance evaluation criteria

This study applied two performance measures, the root mean square error (RMSE) and the skill-spread ratio, to evaluate the proposed HEPS performance. For a well-designed HEPS, the spread of ensemble forecasts will be large enough to cover the prediction uncertainty. This statement implies that the spread should be the same as or larger than the RMSE. The RMSE, which is commonly referred to as skill, measures the difference between the observations and the ensemble mean without considering

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

the direction. The closer the RMSE is to zero, the better the ensemble mean is as a forecast. The RMSE is defined as follows:

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

$$\mu = \frac{1}{m}\sum_{i=1}^{m} P_{peak,i} \tag{2}$$

$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(P_{peak,i} - \mu\right)^2} \tag{3}$$

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

The skill-spread score (hereinafter referred to as the score), which ranges from zero to infinity, is the ratio of the standard deviation of the ensemble peak-flow forecasts to the RMSE (Wilks, 2006). Scores less than one mean that the spread of the ensemble forecasts is large enough to cover the prediction uncertainty. It is defined as follows:

$$\mathrm{Score} = \frac{RMSE}{\sigma} \tag{4}$$

5.2 Model calibration and validation

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

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

flow (n_c). The proposed HEPS used data from five rainfall gauges, including LTGX, YSGZ, C1U610, C0U520 and C1U630 (see Figure 2 for locations), and the Thiessen polygon method (Thiessen, 1911) to estimate the hourly spatial-average rainfall intensities in order to provide rainfall input data to the KW-GIUH model. The topographic data used in KW-GIUH are contained within a digital elevation model with a resolution of 5 m obtained using aerial photographs. Kuo et al. (2016) used in situ observations of flow discharges made at the Hsincheng and Yuanshan Bridges during Typhoons Saola, Soulik, and Soudelor to calibrate the parameters of the KW-GIUH model. Figure 5 shows that the percent errors in the peak discharges of the selected typhoons were 4.59%, 2.07%, and -5.89% at the Hsincheng Bridge, and 14.88%, 5.28%, and -3.05% at the Yuanshan Bridge, respectively. All of the errors in the peak times were less than one hour. The results show that the KW-GIUH model is capable of providing confident predictions for peak time, as well as peak discharge.

The WASH123D model 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. Field measurements at the Hsincheng and Yuanshan Bridges from Kuo et al. (2016) and observed water stages at the Kemalan Bridge were used as the upstream and downstream boundary conditions, respectively. Field hourly records of water-stage at the Zhongshan, Leawood, and Jhuangwei Bridges were used to calibrate the value of Manning's roughness coefficient (*n*) in the WASH123D model and to validate the performance of the model. Figure 6 shows that the percent errors in the peak stage 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 data collection during Typhoon Soudelor. Nevertheless, all of the errors in the peak times were less than one hour. The results show that WASH123D is capable of providing confident predictions of peak times, as well as peak stages.

5.3 Comparison of enveloping rectangles defined using the 'SD-Box' and the 'IQR-Box' methods for supporting the interpretation of ensemble peak-flow results

The proposed HEPS initiates when CWB issues a sea warning and ends 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 compares the forecast peak stages and peak times between the 'SD-Box' and 'IQR-Box' approaches at the Zhongshan, Leawood, and Zhuangwei Bridges. Scores were not calculated for the Leawood Bridge during Typhoon Soudelor due to the lack of complete observations. The scores that are less than one in the table are highlighted. These values indicate that the spread of the ensemble members is large enough to contain the prediction uncertainty. The rectangles defined using the ‘IQR-Box’ method contain 33.3% (31/93) and 52.6% (49/93) of the observed peaks in stage and timing, respectively. Using the ‘SD-Box’ improves the capture rate to 51.6% (48/93) and 64.5% (60/93) for stage and timing, respectively. Among all of the forecasts, there is only one forecast for which the ‘IQR-Box’ score is less than one, and the score of the ‘SD-Box’ is not. This situation occurs at the Zhuangwei Bridge during Typhoon Soudelor. However, the score for the ‘SD-Box’ method is still very close to one (1.01), which means that it nearly captures the observed peak. Overall, the ‘SD-Box’ method yielded average scores of 1.18 for the peak stages and 1.08 for the peak times. In comparison with the ‘IQR-Box’ method, which yielded scores of 2.06 for the peak stages and 2.06 for the peak times, the results show that the enveloping rectangles defined using the ‘SD-Box’ method are more reliable during typhoon events.

5.4 Including 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 for the proposed

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

HEPS during operation. Thus, the descriptive statistics were calculated using insufficient sample sizes (less than 30). The same issue exists in other studies that employ HEPSs (e.g., Yang and Yang, 2014; Zappa et al., 2013). It is difficult to increase the number of ensemble members used in HEPSs, due to the limited computational resources that are available. Therefore, this study proposes a method for including present and previous forecasts in order to expand the sample size during the estimation process.

It must be shown that the forecast performance is independent of time before all available forecasts can be included in the estimation process. The time of concentration of the peak flow at the Zhongshan Bridge is approximately 4 hours. This study calculated the error in the maximum 4-hour rainfall between the average forecasts and the average observations at the watershed upstream of the Zhongshan Bridge. Figure 7 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 1^{st} and 5^{th} forecasts, respectively. However, the 6^{th} forecast was better than the 5^{th}, which implies that there is no trend in the cascading forecasting process. Based on these results, this study

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Hydrology and Earth System Sciences Open Access Discussions EGU

assumed that the performance of the HEPS is independent of lead time during typhoon events. Therefore, it is reasonable to include all available forecasts during an event to expand the sample size.

