1. *The sharpness of the BOX*

The length and width of a Box in Figure 1 indicate the variation of the forecasts in terms of peak-stage and peak-time. The shorter length or width means less varied forecasts in peak-stage or peak-time.
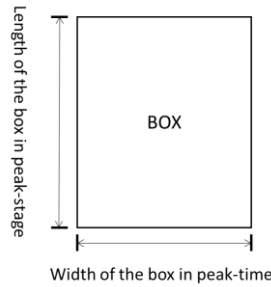

Figure 1 The length and width of a BOX

Tables 1 and 2 compare the length of SD-Box and IQR-Box for any given forecast and inclusion of all the available forecasts at any given time. Figure 2 illustrates the change of the length during different forecast run. SD-Box-single (blue line) shows a longer length than the IQR-Box-single (black line). For example, the biggest difference is about 1.75m for Zhongshan Bridge at the 6th forecast in Saola (2012). Longer length increases the probability of hit.

The difference between the length of SD-Box and IQR-Box is improved by including all available forecasts at any given time. The differences are less than 1m during each forecast run and the probability of hit is increased for both SD-Box-all and IQR-Box-all. Although the length of SD-Box-all is still large than IQR-Box-all, the probability of hit is about two times of that of IQR-Box-all.

Tables 3 and 4 compare the width of SD-Box and IQR-Box for any given forecast and inclusion of all available forecasts at any given time. Figure 3 illustrates the change of the width during different forecast run. Same as the finding from peak-stage, the difference between the widths of peak-time is improved by including all the available forecasts at any given time. Due to the dispersity of rainfall forecast in Saola (2012) as shown in Figure 7 in the original manuscript, the widths of SD-Box-all and IQR-Box-all are much larger than that in other events. It is the evidence that the SD-Box can reflect the characteristics of the forecast of typhoons and increase the users' confidence. However, the SD-Box might over-forecast the peak-time since the width of SD-Box in peak-time is larger than that of IQR-Box. The method still provide useful information about peak-time and increase people' attention to possible floods during typhoons.

Table 1 The length (m) of SD-Box and IQR-Box in peak-stage ***at any given forecast***. The highlighted number means the observation is lied in the length of peak-stage. The last column of the table is the probability of hit of peak-stage during each typhoon event.

Location/Typhoon	Method	Forecast										Hit
		1	2	3	4	5	6	7	8	9	10	
Zhongshan Bridge												
Dujuan (2015)	SD-Box	1.00	1.04	1.32	1.84	1.04	1.12	*1.21*	—	—	—	14%
	IQR-Box	0.83	0.74	1.05	0.88	0.37	1.01	0.70	—	—	—	0%
Soudelor (2015)	SD-Box	*1.72*	*1.74*	1.03	*1.47*	0.94	1.07	0.93	0.59	1.59	—	33%
	IQR-Box	*1.38*	0.93	0.87	1.29	0.71	0.43	0.56	0.25	1.13	—	11%
Soulik (2013)	SD-Box	2.66	1.70	1.70	*1.73*	*1.63*	*1.43*	*1.09*	*2.75*	—	—	63%
	IQR-Box	1.70	1.27	1.22	1.09	*1.19*	*0.97*	*0.61*	*2.07*	—	—	50%
Saola (2012)	SD-Box	*2.72*	*1.56*	*2.06*	*2.58*	*2.58*	*3.30*	2.36	1.57	1.03	*2.46*	70%
	IQR-Box	*2.27*	*0.98*	0.93	*1.83*	1.04	1.55	0.96	1.21	0.82	*2.16*	40%
Leawood Bridge												
Dujuan (2015)	SD-Box	0.77	0.82	0.94	1.49	0.78	0.80	1.68	—	—	—	0%
	IQR-Box	0.69	0.50	0.75	0.67	0.26	0.65	0.90	—	—	—	0%
Soudelor (2015)	SD-Box	—	—	—	—	—	—	—	—	—	—	—
	IQR-Box	—	—	—	—	—	—	—	—	—	—	—
Soulik (2013)	SD-Box	*2.45*	*1.58*	1.52	*1.82*	*1.76*	*1.58*	*1.35*	*2.40*	—	—	88%
	IQR-Box	1.38	0.94	0.87	*0.97*	*1.25*	*1.26*	*0.88*	*1.76*	—	—	63%
Saola (2012)	SD-Box	*2.96*	1.46	2.06	2.65	2.50	*2.90*	*2.18*	*1.17*	*0.78*	2.69	50%
	IQR-Box	2.23	0.83	1.31	2.24	1.46	*1.17*	0.69	*0.89*	*0.60*	2.47	30%
Zhuangwei Bridge												
Dujuan (2015)	SD-Box	0.52	0.46	*0.76*	*1.21*	*0.78*	0.59	1.24	—	—	—	43%
	IQR-Box	0.41	0.33	*0.62*	*0.65*	0.29	0.46	0.87	—	—	—	29%
Soudelor (2015)	SD-Box	0.99	*0.99*	*0.60*	*0.82*	0.63	0.67	*0.51*	0.42	1.17	—	44%
	IQR-Box	0.66	*0.40*	*0.48*	*0.63*	*0.55*	0.23	*0.35*	0.23	0.87	—	56%
Soulik (2013)	SD-Box	*1.64*	*0.90*	0.83	*1.03*	*0.94*	*0.90*	*0.92*	*1.65*	—	—	100%
	IQR-Box	0.91	0.44	0.41	*0.48*	*0.67*	*0.62*	*0.43*	1.27	—	—	50%
Saola (2012)	SD-Box	*1.79*	0.94	1.36	1.71	1.68	*2.37*	*1.63*	*0.98*	*0.84*	1.70	50%
	IQR-Box	1.25	0.44	0.77	1.29	0.66	*1.05*	0.73	*0.83*	*0.65*	1.48	30%

Table 2 The length (m) of SD-Box and IQR-Box in peak-stage *including all the available forecasts at any given time*. The highlighted number means the observation is lied in the length of peak-stage. The last column of the table is the probability of hit of peak-stage during each typhoon event.