Figure 8 illustrates the comparisons between using the 'SD-Box' method with one forecast and using the 'SD-Box' method including all available forecasts (hereinafter indicated as 'SD-Box Single' and 'SD-Box All') at the 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 scores for stage during Typhoon Soudelor ranged from 0 to 5 when the 'SD-Box Single' method was used, but they were below or close to 1 with 'SD-Box All'. The results showed that the inclusion of all available forecasts in the calculation process decreased the variation among the forecasts; in other words, the uncertainty of the forecasts decreased. Figure 9 illustrates the scores of all of the forecasts for the different typhoon events. The 'SD-Box Single' contained 47.1% of the observed peaks in terms of stage (37.3% + 9.8%), whereas 'SD-Box All' contained 63.7% (57.8% + 5.9%) of the observed peaks. Furthermore, the 'SD-Box Single' contained 58.9% (37.3% + 21.6%) of the observed peaks in terms of timing, whereas 'SD-Box All' contained 71.5% (57.8% + 13.7%). The results show that the 'SD-Box All' method can capture more of the observed peaks in terms of both stage

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

and timing. In particular, 'SD-Box All' improved the forecast performance and increased the capture rate from 37.3% to 57.8% for both stage and timing.

6. CONCLUSIONS

This study proposed a HEPS that employs NWP models to perform rainfall forecasts and hydrologic models to produce ensemble flood forecasts during typhoon events. Because the communication of ensemble forecasts is critical for helping end-users to respond, a modified version of the 'Peak-Box' visualization method, which was originally described by Zappa et al. (2013), was also proposed to support the interpretation of ensemble forecast results for operational purposes. 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. A total of 93 forecasts and two performance measures were considered. The results showed that the proposed HEPS is able to provide flood forecasts during the selected typhoon events. In addition, the 'SD-Box' visualization approach, which considers the mean and the standard deviation instead of the 25th and 75th percentiles, captured more of the observed peaks during typhoon events. The average skill-spread scores of the 'SD-Box' method for the selected events were 1.18 and 1.08 in terms of stage and timing, respectively. These results represent a significant improvement over the original 'Peak-Box' method, which resulted in scores of 2.06 for both peak stage and peak

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

timing. Scores of less than one indicate that the spread of the ensemble forecasts is large enough to contain the prediction uncertainty. Since the average score achieved by the 'SD-Box' method was close to one, it has been shown to be more reliable than the original 'Peak-Box' method during typhoon events. The results satisfy the statement "One of the main objectives of ensemble flood forecasts is the representation of the full spectrum of forecast uncertainty and/or predictability in [the] form of different hydrological responses to the input of the various members obtained from an atmospheric EPS" made by Zappa et al. (2013).

Descriptive statistics, such as the quartile deviation and the standard deviation, are susceptible to outliers when calculated using an insufficient number of observations. Adding more ensemble members is expensive in terms of computer resources. This study proposed a method that enables increasing the sample size, leading to statistically significant results. This method involves including present and previous 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 previous 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. This result can be explained by the inclusion of all available forecasts in the calculation process decreasing the uncertainty of the forecasts. 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. (2016) 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. (2013), 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. (2016) 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.

AUTHOR CONTRIBUTION

Ya-Chi Chang, Mei-Ying Lin, Jui-Yi Ho calibrated and verified the parameters of WASH123D, POM and KW-GIUH models. Cheng-Hsin Chen dealt with the data processing of the models and performed the simulations. Sheng-Chi Yang and Ya-Chi

Chang analyzed the results of HEPS and developed a new approach for improved interpretation during typhoon events. Sheng-Chi Yang, Tsun-Hua Yang, Ya-Chi Chang and Kwan-Tun Lee prepared the manuscript with contributions from all co-authors.

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

REFERENCES

Bartholmes, J. C., Thielen, J., Ramos, M. H., and Gentilini, S.: The european flood alert system EFAS-Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts, Hydrology and Earth System Sciences, 13, 141-153, 2009.

Blumberg, A. F. and Mellor, G. L.: A Description of a Three-Dimensional Coastal Ocean Circulation Model, in Three-Dimensional Coastal Ocean Models, American Geophysical Union, Washington, D.C., 1987.

Bruen, M., Krahe, P., Zappa, M., Olsson, J., Vehvilainen, B., Kok, K., and Daamen, K.: Visualizing flood forecasting uncertainty: some current European EPS platforms-COST731 working group 3, Atmospheric Science Letters, 11, 92-99, 2010.

Chiou, M. D.: Characteristic and numerical simulation of astronomic tide and storm surge in Taiwan water, Ph. D., Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Tainan, Taiwan, 135 pp., 2010.

Cloke, H. L. and Pappenberger, F.: Ensemble flood forecasting: A review, J. Hydrol., 375, 613-626, 2009.

Coughlan de Perez, E., van den Hurk, B., van Aalst, M. K., Amuron, I., Bamanya, D., Hauser, T., Jongman, B., Lopez, A., Mason, S., Mendler de Suarez, J.,

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Pappenberger, F., Rueth, A., Stephens, E., Suarez, P., Wagemaker, J., and Zsoter, E.: Action-based flood forecasting for triggering humanitarian action, Hydrology and Earth System Sciences, 20, 3549-3560, 2016.

Cuo, L., Pagano, T.C. and Wang, Q.J.: A review of quantitative precipitation forecasts and their use in short-to medium-range streamflow forecasting, 12, 713-728, 2011.

Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J., and Ramos, M.-H.: Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting, Environmental Hazards, 7, 115-127, 2007.

Demeritt, D., Nobert, S., Cloke, H., and Pappenberger, F.: Challenges in communicating and using ensembles in operational flood forecasting, Meteorological applications, 17, 209-222, 2010.

Frick, J. and Hegg, C.: Can end-users' flood management decision making be improved by information about forecast uncertainty?, Atmospheric Research, 100, 296-303, 2011.

Fundel, F. and Zappa, M.: Hydrological ensemble forecasting in mesoscale catchments: Sensitivity to initial conditions and value of reforecasts, Water Resources Research, 47, W09520, 2011.

Hostache, R., Matgen, P., Montanari, A., Montanari, M., Hoffmann, L., and Pfister, L.: Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmospheric Research, 100, 263-274, 2011.

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., and Feng, L.: Ensemble forecasting of typhoon rainfall and floods over a mountainous watershed in Taiwan, J. Hydrol., 506, 55-68, 2013.

Huang Jr, C., Yu, C. K., Lee, J. Y., Cheng, L. W., Lee, T. Y., and Kao, S. J.: Linking typhoon tracks and spatial rainfall patterns for improving flood lead time predictions over a mesoscale mountainous watershed, Water Resources Research, 48, 2012.

Jang, J. H., Yu, P. S., Yeh, S. H., Fu, J. C., and Huang, C. J.: A probabilistic model for real‐time flood warning based on deterministic flood inundation mapping, Hydrological processes, 26, 1079-1089, 2012.

Jaun, S., Ahrens, B., Walser, A., Ewen, T., and Schär, C.: A probabilistic view on the

August 2005 floods in the upper Rhine catchment, Natural Hazards and Earth System Science, 8, 281-291, 2008.

Kuo, C. W., Hong, J. H., Wang, H. W., Wang, Y. C., Tsun, S. C., and Li, S. C.: Comparisons of velocity profile extrapolation methods for moving-boat ADCP flow measurements, Taiwan Water Conservancy, 35-46, 2016.

Lee, K. T., Chang, C. H., Yang, M. S., and Yu, W. S.: Reservoir attenuation of floods from ungauged watersheds, Hydrological Sciences Journal, 46, 349-362, 2001.

Lee, K. T., Chung, Y. R., Lau, C. C., Meng, C. C., and Chiang, S.: A windows-based inquiry system for design discharge based on geomorphic runoff modeling, Computers and Geosciences, 32, 203-211, 2006.

Lee, K. T. and Yen, B. C.: Geomorphology and kinematic-wave-base hydrograph derivation, Journal of Hydraulic Engineering ASCE, 123, 73-80, 1997.

Li, M. H., Yang, M. J., Soong, R., and Huang, H. L.: Simulating typhoon floods with gauge data and mesoscale-modeled rainfall in a mountainous watershed, Journal of Hydrometeorology, 6, 306-323, 2005.

Ou, S., Liu, J., Tsai, C., and Hsu, T.: Numerical studies of typhoon-induced storm surge using POM and finite element depth-averaged model in Taiwan, Chinese-German Joint Symposium on Hydraulic and Ocean Engineering, Darmstadt, 2008.

Pagano, T. C., Wood, A. W., Ramos, M.-H., Cloke, H. L., Pappenberger, F., Clark, M. P., Cranston, M., Kavetski, D., Mathevet, T., and Sorooshian, S.: Challenges of operational river forecasting, Journal of Hydrometeorology, 15, 1692-1707, 2014.

Palmer, T. N.: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models, Quarterly Journal of the Royal Meteorological Society, 127, 279-304, 2001.

Pappenberger, F., Beven, K. J., Hunter, N. M., Bates, P. D., Gouweleeuw, B. T., Thielen, J., and de Roo, A. P. J.: Cascading model uncertainty form medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS), Hydrology and Earth System Sciences, 9, 381-393, 2005.

Pappenberger, F., Cloke, H. L., Persson, A., and Demeritt, D.: HESS Opinions On forecast (in)consistency in a hydro-meteorological chain: curse or blessing?, Hydrology and Earth System Sciences, 15, 2011.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Pappenberger, F., Stephens, E., Thielen, J., Salamon, P., Demeritt, D., van Andel, S. J., Wetterhall, F., and Alfieri, L.: Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication, Hydrological Processes, 27, 132-146, 2013.

Ramos, M. H., Mathevet, T., Thielen, J., and Pappenberger, F.: Communicating uncertainty in hydro - meteorological forecasts: mission impossible?, Meteorological Applications, 17, 223-235, 2010.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, Water Resources Research, 46, W05521, 2010.

Rossa, A., Liechti, K., Zappa, M., Bruen, M., Germann, U., Haase, G., Keil, C., and Krahe, P.: The COST 731 Action: A review on uncertainty propagation in advanced hydro-meteorological forecast systems, Atmospheric Research, 100, 150-167, 2011.

Shih, D. S., Yeh, G. T., and Cheng, J. R. C.: Model assessments of precipitation with a unified regional circulation rainfall and hydrological watershed model, Journal of Hydrologic Engineering ASCE, 17, 43-54, 2012.

Thielen, J., Bartholmes, J., Ramos, M. H., and de Roo, A. P. J.: The European Flood Alert System; Part 1: Concept and development, Hydrology and Earth System Sciences, 13, 125-140, 2009.

Thiessen, A. H.: Precipitation averages for large areas, Monthly Weather Review, 39, 1082-1084, 1911.

Thirel, G., Martin, E., Mahfouf, J. F., Massart, S., Ricci, S., and Habets, F.: A past discharges assimilation system for ensemble streamflow forecasts over France-Part 1: Description and validation of the assimilation system, Hydrology and Earth System Sciences, 14, 1623-1637, 2010.

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., and Mysiak, J. (Eds.), Springer, Dordrecht, Netherlands, 2009.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Elsevier, Amsterdam, 2006.

Yang, S. C. and Yang, T. H.: Uncertainty assessment: Reservoir inflow forecasting with ensemble precipitation forecasts and HEC-HMS, Advances in Meteorology, 2014, 2014.

Yang, T. H., Hwang, G. D., Tsai, C.-C., and Ho, J.-Y.: Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts, Hydrology and Earth System Sciences, 20, 4731, 2016.

Yeh, G. T., Cheng, H. P., Cheng, J. R., and Lin, J. H.: A numerical model to simulate flow and contaminant and sediment 1 transport in watershed systems (WASH12D). Technical Rep. CHL-98-15, Waterways Experiment Station, U. S. Army Corps of Engineers, Vicksburg, MS 39180-6199., 1998. 1998.

Yeh, G. T., Huang, G. B., Zhang, F., Cheng, H. P., and 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 Rep. submitted to EPA, Dept. of Civil and Environmental Engineering, 2006.

Yeh, G. T., Shih, D. S., and Cheng, J. R. C.: An integrated media, integrated processes watershed model, Computers and Fluids, 45, 2-13, 2011.

Zappa, M., Beven, K. J., Bruen, M., Cofino, A. S., Kok, K., Martin, E., Nurmi, P., Orfila, B., Roulin, E., Schroter, K., Seed, A., Szturc, J., Vehvilainen, B., Germann, U., and Rossa, A.: Propagation of uncertainty from observing systems and NWP into hydrological models: COST-731 Working Group 2, Atomospheric Science Letter, 11, 83-91, 2010.