| Location/Typhoon	Method	Forecast										Hit
		1	2	3	4	5	6	7	8	9	10	
Zhongshan Bridge												
Dujuan (2015)	SD-Box	1.00	1.00	1.17	1.41	1.52	1.47	2.03	—	—	—	0%
	IQR-Box	0.83	0.84	0.74	0.90	1.21	1.06	1.25	—	—	—	0%
Soudelor (2015)	SD-Box	*1.72*	*1.92*	*1.87*	*1.78*	*1.77*	*1.80*	*1.71*	1.66	*2.16*	—	89%
	IQR-Box	*1.38*	*1.56*	*1.46*	*1.30*	1.27	1.13	1.03	0.90	1.29	—	44%
Soulik (2013)	SD-Box	2.66	2.22	2.04	2.01	*1.98*	*1.97*	*1.94*	*2.22*	—	—	50%
	IQR-Box	1.70	1.50	1.43	1.37	1.48	1.43	*1.39*	*1.51*	—	—	25%
Saola (2012)	SD-Box	*2.72*	*2.19*	*2.20*	*2.29*	*2.34*	*2.71*	*2.95*	*2.88*	*2.80*	*2.81*	100%
	IQR-Box	*2.27*	*1.51*	*1.52*	*1.74*	*1.73*	*2.12*	*2.31*	*2.22*	*2.11*	*2.26*	100%
Leawood Bridge												
Dujuan (2015)	SD-Box	0.77	0.78	0.88	1.08	1.16	1.12	*1.95*	—	—	—	14%
	IQR-Box	0.69	0.55	0.49	0.68	0.87	0.80	0.92	—	—	—	0%
Soudelor (2015)	SD-Box	—	—	—	—	—	—	—	—	—	—	—
	IQR-Box	—	—	—	—	—	—	—	—	—	—	—
Soulik (2013)	SD-Box	*2.45*	*2.04*	*1.86*	*1.89*	*1.90*	*1.91*	*1.90*	*2.08*	—	—	100%
	IQR-Box	1.38	1.17	1.10	1.02	1.29	*1.40*	*1.36*	*1.45*	—	—	38%
Saola (2012)	SD-Box	*2.96*	*2.33*	2.28	2.37	2.39	*2.63*	*2.80*	2.72	2.64	2.70	70%
	IQR-Box	2.23	1.44	1.77	1.91	1.96	2.13	*2.31*	*2.11*	*1.94*	*2.13*	40%
Zhuangwei Bridge												
Dujuan (2015)	SD-Box	0.52	0.48	0.64	*0.84*	*0.92*	*0.89*	*1.43*	—	—	—	57%
	IQR-Box	0.41	0.35	0.32	0.45	*0.73*	*0.60*	0.68	—	—	—	29%
Soudelor (2015)	SD-Box	0.99	*1.10*	*1.06*	*1.00*	*1.01*	*1.02*	0.96	0.94	1.38	—	89%
	IQR-Box	0.66	*0.74*	*0.67*	*0.60*	*0.63*	*0.67*	*0.63*	*0.58*	*0.68*	—	89%
Soulik (2013)	SD-Box	*1.64*	*1.31*	*1.17*	*1.15*	*1.14*	*1.14*	*1.14*	*1.34*	—	—	100%
	IQR-Box	0.91	0.68	0.58	0.51	*0.59*	*0.66*	*0.68*	*0.78*	—	—	50%
Saola (2012)	SD-Box	*1.79*	*1.41*	1.41	1.48	1.52	*1.82*	1.99	1.92	1.86	1.89	70%
	IQR-Box	1.25	0.80	0.95	1.07	1.10	1.19	*1.48*	*1.29*	*1.23*	*1.33*	40%
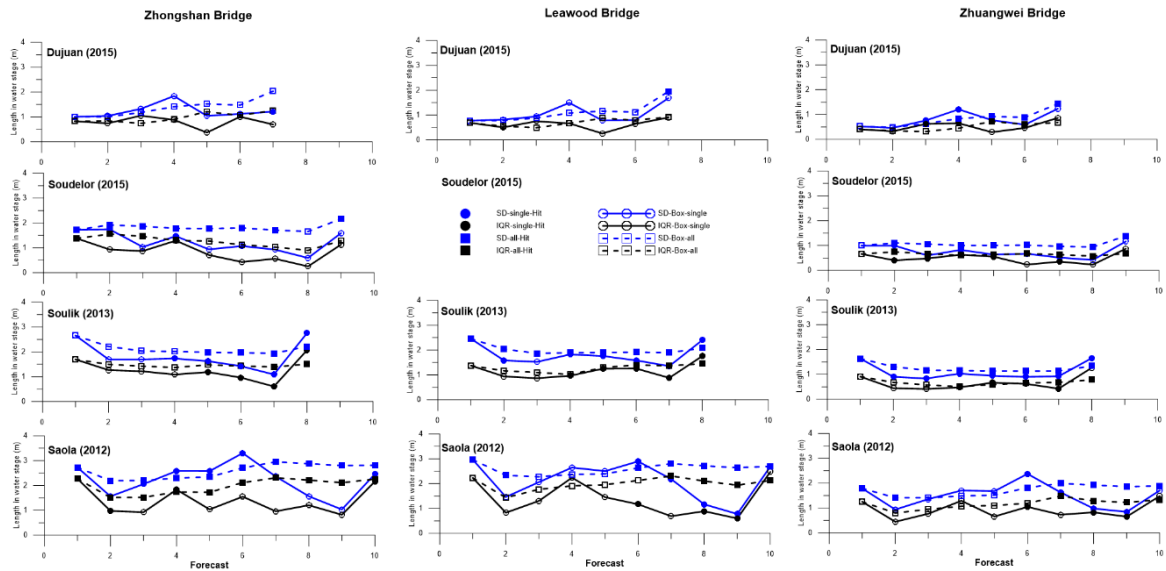
Figure 2 Changes in length of peak-stage for SD-Box (blue) and IQR-Box (black) considering only one forecast (solid lines) and inclusion of all the available forecasts at any given time (dotted lines). The solid circles and squares mean the observations were forecasted.

Table 3 The width (hr) of SD-Box and IQR-Box in peak-time *at any given forecast*. The highlighted number means the observation is lied in the width of peak-time. The last column of the table is the probability of hit of peak-time during each typhoon event.

Location/Typhoon	Method	Forecast										Hit
		1	2	3	4	5	6	7	8	9	10	
Zhongshan Bridge												
Dujuan (2015)	SD-Box	3.93	3.03	1.89	2.19	3.14	2.74	67.81	–	–	–	14%
	IQR-Box	1.50	1.50	1.00	2.00	3.00	3.00	67.00	–	–	–	14%
Soudelor (2015)	SD-Box	7.24	5.51	3.10	3.95	1.85	2.63	3.00	2.62	23.60	–	44%
	IQR-Box	5.00	3.00	2.00	4.00	1.00	2.00	0.00	1.00	1.25	–	22%
Soulik (2013)	SD-Box	3.78	4.01	2.92	2.41	3.12	4.42	5.17	4.97	–	–	38%
	IQR-Box	2.00	1.75	2.00	1.00	2.50	1.00	1.00	3.00	–	–	38%
Saola (2012)	SD-Box	23.80	24.10	21.12	39.44	20.74	11.25	12.88	15.40	6.71	37.01	100%
	IQR-Box	15.50	14.00	10.50	22.00	15.00	7.00	10.50	9.00	5.00	4.50	90%
Leawood Bridge												
Dujuan (2015)	SD-Box	3.56	3.32	2.37	2.88	2.60	2.80	68.87	–	–	–	71%
	IQR-Box	2.00	1.50	1.50	2.50	2.00	2.50	67.00	–	–	–	71%
Soudelor (2015)	SD-Box	–	–	–	–	–	–	–	–	–	–	–
	IQR-Box	–	–	–	–	–	–	–	–	–	–	–
Soulik (2013)	SD-Box	24.51	3.55	2.92	2.18	2.83	10.52	6.68	5.83	–	–	75%
	IQR-Box	2.75	1.00	2.00	1.00	2.00	1.75	1.00	4.00	–	–	50%
Saola (2012)	SD-Box	23.05	23.54	34.90	46.32	37.66	23.46	12.16	15.48	6.56	37.12	50%
	IQR-Box	15.50	10.00	12.50	34.50	17.00	8.50	10.00	9.00	5.00	4.50	30%
Zhuangwei Bridge												
Dujuan (2015)	SD-Box	3.52	3.24	2.26	2.57	2.54	2.70	69.62	–	–	–	14%
	IQR-Box	1.50	1.00	1.50	1.50	2.00	2.50	69.00	–	–	–	0%
Soudelor (2015)	SD-Box	8.02	5.75	3.96	4.29	1.97	3.04	3.11	2.62	19.32	–	44%
	IQR-Box	8.00	4.00	3.00	3.00	2.00	2.00	1.00	1.00	6.00	–	44%
Soulik (2013)	SD-Box	4.40	3.63	3.18	2.81	2.62	4.62	7.17	6.78	–	–	88%
	IQR-Box	3.00	2.00	2.00	1.00	1.75	1.00	2.00	5.00	–	–	63%
Saola (2012)	SD-Box	23.69	23.44	16.91	38.52	19.89	10.74	11.90	15.48	6.33	37.54	90%
	IQR-Box	14.50	10.00	7.00	7.00	13.75	7.00	9.50	9.00	5.00	2.00	60%

Table 4 The width (hr) of SD-Box and IQR-Box in peak-time *including all the available forecasts at any given time*. The highlighted number means the observation is lied in the width of peak-time. The last column of the table is the probability of hit of peak-time during each typhoon event.