Zappa, M., Fundel, F., and Jaun, S.: A 'Peak-Box' approach for supporting interpretation and verification of operational ensemble peak-flow forecasts, Hydrological Processes, 27, 117-131, 2013.

Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F.: Superposition of three sources of uncertainties in operational flood forecasting chains, Atmospheric Research, 100, 246-262, 2011.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Hydrology and Earth System Sciences
Discussions

FIGURES

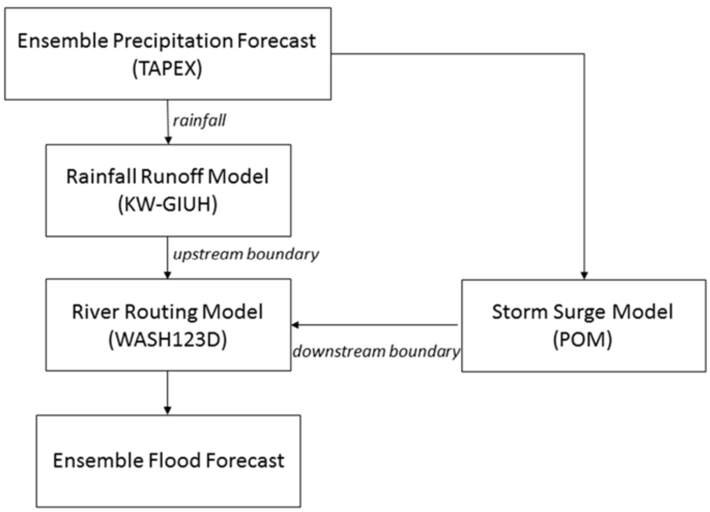


Figure 1 Flowchart describing the flow of data processing within the Yilan River HEPS.

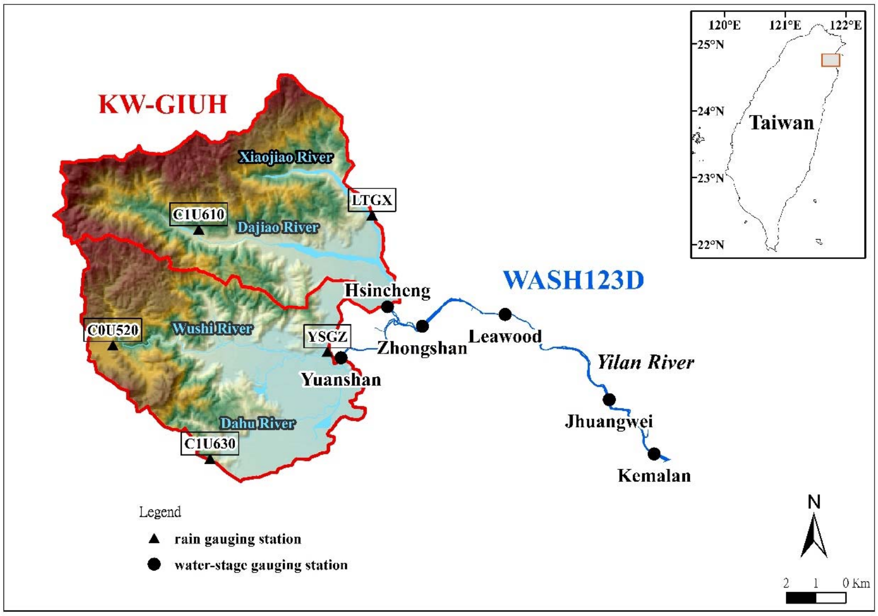
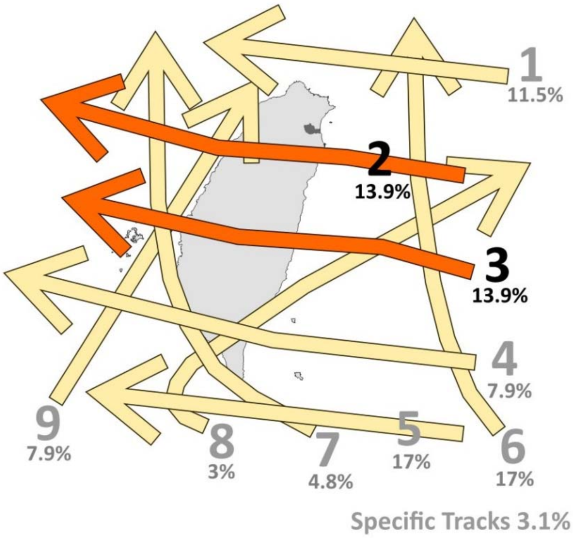


Figure 2 Study area and locations of streamflow gauges. Black dots and triangles indicate the locations of water-stage gauging stations and rain gauge stations, respectively.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017



Redrawn from Kuo et al. (2012)

596 **Figure 3** Schematic diagram showing the tracks of typhoons invading Taiwan. The percentages shown in the figure are the statistical results from 1958 through 2006 obtained from the Central Weather Bureau (CWB). The dark gray polygon located in northern Taiwan indicates the Yilan River catchment. Type-2 and Type-3 typhoons bring heavy rainfall to the Yilan River catchment.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

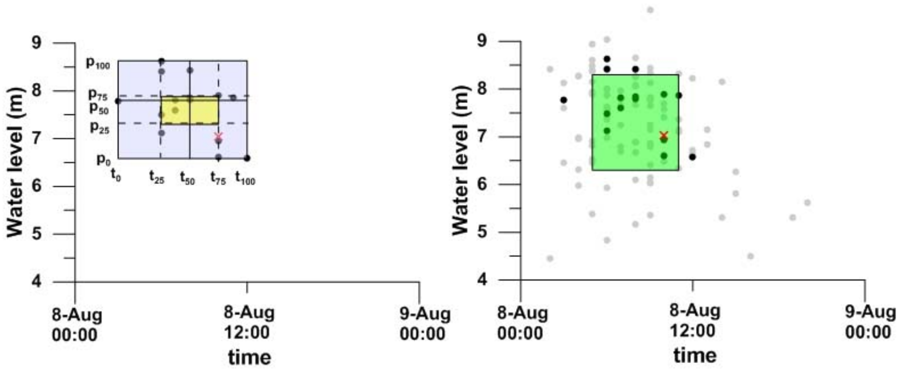


Figure 4 The left panel shows a graphical explanation of the 'Peak-Box' approach. The outer rectangle is the 'Peak-Box,' and the internal rectangle (the yellow area) is the 'IQR-Box'. The solid dots represent all of the ensemble forecasts. The right panel shows a graphic explanation of the proposed extension of the 'Peak-Box' approach. The enveloping rectangle is the 'SD-Box' (the green area). The solid black and gray dots represent current and previous peak-flow forecasts, respectively.

Hydrology and Earth System Sciences
Open Access
Discussions

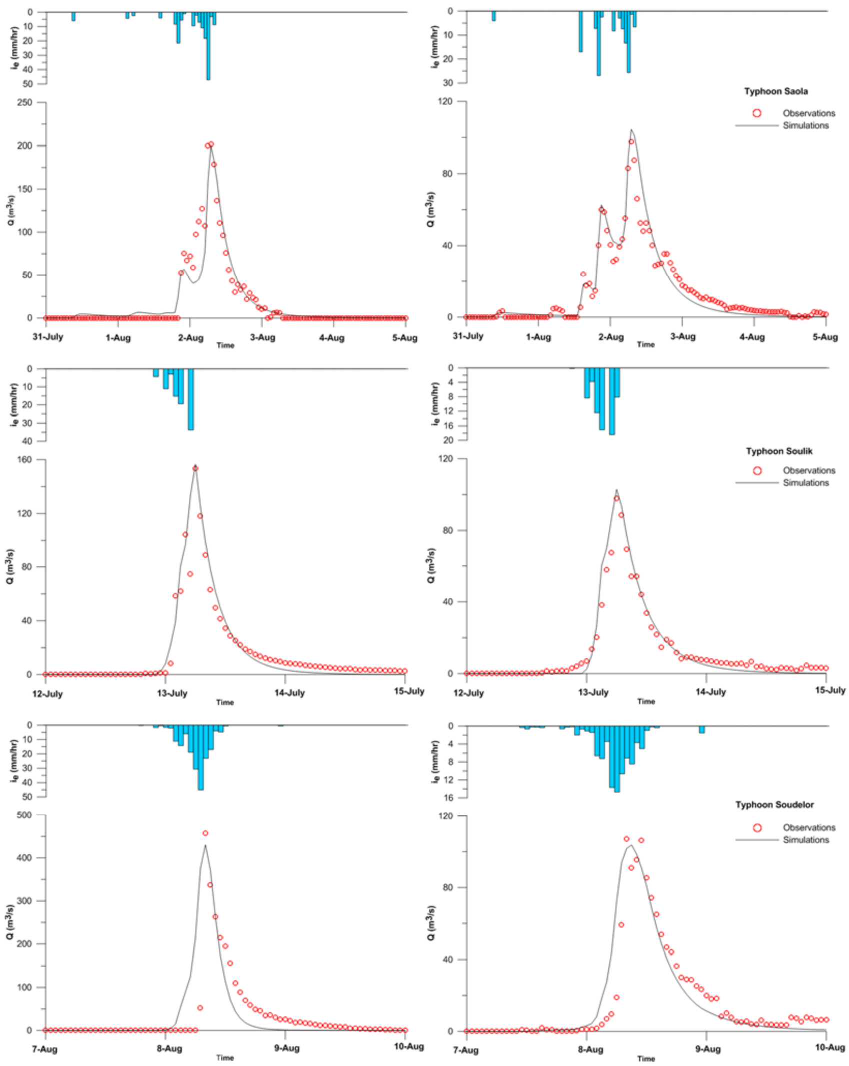


Figure 5 Comparison of simulated discharges (red circles) and recorded discharges (solid lines) for model calibration (Typhoons Saola and Soulik) and validation (Typhoon Soudelor) experiments at Hsinsheng (left) and Yuanshen (right). The blue bars are the hourly spatial-average rainfall intensities measured in the watershed upstream of Hsinsheng and Yuanshen.

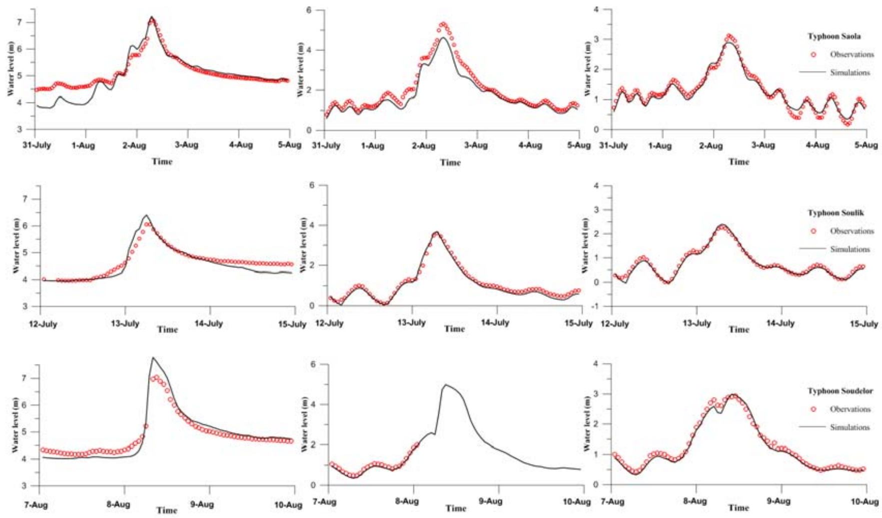
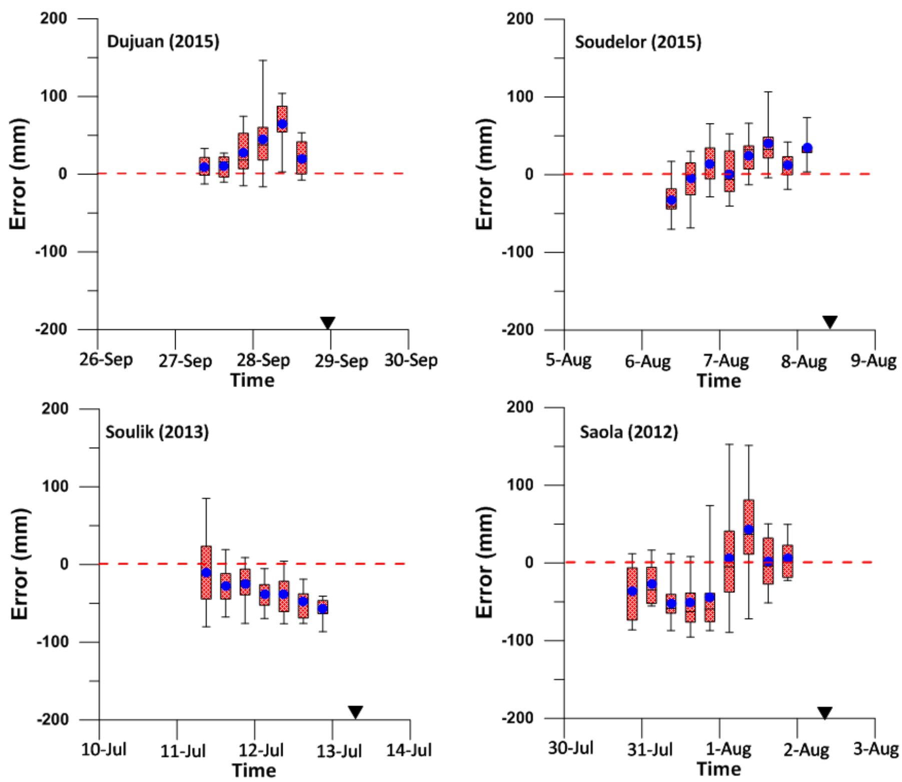