| Location/Typhoon	Method	Forecast										Hit
		1	2	3	4	5	6	7	8	9	10	
Zhongshan Bridge												
Dujuan (2015)	SD-Box	3.93	3.47	3.43	3.24	3.49	3.38	30.81	—	—	—	14%
	IQR-Box	1.50	1.75	2.00	2.00	2.00	2.00	3.00	—	—	—	0%
Soudelor (2015)	SD-Box	7.24	6.33	5.44	5.09	4.68	4.39	4.24	4.12	10.34	—	56%
	IQR-Box	5.00	3.75	3.00	3.00	2.00	2.75	2.50	3.00	3.00	—	11%
Soulik (2013)	SD-Box	3.78	3.85	3.62	4.03	3.86	9.72	9.28	8.85	—	—	38%
	IQR-Box	2.00	2.00	2.00	3.00	3.00	4.00	4.00	4.00	—	—	0%
Saola (2012)	SD-Box	23.80	23.49	22.41	27.22	26.25	27.50	25.93	24.79	23.50	25.60	100%
	IQR-Box	15.50	14.75	14.00	19.00	18.00	16.00	14.00	14.00	13.50	13.00	100%
Leawood Bridge												
Dujuan (2015)	SD-Box	3.56	3.42	3.39	3.27	3.54	3.42	30.96	—	—	—	100%
	IQR-Box	2.00	2.00	2.00	2.00	2.50	2.00	2.00	—	—	—	100%
Soudelor (2015)	SD-Box	—	—	—	—	—	—	—	—	—	—	—
	IQR-Box	—	—	—	—	—	—	—	—	—	—	—
Soulik (2013)	SD-Box	24.51	17.43	14.25	12.77	11.52	14.76	13.95	13.20	—	—	100%
	IQR-Box	2.75	2.00	1.75	3.00	3.00	4.00	4.00	3.25	—	—	66%
Saola (2012)	SD-Box	23.05	22.79	27.39	32.55	33.29	34.25	32.22	30.61	28.92	30.43	100%
	IQR-Box	15.50	15.25	14.00	22.00	19.50	16.00	16.00	15.25	14.00	12.75	100%
Zhuangwei Bridge												
Dujuan (2015)	SD-Box	3.52	3.36	3.54	3.34	3.44	3.35	31.03	—	—	—	14%
	IQR-Box	1.50	2.00	3.00	2.25	3.00	2.00	2.00	—	—	—	0%
Soudelor (2015)	SD-Box	8.02	6.84	6.07	5.65	5.23	4.92	4.70	4.51	7.89	—	100%
	IQR-Box	8.00	5.75	4.00	4.00	4.00	3.00	3.00	2.00	2.50	—	56%
Soulik (2013)	SD-Box	4.40	3.96	3.74	4.08	3.83	11.01	10.54	10.19	—	—	100%
	IQR-Box	3.00	2.25	2.75	3.00	3.00	3.00	3.75	3.00	—	—	100%
Saola (2012)	SD-Box	23.69	23.10	20.97	26.12	25.01	27.44	25.76	24.63	23.36	25.69	100%
	IQR-Box	14.50	15.00	15.00	16.50	15.50	14.00	13.25	13.00	13.00	13.00	100%
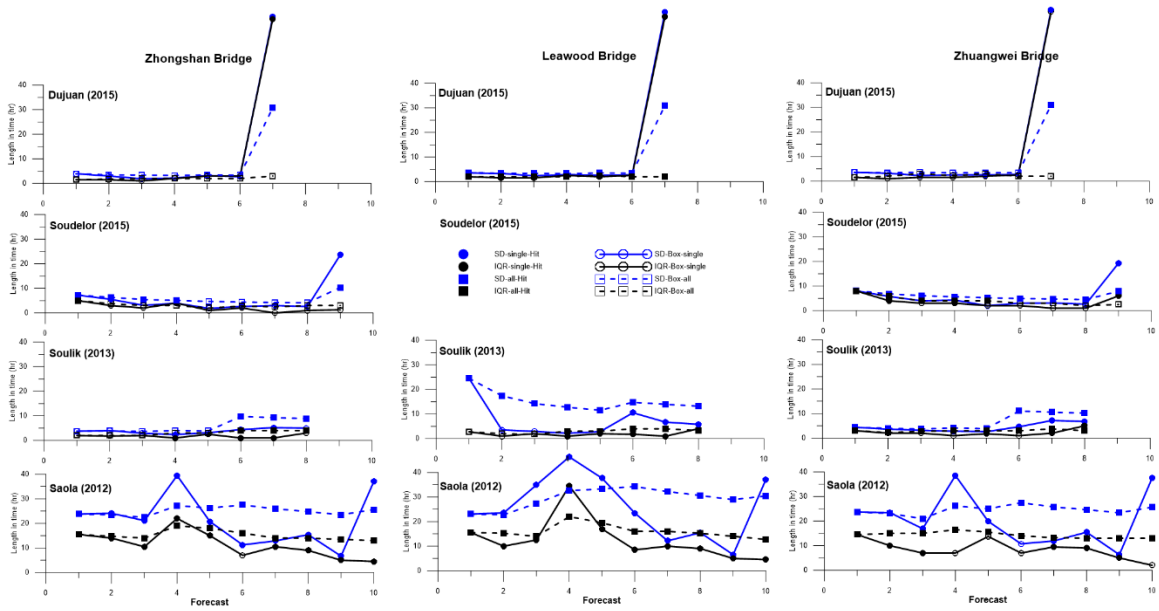
Figure 3 Changes in width of peak-time for SD-Box (blue) and IQR-Box (black) considering only one forecast (solid lines) and inclusion of all the available forecasts (dotted lines) at any given time. The solid circles and squares mean the observations are forecasted.
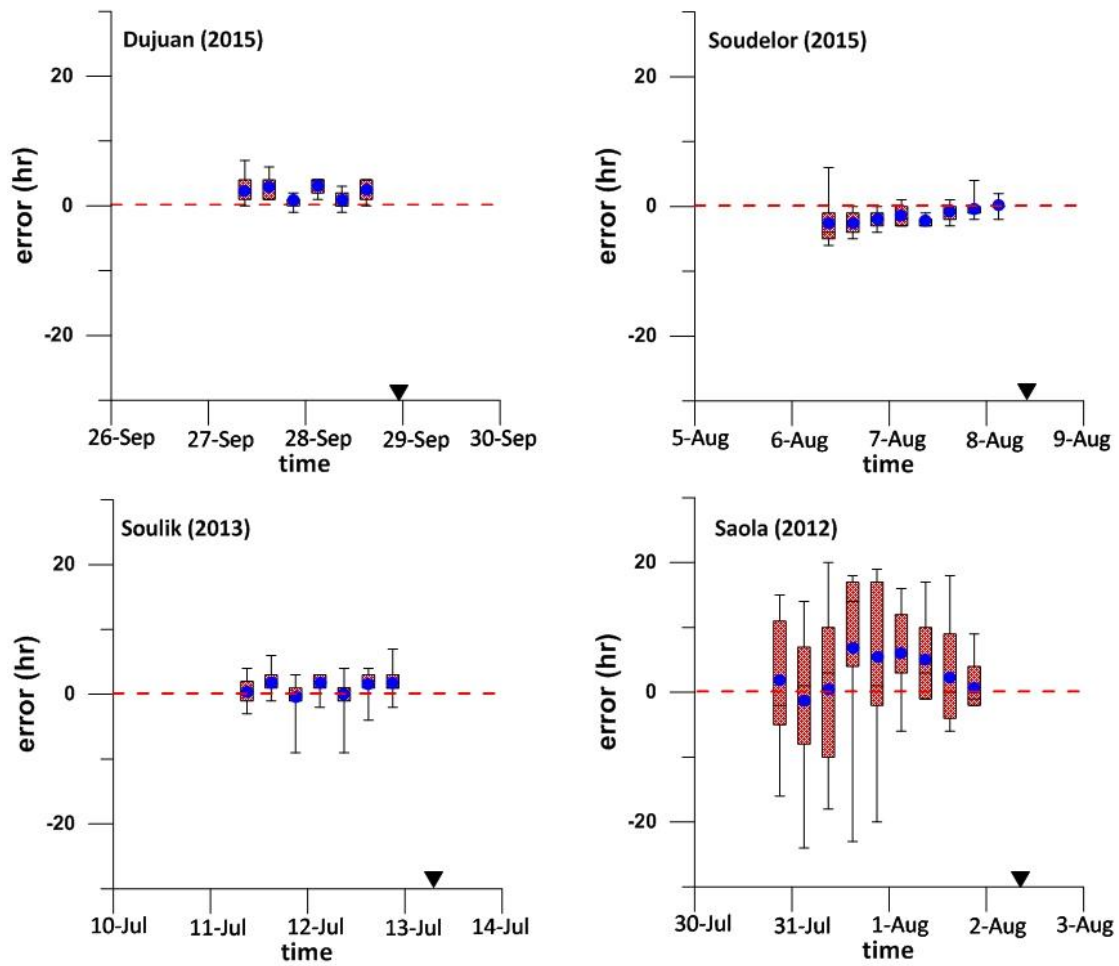
Figure 7 (in original manuscript) Box-and-whisker plot at the upstream watershed of Zhongshan Bridge during four selected typhoon events. The blue dots indicates the ensemble means. The inverted triangle indicates the time of occurrence of maximum 4-hour rainfall. The results show that there is no obvious trend in lead time for both magnitude and timing errors.

2. The modified figure 4 in the manuscript

Hydrology and
Earth System
Sciences
Discussions

Open Access

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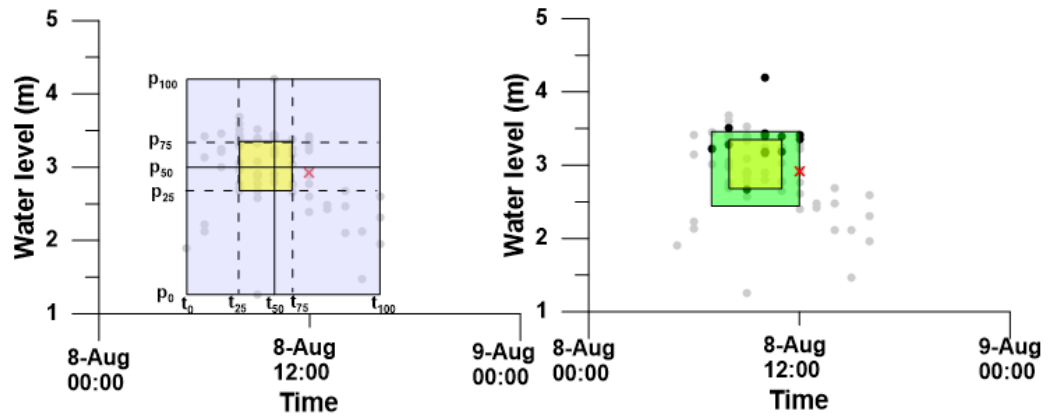
Development of a hydrological ensemble prediction system and a visualization approach for improved interpretation during typhoon events

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

6 [1]*Taiwan Typhoon and Flood Research Institute (TTFRI), National Applied Research Laboratories*
7 *(NARLabs), Taipei, Taiwan*

8 [2]*Department of River and Harbor Engineering, National Taiwan Ocean University, Keelung, Taiwan*

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

11 *E-mail: rachel.ev91@gmail.com*

12

13   ABSTRACT

14   Typhoons are accompanied by heavy rainfall and cause loss of life and property.

15 Hydrological ensemble prediction systems can provide decision makers with

16 hydrological information, such as peak stage and peak time, with some lead time. This

17 information assists decision makers in taking the necessary measures to prevent and

18 mitigate disasters. This study proposes a hydrological ensemble prediction system that

19 includes numerical weather models that perform rainfall forecasts and hydrologic

20 models that produce assessments of surface runoff and the associated flooding.

21 However, the spatiotemporal uncertainty associated with the numerical models and the

22 difficulty in interpreting the model results hinder effective decision making during

23 emergency response situations. Thus, this study also presents an extension of the 'Peak-

Hydrology and
Earth System
Sciences

Open Access

Discussions

24   Box' visualization methodology that assists in interpreting the forecast results for

25   operational purposes. A small watershed with area of 100 km^2 and four typhoons that

26   occurred from 2012 to 2015 were selected to evaluate the performance of these tools.

27   The results showed that the modified visualization approach improved the intelligibility

28   of forecasts of the peak stages and peak times compared to that of approaches

29   previously described in the literature. The new approach includes all available forecasts

30   to increase the sample size. The capture rate is greater than 50%, which is considered

31   practical for decision makers. The proposed system and the modified visualization

32   approach have demonstrated their potential for both decreasing the uncertainty of

33   numerical rainfall forecasts and improving the performance of flood forecasts.

34

35   **KEY WORDS** Hydrological ensemble prediction system; peak flow; decision support;

36   visualization.

37

1. INTRODUCTION

38

39   Numerical weather prediction (NWP) models generate different precipitation

40   forecasts for specified locations and times due to the incompleteness of the input

41   observations, the approximate nature of the forecast models and their parameterizations,

42   and the random errors that result from perturbing the initial atmospheric conditions

43   (Palmer, 2001; Hostache et al., 2011). Ensemble prediction systems (EPSs), which

44   consist of an adequate number of equiprobable NWP models, have been established to

45   provide probabilistic precipitation forecasts instead of a single deterministic forecast

46   (Cloke and Pappenberger, 2009). An EPS provides predictions with greater skill than

47   those obtained from individual runs of NWP models or deterministic model runs,

48   especially for longer lead times (Demeritt et al., 2007; Cuo et al., 2011).

49   A hydrological ensemble prediction system (HEPS) is an integrated system that

50   couples an EPS with catchment-scale hydrological models to provide flood forecasts

51   with sufficient lead time. The importance of such systems in disaster mitigation, water

52   resource management, and hydropower dam and lake operation is growing

53   (Pappenberger et al., 2005; Cloke and Pappenberger, 2009; Zappa et al., 2010, 2013;

54   Yang and Yang, 2014). However, uncertainties stemming from factors including

55   boundary conditions, initial conditions, and model parameter values affect the forecast

56   accuracy of these systems. The precipitation forecasts of NWP models dominate the

Hydrology and
Earth System
Sciences
Open Access
EGU
Discussions

57   overall uncertainty of these systems (Zappa et al., 2011; Rossa et al., 2011). It is

58   necessary to develop guidelines and tools for communicating the uncertainties

59   associated with complex HEPSs (e.g., Jaun et al., 2008; Thielen et al., 2009; Bartholmes

60   et al., 2009; Todini, 2009; Bruen et al., 2010; Renard et al., 2010; Thirel et al., 2010;

61   Zappa et al., 2010, 2013; Frick and Hegg, 2011; Pappenberger et al., 2011a, 2011b;

62   Fundel and Zappa, 2011; Pappenberger et al., 2013).

63   Effective communication of ensemble forecasts means that clear expression of the

64   uncertainties associated with HEPS is important so that end-users can easily respond to

65   the information provided during operations (Demeritt et al., 2010; Ramos et al., 2010;

66   Pappenberger et al., 2013; Zappa et al., 2013; Pagano et al., 2014). Pagano et al. (2014)

67   noted that defining effective methods for the communication of ensemble forecasts is a

68   challenge for future operational river forecasting and represents a future research

69   opportunity. Pappenberger et al. (2013) argued that the uncertainty information

70   provided by HEPSs sometimes results in resistance on the part of the public if experts

71   or nonexperts cannot easily understand the information provided. At present, HEPSs

72   still rely on conventional visualization techniques, such as 'spaghetti diagrams' or box

73   plots, to display the distributions of forecast results. Pappenberger et al. (2013) focused

74   on expert users of HEPSs and the communication among these experts and identified

75   key information for the public, such as discharge, lead time, warning levels, return

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

76   periods, worst/best scenario, etc. Zappa et al. (2013) proposed the 'Peak-Box'

77   visualization approach to support the interpretation and verification of HEPS results.

78   This approach has been used to obtain quantitative and qualitative insights, such as the

79   timing, water level, and discharge associated with peak flow. This information is crucial

80   for end-users and decision makers. Zappa et al. (2013) applied an operational HEPS,

81   namely, the IFKIS-HYDRO hydrological nowcasting system, to five different basins in

82   Switzerland to evaluate the performance of the 'Peak-Box' methodology. The sizes of

83   the basins ranged from 186 km^2 to 1696 km^2. The study found that, of 485 operational

84   forecasts performed from June 2007 through November 2008, 30% to 55% of the

85   observed peaks fell outside the 'Peak-Box'.

86   Typhoons are common natural events that cause severe damage in countries at the

87   edge of the northwestern Pacific Ocean, such as Japan, the Philippines, and Taiwan. For

88   example, based on records covering 1958 to 2010, an average of 3.4 typhoons affect

89   Taiwan annually, and these events cause an annual average loss of more than 500

90   million U.S. dollars (Li et al., 2004). Typhoon-related flood events cause these losses.

91   If they provide early warnings with sufficient lead time, flood forecasts from a HEPS

92   can help authorities prepare disaster prevention and mitigation measures. A customized

93   visualization method for typhoons is also necessary to make the ensemble flood

94   forecasts generated by HEPS meaningful for emergency responders. Therefore, this

95   study presents a HEPS that can provide ensemble flood forecasts during typhoon events

96   and proposes a customized visualization approach especially for typhoons to simplify

97   the forecast information. This approach is an extension of the one presented by Zappa

98   et al. (2013); it has been modified to increase the percentage of observed peaks that fall

99   within the predicted range during typhoon events. The remainder of this paper is

100   organized as follows. Section 2 includes the details of the proposed HEPS. Section 3

101   briefly describes the study area and typhoon events used in the study. Section 4

102   compares the original 'Peak-Box' approach with the proposed extended version. Finally,

103   Sect. 5 and 6 present the results, discussion, and conclusions.

104   ## 2.   SETUP OF THE HYDROLOGICAL ENSEMBLE PREDICTION

105   ## SYSTEM

106   This study proposes a HEPS that integrates various models. These models include

107   NWP models that provide ensemble precipitation forecasts, a rainfall-runoff model that

108   generates upstream boundary conditions, a storm surge model that generates

109   downstream boundary conditions, and a flood routing model that simulates river flows.