Figure 6 Comparison of simulated (red circles) and recorded (solid lines) water levels for model calibration (Typhoons Saola and Soulik) and validation (Typhoon Soudelor) experiments at Zhongshan (left), Leawood (central) and Jhungwei (right).

(a) Magnitude error of maximum 4-hour rainfall



(b) Timing error of maximum 4-hour rainfall

Hydrology and Earth System Sciences
Discussions
Open Access

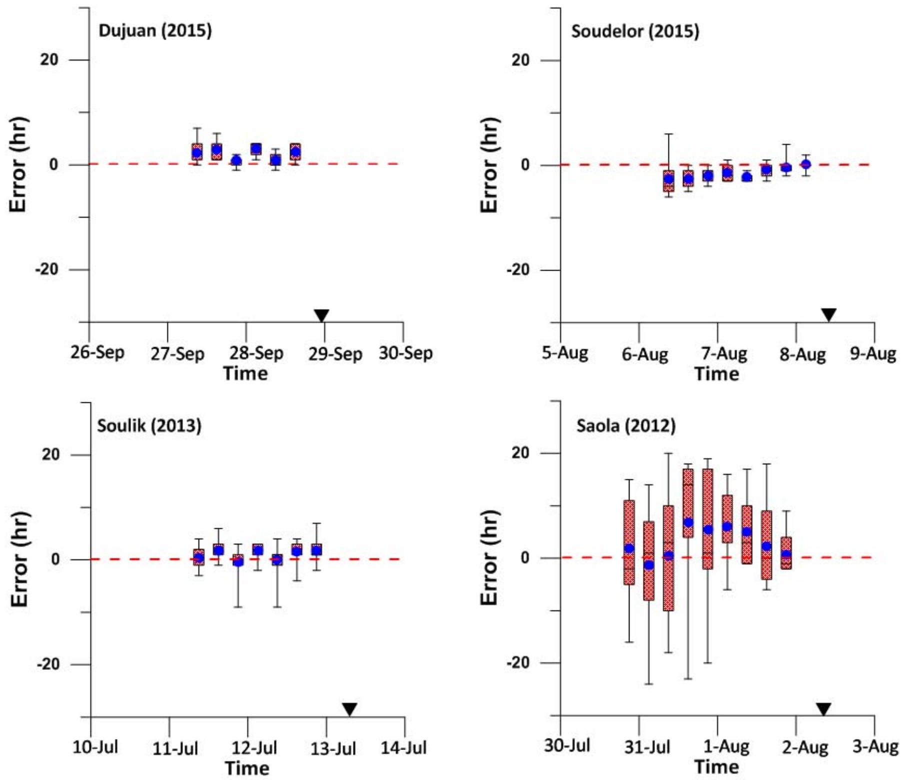
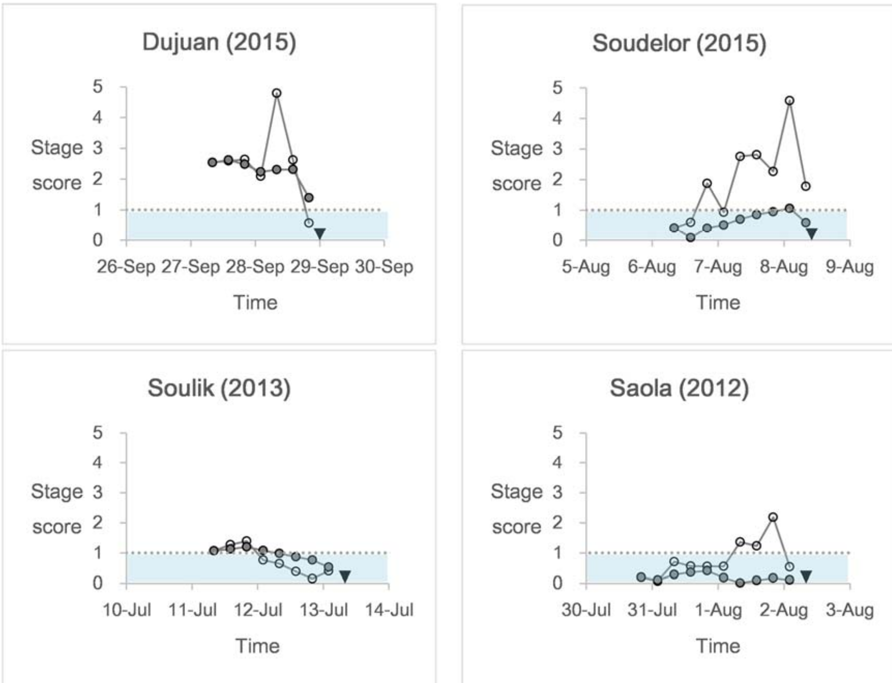