110   The data processing is shown in Figure 1. The HEPS produces ensemble flood forecasts

111   with a 72-hour lead time four times a day. The models used in the HEPS are described

112   in the following subsections.

113   *2.1 Ensemble precipitation forecasts*

114   The Taiwan Cooperative Precipitation Ensemble Forecast Experiment (TAPEX)

115   began in 2010. It is a collective effort among academic institutes and government

116   agencies, such as the National Taiwan University, the National Central University, the

117   National Taiwan Normal University, the Chinese Culture University, the Central

118   Weather Bureau (CWB), the National Center for High-Performance Computing, the

119   Taiwan Typhoon and Flood Research Institute (TTFRI), and the National Science and

120   Technology Center for Disaster Reduction. TAPEX is the first attempt to design a high-

121   resolution (5-km) numerical ensemble model in Taiwan. This effort applies various

122   NWP models, such as the Weather Research and Forecasting Model (WRF), the Fifth-

123   Generation Penn State/NCAR Mesoscale Model (MM5), the Cloud-Resolving Storm

124   Simulator (CReSS), and the Hurricane Weather Research and Forecasting Model

125   (HWRF). It also considers different setups in terms of the model initial conditions, data

126   assimilation processes and model physics. TAPEX generates four runs a day and

127   provides ensemble predictions of the wind and pressure fields and quantitative

128   estimates of precipitation with a lead time of 72 hours. Further information can be found

129   in Hsiao et al. (2013). A typhoon's average impact duration is 73.68 hours (Huang et

130   al., 2012) and the average lag between observed peak precipitation and flooding in

131   Taiwan is between 2 and 10 hours (Jang et al., 2012). This study focuses on a one-way

7

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

132   coupling in which TAPEX provides rainfall forecast to the rainfall-runoff model;

133   feedbacks from the rainfall-runoff model to TAPEX are not considered.

134   *Rainfall-runoff model*

135   The HEPS uses the surface runoff forecast generated by a kinematic-wave-based

136   geomorphologic instantaneous unit hydrograph model (the KW-GIUH model) as its

137   upstream boundary condition. The KW-GIUH model, which was developed by Lee and

138   Yen (1997), can reflect the effects of watershed geomorphology, land cover conditions,

139   soil characteristics and rainfall intensity on runoff. It has been successfully applied to

140   many Taiwanese catchments (Lee et al., 2001; 2006).

141   *2.3 Storm surge model*

142   Storm surges are abnormal increases in water levels above those expected from

143   astronomical tides. They are generated by strong winds and atmospheric pressure

144   changes and affect water levels downstream (near estuaries) during typhoons. The

145   HEPS uses the storm surge and tide forecasts generated by the Princeton Ocean Model

146   (POM) and the TOPEX-POSEIDON global tidal model (TPXO6.2) as downstream

147   boundary conditions. The POM model, which was developed by Blumberg and Mellor

148   (1987), is a three-dimensional, nonlinear, primitive equation finite difference ocean

149   model. It has been applied to simulate a wide range of ocean problems, including

150   coastal storm surge in Taiwan (Ou et al., 2008; Chiou, 2010). In this study, the TAPEX

151   model provides ensemble pressure field and wind field forecasts to POM and the

152   TPXO6.2 model and obtains tidal level predictions. As with TAPEX, it generates four

153   runs a day, and each run has a 72-hour lead time.

154   *2.4 Flood routing model*

155   The Numerical Model Simulating Water Flow and Contaminant and Sediment

156   Transport in WAterSHed Systems of 1D Stream/River Networks, 2D Overland

157   Regimes, and 3D Subsurface Media (WASH123D) was developed by Yeh et al. (1998)

158   to simulate one-dimensional channel networks, two-dimensional overland flow, and

159   three-dimensional variably saturated subsurface flow. It has been applied successfully

160   in Taiwan and around the world, and it was chosen by the US Army Corps of Engineers

161   as the core computational code used in modeling the Lower East Coast (LEC) Wetland

162   Watershed (e.g., Yeh et al., 2006; Yeh and Shih., 2011; Shih et al., 2012; Hsiao et al.,

163   2013). The HEPS uses the one-dimensional channel model of WASH123D as its flood

164   routing model to simulate water stages in rivers.

165   3. STUDY AREA AND TYPHOON EVENTS

166   *3.1 Study area*

167   This study selected the Yilan River in northeastern Taiwan as the study area

168   (Figure 2). The river flows through the city of Yilan and has a main stream length of

169   approximately 24.4 km and a watershed area of 149.06 km^2. It has four main tributaries,

Hydrology and
Earth System
Sciences

Open Access

Discussions

EGU

170   which are the Wushi River, the Dahu River, the Dajiao River and the Xiaojiao River.

171   The Water Resource Agency (WRA) and TTFRI have selected this river as one of two

172   watersheds where long-term monitoring experiments are being carried out (the other is

173   the Dianbao Creek basin in southwestern Taiwan). The purpose of the experimental

174   watersheds is to generate long-term and high-density hydrological monitoring data that

175   can be used for scientific studies, including the development of hydrological and

176   hydraulic models and the study of environmental changes. In total, 11 rainfall gauging

177   stations, 16 water-stage gauging stations, five river-velocity gauging stations, and 36

178   inundation-depth gauging stations have been installed in the Yilan River Basin. Figure

179   2 shows the locations of the water-stage and rainfall gauging stations that collected the

180   data that we used in this study. The monitoring data have been carefully collected and

181   processed. For full information and to download the available data, please refer to the

182   official website (http://wraew.ttfri.narl.org.tw/index.php).

183   TAPEX provides 72-hour rainfall forecasts for five rainfall gauges in the upstream

184   portion of the Yilan River Basin. The KW-GIUH model calculates the surface runoff

185   and estimates river flow at the Hsincheng and Yuanshan Bridges. This study uses the

186   POM and TPXO6.2 models to forecast the tides at Suao and to estimate the water stages

187   at the Kemalan Bridge. WASH123D then generates ensemble flow forecasts using

188   flows at the bridges mentioned above as the upstream boundary condition and the water

Hydrology and
Earth System
Sciences

Open Access

EGU

Discussions

189   stage at the Kemalan Bridge as the downstream boundary condition. The detailed

190   locations of these places are shown in Figure 2.

191   *3.2 Typhoon events*

192   Figure 3 shows the tracks of the different typhoons that have affected Taiwan,

193   according to historical records (Huang et al., 2012). Of the ten categories, Type-2 and

194   Type-3 typhoons account for approximately 28% of all typhoons and bring heavy

195   rainfall to the Yilan River Basin. For instance, a rainfall of 158 mm in 4 hours was

196   observed at rainfall gauging station C1U610 (shown in Figure 2) during Typhoon

197   Soulik. Table 1 shows all of the typhoons that invaded Taiwan from 2012 through 2015.

198   Five of these events are Type-2 and Type-3 typhoons, which have the biggest impact

199   on the Yilan River Basin. Therefore, this study selected the typhoons Saola (2012),

200   Soulik (2013), Soudelor (2015), and Dujuan (2015) to calibrate the HEPS and test its

201   performance. Typhoon Matmo, a Type-3 typhoon that occurred in 2014, was not

202   included due to its weak intensity. This study used historical observations of rainfall,

203   river stage, and tide to validate the parameters in the proposed HEPS.

204   4. A VISUALIZATION APPROACH FOR SUPPORTING THE

205   INTERPRETATION OF OPERATIONAL ENSEMBLE PEAK-

Hydrology and
Earth System
Sciences

Open Access

EGU

Discussions

206   FLOW FORECASTS DURING TYPHOON EVENTS

207   This study modified the 'Peak-Box' approach originally proposed by Zappa et al.

208   (2013) to provide better communication of HEPS forecasts during typhoon events.

209   Figure 4 compares the two approaches, and the modifications are described in detail

210   below. The purpose of the modifications is to develop a visualization approach that

211   simplifies the ensemble flow forecast information for use in formulating emergency

212   responses during typhoon events.

213   a. **Remove the horizontal and the vertical lines.** The horizontal and vertical lines

214   that indicate the medians of ensemble forecasts in the original 'Peak-Box' approach

215   are removed to prevent some information from being misused. Although

216   uncertainties exist in the HEPS, Pappenberger et al. (2013) noted a considerable

217   desire on the part of end-users to reduce probabilistic forecasts to deterministic

218   actions. The two lines may lead end-users to believe that the information provided

219   represents a single deterministic forecast, rather than a probabilistic one.

220   b. **Remove the outer rectangle.** In the original 'Peak-Box' approach, two rectangles

221   are displayed. The outer rectangle is the 'Peak-Box,' which highlights all

222   possibilities from the ensemble forecast, and the inner rectangle is the 'IQR-Box'

223   that emphasizes the 25th and 75th percentiles of the peak times and peak discharges

224   of the ensemble forecast. Zappa et al. (2013) argued that the outer rectangle

225   provides the forecaster with additional information. However, this argument does

226   not hold during typhoons, when the availability of too much data may obscure

227   critical information. Therefore, only one rectangle is shown in the study. This

228   rectangle indicates where the observed peak stage is likely to occur.

229   c.   **Use the mean and the standard deviation to define the rectangle.** This study

230   defines an 'SD-Box' that uses the mean (μ) and the standard deviation (σ), instead

231   of the first and third quartiles, to define the enveloping rectangle. As shown in the

232   right panel of Figure 4, the lower left coordinate of the 'SD-Box' is defined as the

233   mean peak time minus one standard deviation ($\mu_t - \sigma_t$) and the mean peak stage

234   minus one standard deviation $(\mu_h - \sigma_h)$ produced by all of the ensemble members.

235   The upper right coordinate is defined as the mean peak time plus one standard

236   deviation ($\mu_t + \sigma_t$) and the mean peak stage plus one standard deviation ($\mu_h +$

237   σ_h) of all of the ensemble members. In principle, the 'IQR-Box' should contain

238   25% (50% of the peak discharge times and 50% of the peak times) of all forecasts.

239   In practice, it contained from 12.5% to 37.5%, due to the distribution of ensemble

240   members (Zappa et al., 2013). Using the mean and the standard deviation (the 'SD-

241   Box') results in a larger area, includes 46.60% of the ensemble forecasts (68.27%

242   of peak water level times and 68.27% of the peak times) and has a greater chance

243   of including the observed peaks.

Hydrology and
Earth System
Sciences
Discussions

244   d. **Include all forecasts with different lead times in the rectangle.** Descriptive

245   statistics, such as the quartile deviation and the standard deviation, are susceptible

246   to outliers when calculated using insufficient sample sizes. Adding extra ensemble

247   members to produce more forecasts consumes computer resources. Yang et al.

248   (2016) showed that the performance of NWP models is independent of the length

249   of the lead time during typhoon events. Therefore, in order to expand the sample

250   size, this study includes present (*t*) and previous forecasts (*t-1, t-2, t-3... t-n*, where

251   *n* is the number of available forecasts when the system is initiated) to provide

252   ensemble flow forecasts. As shown in the right panel of Figure 4, the green area

253   illustrates the 'SD-Box'. The black and gray solid dots represent the current and

254   previous peak-flow forecasts, respectively.

5. RESULTS AND DISCUSSION

256   *5.1 Performance evaluation criteria*

257   This study applied two performance measures, the root mean square error (RMSE)

258   and the skill-spread ratio, to evaluate the proposed HEPS performance. For a well-

259   designed HEPS, the spread of ensemble forecasts will be large enough to cover the

260   prediction uncertainty. This statement implies that the spread should be the same as or

261   larger than the RMSE. The RMSE, which is commonly referred to as skill, measures

262   the difference between the observations and the ensemble mean without considering

263   the direction. The closer the RMSE is to zero, the better the ensemble mean is as a

264   forecast. The RMSE is defined as follows:

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

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

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

268   where μ is the ensemble mean of ensemble peak-flow forecasts; O_{peak} is the observation

269   of peak flow; $P_{peak,i}$ is the prediction of peak flow of the i_{th} member; m is the number of

270   ensemble members; and σ is the standard deviation of ensemble peak-flow forecasts.

271   The skill-spread score (hereinafter referred to as the score), which ranges from

272   zero to infinity, is the ratio of the standard deviation of the ensemble peak-flow forecasts

273   to the RMSE (Wilks, 2006). Scores less than one mean that the spread of the ensemble

274   forecasts is large enough to cover the prediction uncertainty. It is defined as follows:

275                           $$\text{Score} = \frac{RMSE}{\sigma}$$ (4)

276   *Model calibration and validation*

277   Two parameters in the proposed HEPS KW-GIUH model have been calibrated

278   using in situ observations made during typhoon events. These parameters are the

279   roughness coefficient for overland flow (n_0) and the roughness coefficient for channel

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

280   flow (n_c). The proposed HEPS used data from five rainfall gauges, including LTGX,

281   YSGZ, C1U610, C0U520 and C1U630 (see Figure 2 for locations), and the Thiessen

282   polygon method (Thiessen, 1911) to estimate the hourly spatial-average rainfall

283   intensities in order to provide rainfall input data to the KW-GIUH model. The

284   topographic data used in KW-GIUH are contained within a digital elevation model with

285   a resolution of 5 m obtained using aerial photographs. Kuo et al. (2016) used in situ

286   observations of flow discharges made at the Hsincheng and Yuanshan Bridges during

287   Typhoons Saola, Soulik, and Soudelor to calibrate the parameters of the KW-GIUH

288   model. Figure 5 shows that the percent errors in the peak discharges of the selected

289   typhoons were 4.59%, 2.07%, and -5.89% at the Hsincheng Bridge, and 14.88%, 5.28%,

290   and -3.05% at the Yuanshan Bridge, respectively. All of the errors in the peak times

291   were less than one hour. The results show that the KW-GIUH model is capable of

292   providing confident predictions for peak time, as well as peak discharge.

293   The WASH123D model adopted the most recent available cross-sectional

294   bathymetry of the Yilan River, which was measured in 2010, as its input topography

295   data. The upstream boundary of the model is set at the Hsincheng and Yuanshan Bridges,

296   and the downstream boundary of the model is set at the Kemalan Bridge. Field

297   measurements at the Hsincheng and Yuanshan Bridges from Kuo et al. (2016) and

298   observed water stages at the Kemalan Bridge were used as the upstream and

299   downstream boundary conditions, respectively. Field hourly records of water-stage at

300   the Zhongshan, Leawood, and Jhuangwei Bridges were used to calibrate the value of

301   Manning's roughness coefficient (n) in the WASH123D model and to validate the

302   performance of the model. Figure 6 shows that the percent errors in the peak stage for

303   Typhoons Saola, Soulik, and Soudelor, were 2.1%, 5.7%, and 10.6% at Zhongshan,

304   12.9% and 2.2% at Leawood, and 7.4%, 6.0%, and 2.1% at Jhuangwei, respectively.

305   There was one data gap at Leawood due to incomplete data collection during Typhoon

306   Soudelor. Nevertheless, all of the errors in the peak times were less than one hour. The

307   results show that WASH123D is capable of providing confident predictions of peak

308   times, as well as peak stages.

309   *5.3 Comparison of enveloping rectangles defined using the 'SD-Box' and the 'IQR-Box'*

310   *methods for supporting the interpretation of ensemble peak-flow results*

311   The proposed HEPS initiates when CWB issues a sea warning and ends when the

312   next ensemble forecast is six hours less than the left edge of the 'SD-Box'. In that regard,

313   93 forecasts are available for the four selected typhoons. Table 2 compares the forecast

314   peak stages and peak times between the 'SD-Box' and 'IQR-Box' approaches at the

315   Zhongshan, Leawood, and Zhuangwei Bridges. Scores were not calculated for the

316   Leawood Bridge during Typhoon Soudelor due to the lack of complete observations.

317   The scores that are less than one in the table are highlighted. These values indicate that

318   the spread of the ensemble members is large enough to contain the prediction

319   uncertainty. The rectangles defined using the 'IQR-Box' method contain 33.3% (31/93)

320   and 52.6% (49/93) of the observed peaks in stage and timing, respectively. Using the

321   'SD-Box' improves the capture rate to 51.6% (48/93) and 64.5% (60/93) for stage and

322   timing, respectively. Among all of the forecasts, there is only one forecast for which the

323   'IQR-Box' score is less than one, and the score of the 'SD-Box' is not. This situation

324   occurs at the Zhuangwei Bridge during Typhoon Soudelor. However, the score for the

325   'SD-Box' method is still very close to one (1.01), which means that it nearly captures

326   the observed peak. Overall, the 'SD-Box' method yielded average scores of 1.18 for the

327   peak stages and 1.08 for the peak times. In comparison with the 'IQR-Box' method,

328   which yielded scores of 2.06 for the peak stages and 2.06 for the peak times, the results

329   show that the enveloping rectangles defined using the 'SD-Box' method are more

330   reliable during typhoon events.