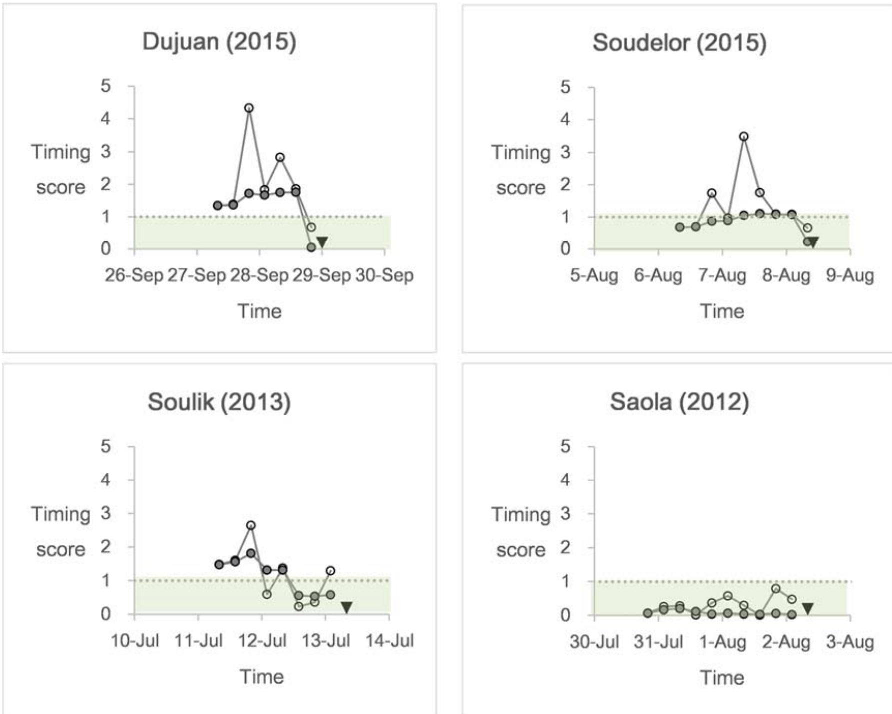
Figure 7 Box-and-whisker plot at the watershed upstream of the Zhongshan Bridge during the four selected typhoon events. The blue dots indicate the ensemble means. The inverted triangles indicate the time of occurrence of the maximum 4-hour rainfall. The results show that there is no obvious trend in lead time for the errors in either the stage or timing.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

(a) Scores for peak-stage forecasts



(b) Scores in peak-timing forecasts



622 **Figure 8** The scores 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.

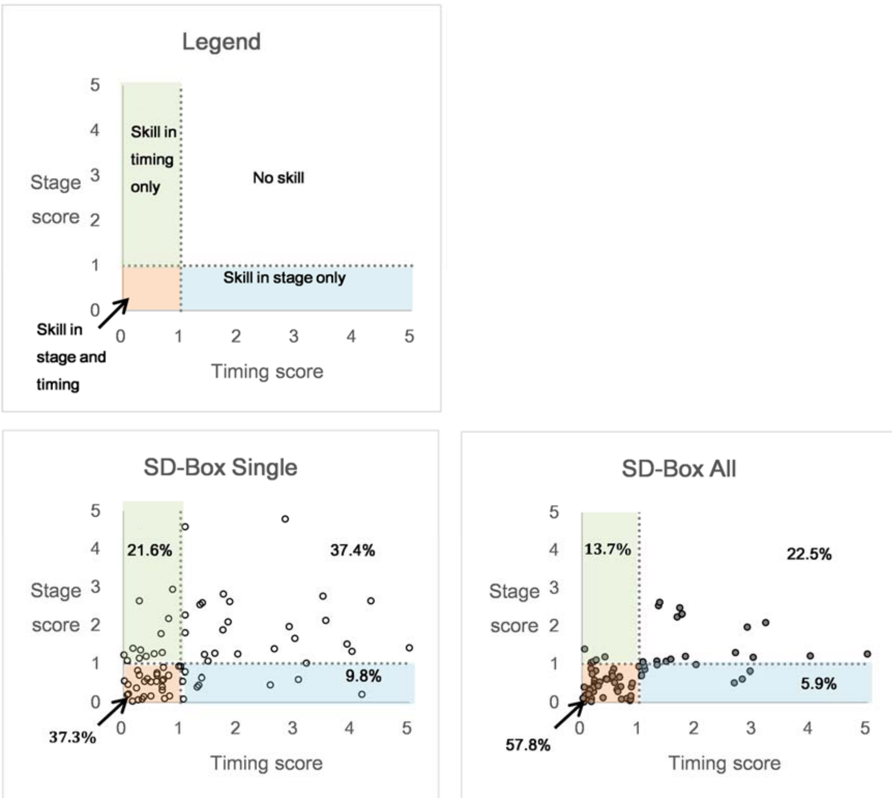


Figure 9 Comparison of scores obtained for 'SD-Box Single' and 'SD-Box All'. The results show that the 'SD-Box All' approach significantly improves the performance compared with the results obtained using the 'SD-Box Single' method.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

TALBLES

Table 1 All typhoons that invaded Taiwan during 2012 through 2015. A total of four typhoons of Type-2 and Type-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 in this study. Typhoon Matmo, a Type-3 typhoon that occurred in 2014, was not selected due to its weak typhoon 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 Jun 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
		—	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 Jun 2012
TALIM	9	—	19-21 Jun 2012

(Source: Central Weather Bureau, Taiwan)

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

Table 2 Comparisons of scores in peak stage and peak time between the 'IQR-Box' and 'SD-Box' approaches. Scores less than one (highlighted) indicate that the enveloping rectangle did contain the observed peak.