331   *5.4 Including all forecasts with different lead times during an event to expand the*

332   *sample size*

333   The sample size has a strong effect in terms of determining whether a result is

334   statistically significant. In other words, the number of available ensemble members is

335   important for both the 'SD-Box' and 'IQR-Box' methods. For example, the number of

336   available ensemble members for each forecast ranged from 11 to 14 for the proposed

337   HEPS during operation. Thus, the descriptive statistics were calculated using

338   insufficient sample sizes (less than 30). The same issue exists in other studies that

339   employ HEPSs (e.g., Yang and Yang, 2014; Zappa et al., 2013). It is difficult to increase

340   the number of ensemble members used in HEPSs, due to the limited computational

341   resources that are available. Therefore, this study proposes a method for including

342   present and previous forecasts in order to expand the sample size during the estimation

343   process.

344   It must be shown that the forecast performance is independent of time before all

345   available forecasts can be included in the estimation process. The time of concentration

346   of the peak flow at the Zhongshan Bridge is approximately 4 hours. This study

347   calculated the error in the maximum 4-hour rainfall between the average forecasts and

348   the average observations at the watershed upstream of the Zhongshan Bridge. Figure 7

349   shows that there is no obvious trend in the errors in stage and timing, regardless of the

350   length of the lead time. The correlation coefficients were -0.09 and 0.11, respectively,

351   and these values indicate that no significant correlations exist between errors in stage

352   or timing on the one hand and lead time on the other. For example, the best and worst

353   forecasts during Typhoon Dujuan in terms of stage error were the 1st and 5th forecasts,

354   respectively. However, the 6th forecast was better than the 5th, which implies that there

355   is no trend in the cascading forecasting process. Based on these results, this study

Hydrology and
Earth System
Sciences
Discussions

Open Access

356   assumed that the performance of the HEPS is independent of lead time during typhoon

357   events. Therefore, it is reasonable to include all available forecasts during an event to

358   expand the sample size.

359   Figure 8 illustrates the comparisons between using the 'SD-Box' method with one

360   forecast and using the 'SD-Box' method including all available forecasts (hereinafter

361   indicated as 'SD-Box Single' and 'SD-Box All') at the Zhongshan Bridge. The

362   performance of 'SD-Box All' was more consistent than that of 'SD-Box Single' in terms

363   of both stage and timing. For example, the scores for stage during Typhoon Soudelor

364   ranged from 0 to 5 when the 'SD-Box Single' method was used, but they were below

365   or close to 1 with 'SD-Box All'. The results showed that the inclusion of all available

366   forecasts in the calculation process decreased the variation among the forecasts; in other

367   words, the uncertainty of the forecasts decreased. Figure 9 illustrates the scores of all

368   of the forecasts for the different typhoon events. The 'SD-Box Single' contained 47.1%

369   of the observed peaks in terms of stage (37.3% + 9.8%), whereas 'SD-Box All'

370   contained 63.7% (57.8% + 5.9%) of the observed peaks. Furthermore, the 'SD-Box

371   Single' contained 58.9% (37.3% + 21.6%) of the observed peaks in terms of timing,

372   whereas 'SD-Box All' contained 71.5% (57.8% + 13.7%). The results show that the

373   'SD-Box All' method can capture more of the observed peaks in terms of both stage

Hydrology and
Earth System
Sciences
Discussions

Open Access

374   and timing. In particular, 'SD-Box All' improved the forecast performance and

375   increased the capture rate from 37.3% to 57.8% for both stage and timing.

6. CONCLUSIONS

377   This study proposed a HEPS that employs NWP models to perform rainfall

378   forecasts and hydrologic models to produce ensemble flood forecasts during typhoon

379   events. Because the communication of ensemble forecasts is critical for helping end-

380   users to respond, a modified version of the 'Peak-Box' visualization method, which was

381   originally described by Zappa et al. (2013), was also proposed to support the

382   interpretation of ensemble forecast results for operational purposes. Four typhoon

383   events during the period 2012-2015 and observations collected in the Yilan

384   Experimental Watershed were used to evaluate the performance of these techniques. A

385   total of 93 forecasts and two performance measures were considered. The results

386   showed that the proposed HEPS is able to provide flood forecasts during the selected

387   typhoon events. In addition, the 'SD-Box' visualization approach, which considers the

388   mean and the standard deviation instead of the 25th and 75th percentiles, captured more

389   of the observed peaks during typhoon events. The average skill-spread scores of the

390   'SD-Box' method for the selected events were 1.18 and 1.08 in terms of stage and

391   timing, respectively. These results represent a significant improvement over the original

392   'Peak-Box' method, which resulted in scores of 2.06 for both peak stage and peak

393   timing. Scores of less than one indicate that the spread of the ensemble forecasts is large

394   enough to contain the prediction uncertainty. Since the average score achieved by the

395   'SD-Box' method was close to one, it has been shown to be more reliable than the

396   original 'Peak-Box' method during typhoon events. The results satisfy the statement

397   "One of the main objectives of ensemble flood forecasts is the representation of the full

398   spectrum of forecast uncertainty and/or predictability in [the] form of different

399   hydrological responses to the input of the various members obtained from an

400   atmospheric EPS" made by Zappa et al. (2013).

401   Descriptive statistics, such as the quartile deviation and the standard deviation, are

402   susceptible to outliers when calculated using an insufficient number of observations.

403   Adding more ensemble members is expensive in terms of computer resources. This

404   study proposed a method that enables increasing the sample size, leading to statistically

405   significant results. This method involves including present and previous available

406   forecasts in the calculation process. For example, the proposed HEPS generated 11

407   available ensemble members at each forecast during Typhoon Dujuan. By including all

408   of the present and previous available forecasts (the 'SD-Box All' method), the sample

409   size increased to 22 for the second forecast, 33 for the third forecast, and so on. The

410   results showed that the 'SD-Box All' made more consistent predictions. This result can

411   be explained by the inclusion of all available forecasts in the calculation process

412   decreasing the uncertainty of the forecasts. As a result, the rectangles defined by the

413   'SD-Box All' method contained 57.8% of the observed peaks in stage and timing.

414   Coughlan de Perez et al. (2016) suggested that a HEPS that produces a false alarm rate

415   below 50% is tolerable for decision makers in terms of the economic and practical

416   consequences of taking action. However, this study assumed that the forecast

417   performance of the proposed HEPS is independent of the length of the lead time and

418   conducted an experiment to prove it. Other studies, such as that of Zappa et al. (2013),

419   have claimed that the most accurate forecasts were obtained for lead times of two or

420   more days. Such statements imply that the performance of HEPSs do not improve with

421   shorter lead times or are independent of lead time, and Yang et al. (2016) found that the

422   best performance is obtained before a typhoon makes landfall. This assumption is still

423   susceptible to the topography of the applied area and the type of extreme event being

424   considered. Further investigation of various conditions must be performed before firm

425   conclusions can be drawn. Regardless, the proposed HEPS and the modified

426   visualization approach have been shown to produce convincing peak-stage and peak-

427   timing forecasts for operational purposes during a typhoon.

428   AUTHOR CONTRIBUTION

429   Ya-Chi Chang, Mei-Ying Lin, Jui-Yi Ho calibrated and verified the parameters of

430   WASH123D, POM and KW-GIUH models. Cheng-Hsin Chen dealt with the data

431   processing of the models and performed the simulations. Sheng-Chi Yang and Ya-Chi

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

432   Chang analyzed the results of HEPS and developed a new approach for improved

433   interpretation during typhoon events. Sheng-Chi Yang, Tsun-Hua Yang, Ya-Chi Chang

434   and Kwan-Tun Lee prepared the manuscript with contributions from all co-authors.

443   REFERENCES

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

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

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

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

456   Cloke, H. L. and Pappenberger, F.: Ensemble flood forecasting: A review, J. Hydrol.,
457   375, 613-626, 2009.

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

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

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

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

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

471   Frick, J. and Hegg, C.: Can end-users' flood management decision making be improved
472   by information about forecast uncertainty?, Atmospheric Research, 100, 296-303,
473   2011.

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

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

481   Hsiao, L. F., Yang, M. J., Lee, C. S., Kuo, H. C., Shih, D. S., Tsai, C. C., Wang, C. J.,
482   Chang, L. Y., Chen, D. Y. C., and Feng, L.: Ensemble forecasting of typhoon
483   rainfall and floods over a mountainous watershed in Taiwan, J. Hydrol., 506, 55-
484   68, 2013.