(a) Scores in peak-stage forecasts

Location/Typhoon	Method	Forecast									
		1	2	3	4	5	6	7	8	9	10
Zhongshan Bridge											
Dujuan (2015)	SD-Box	2.54	2.59	2.64	2.09	4.79	2.62	0.57	–	–	–
	IQR-Box	2.83	3.78	3.30	4.53	14.03	2.87	1.07	–	–	–
Soudelor (2015)	SD-Box	0.41	0.60	1.88	0.93	2.76	2.82	2.27	4.59	1.78	–
	IQR-Box	0.22	1.26	2.20	1.14	3.39	7.00	4.07	10.60	2.58	–
Soulik (2013)	SD-Box	1.07	1.27	1.39	0.76	0.64	0.38	0.15	0.40	–	–
	IQR-Box	1.86	1.76	1.94	1.29	0.87	0.36	0.65	0.56	–	–
Saola (2012)	SD-Box	0.20	0.07	0.71	0.56	0.55	0.55	1.36	1.23	2.18	0.54
	IQR-Box	0.14	0.01	1.81	0.79	1.70	1.42	3.66	1.90	2.45	0.48
Leawood Bridge											
Dujuan (2015)	SD-Box	1.21	1.27	1.75	1.24	3.48	1.48	1.67	–	–	–
	IQR-Box	1.10	2.29	2.17	2.98	11.15	1.84	3.23	–	–	–
Soudelor (2015)	SD-Box	–	–	–	–	–	–	–	–	–	–
	IQR-Box	–	–	–	–	–	–	–	–	–	–
Soulik (2013)	SD-Box	0.79	0.95	1.06	0.36	0.20	0.10	0.27	0.54	–	–
	IQR-Box	1.76	1.79	2.06	0.75	0.31	0.16	0.09	0.76	–	–
Saola (2012)	SD-Box	0.93	1.25	1.66	1.32	1.41	0.16	0.29	0.22	0.04	1.36
	IQR-Box	1.14	2.12	2.71	1.60	2.51	0.00	1.32	0.01	0.28	1.36
Zhuangwei Bridge											
Dujuan (2015)	SD-Box	1.97	2.13	0.60	0.21	0.46	1.51	2.94	–	–	–
	IQR-Box	2.76	2.88	0.73	0.35	1.62	1.93	4.29	–	–	–
Soudelor (2015)	SD-Box	1.19	0.17	0.45	0.10	1.01	1.24	0.55	1.81	2.64	–
	IQR-Box	1.47	0.23	0.31	0.00	0.87	3.30	0.85	3.03	3.69	–
Soulik (2013)	SD-Box	0.62	0.71	0.79	0.17	0.03	0.32	0.47	0.90	–	–
	IQR-Box	1.45	1.53	1.77	0.49	0.00	0.40	0.45	1.18	–	–
Saola (2012)	SD-Box	0.82	1.08	1.40	1.14	1.26	0.09	0.70	0.09	0.22	1.29
	IQR-Box	1.06	2.39	2.55	1.57	3.42	0.39	1.77	0.37	0.03	1.39

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-264, 2017
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 29 May 2017

644 **(b) Scores in peak-timing forecasts**

Location/Typhoon	Method	Forecast 1	2	3	4	5	6	7	8	9	10
Zhongshan Bridge											
Dujuan (2015)	SD-Box	1.34	1.38	4.33	1.83	2.83	1.86	0.68	—	—	—
	IQR-Box	3.67	3.00	9.00	2.00	3.00	1.67	0.94	—	—	—
Soudelor (2015)	SD-Box	0.68	0.70	1.74	0.97	3.49	1.75	1.08	1.08	0.66	—
	IQR-Box	1.00	1.67	3.00	1.00	7.00	2.00	—	3.00	5.40	—
Soulik (2013)	SD-Box	1.48	1.60	2.64	0.59	1.37	0.23	0.36	1.29	—	—
	IQR-Box	3.00	3.57	4.00	1.00	2.00	1.00	1.00	2.33	—	—
Saola (2012)	SD-Box	0.07	0.26	0.28	0.02	0.37	0.58	0.30	0.01	0.79	0.48
	IQR-Box	0.10	0.29	0.81	0.18	0.67	1.14	0.33	0.11	1.00	0.56
Leawood Bridge											
Dujuan (2015)	SD-Box	0.46	0.11	1.69	0.32	2.24	0.58	0.71	—	—	—
	IQR-Box	1.00	0.33	3.00	0.20	3.00	0.60	1.00	—	—	—
Soudelor (2015)	SD-Box	—	—	—	—	—	—	—	—	—	—
	IQR-Box	—	—	—	—	—	—	—	—	—	—
Soulik (2013)	SD-Box	0.40	1.17	1.96	0.39	0.71	0.11	0.09	0.96	—	—
	IQR-Box	1.18	5.00	3.00	1.00	1.00	1.00	1.00	1.50	—	—
Saola (2012)	SD-Box	0.04	0.09	0.34	0.17	0.04	0.11	0.46	0.07	0.67	0.53
	IQR-Box	0.29	0.10	0.76	0.22	0.53	0.88	0.50	0.00	0.80	1.00
Zhuangwei Bridge											
Dujuan (2015)	SD-Box	2.90	3.54	3.06	4.17	2.57	3.91	0.86	—	—	—
	IQR-Box	6.33	11.00	5.00	7.00	3.00	4.20	1.13	—	—	—
Soudelor (2015)	SD-Box	0.40	0.48	1.32	0.72	3.20	1.42	1.04	1.08	0.28	—
	IQR-Box	0.50	1.00	1.67	1.00	3.00	2.00	3.00	3.00	0.00	—
Soulik (2013)	SD-Box	0.42	0.59	1.08	0.81	0.16	0.68	0.08	0.70	—	—
	IQR-Box	0.33	1.00	2.00	3.00	0.14	3.00	1.00	1.00	—	—
Saola (2012)	SD-Box	0.25	0.07	0.17	0.28	0.54	1.05	0.79	0.33	0.09	0.68
	IQR-Box	0.72	0.50	0.00	3.43	1.07	1.71	1.00	0.44	0.00	5.00

645