485   Huang Jr, C., Yu, C. K., Lee, J. Y., Cheng, L. W., Lee, T. Y., and Kao, S. J.: Linking
486   typhoon tracks and spatial rainfall patterns for improving flood lead time
487   predictions over a mesoscale mountainous watershed, Water Resources Research,
488   48, 2012.

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

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

Hydrology and
Earth System
Sciences
Discussions

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

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

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

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

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

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

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

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

514   Palmer, T. N.: A nonlinear dynamical perspective on model error: A proposal for non-
515   local stochastic-dynamic parametrization in weather and climate prediction
516   models, Quarterly Journal of the Royal Meteorological Society, 127, 279-304,
517   2001.

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

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

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

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

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

536   Rossa, A., Liechti, K., Zappa, M., Bruen, M., Germann, U., Haase, G., Keil, C., and
537   Krahe, P.: The COST 731 Action: A review on uncertainty propagation in
538   advanced hydro-meteorological forecast systems, Atmospheric Research, 100,
539   150-167, 2011.

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

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

546   Thiessen, A. H.: Precipitation averages for large areas, Monthly Weather Review, 39,
547   1082-1084, 1911.

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

552   Todini, E.: Predictive uncertainty assessment in real time flood forecasting. In:
553   Uncertainties in Environmental Modelling and Consequences for Policy Making,
554   Baveye, P. C., Laba, M., and Mysiak, J. (Eds.), Springer, Dordrecht, Netherlands,
555   2009.

556   Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Elsevier, Amsterdam,
557   2006.

558   Yang, S. C. and Yang, T. H.: Uncertainty assessment: Reservoir inflow forecasting with

559   ensemble precipitation forecasts and HEC-HMS, Advances in Meteorology, 2014,
560   2014.

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

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

568   Yeh, G. T., Huang, G. B., Zhang, F., Cheng, H. P., and Lin, H. C.: WASH123D: A
569   numerical model of flow, thermal transport, and salinity, sediment, and water
570   quality transport in watershed systems of 1-D stream-river network, 2-D overland
571   regime, and 3-D subsurface media. Technical Rep. submitted to EPA, Dept. of
572   Civil and Environmental Engineering, 2006.

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

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

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

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

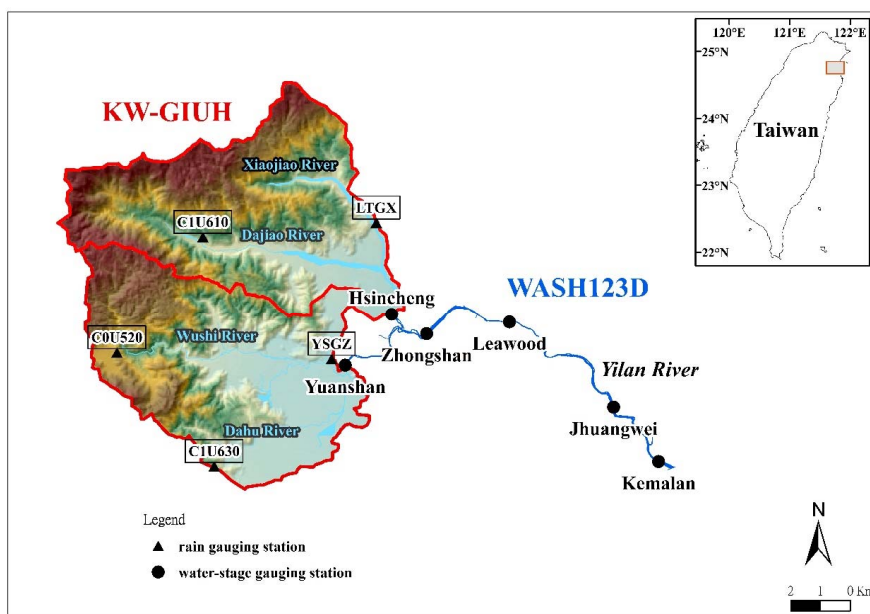586   FIGURES
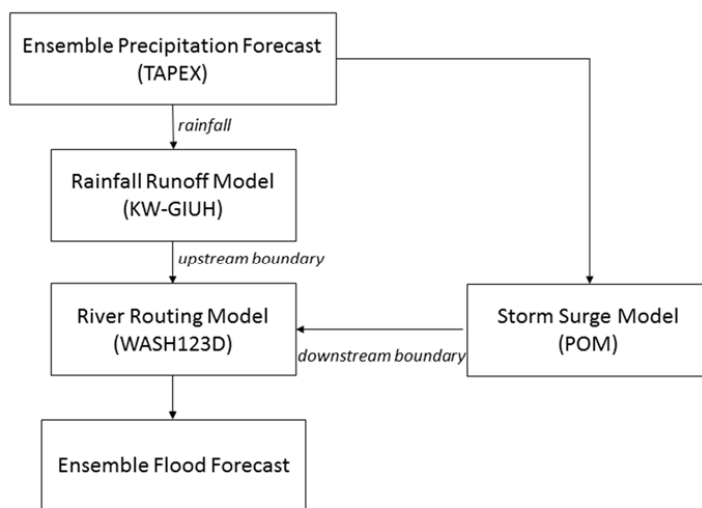
587

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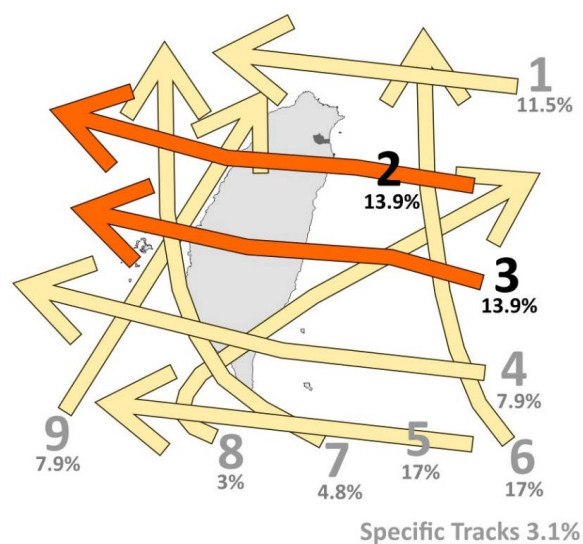589   HEPS.

590

591

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.

595

Redrawn from Kuo et al. (2012)

596   **Figure 3** Schematic diagram showing the tracks of typhoons invading Taiwan. The

597   percentages shown in the figure are the statistical results from 1958 through 2006

598   obtained from the Central Weather Bureau (CWB). The dark gray polygon located in

599   northern Taiwan indicates the Yilan River catchment. Type-2 and Type-3 typhoons

600   bring heavy rainfall to the Yilan River catchment.
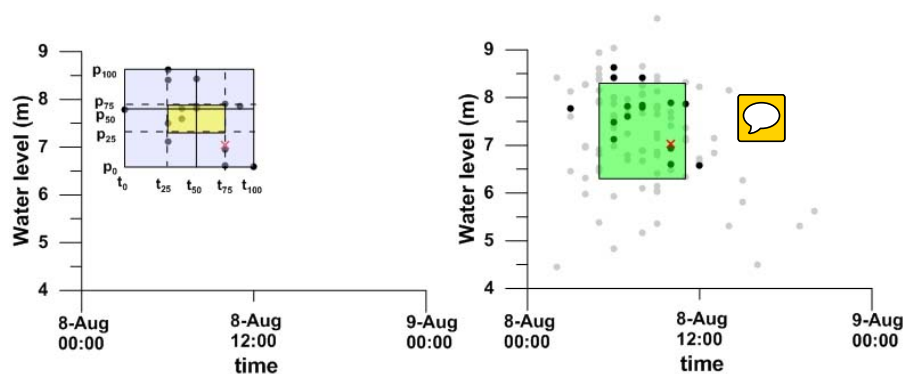
601

602 **Figure 4** The left panel shows a graphical explanation of the 'Peak-Box' approach.

603 The outer rectangle is the 'Peak-Box,' and the internal rectangle (the yellow area) is

604 the 'IQR-Box'. The solid dots represent all of the ensemble forecasts. The right panel

605 shows a graphic explanation of the proposed extension of the 'Peak-Box' approach.

606 The enveloping rectangle is the 'SD-Box' (the green area). The solid black and gray

607 dots represent current and previous peak-flow forecasts, respectively.
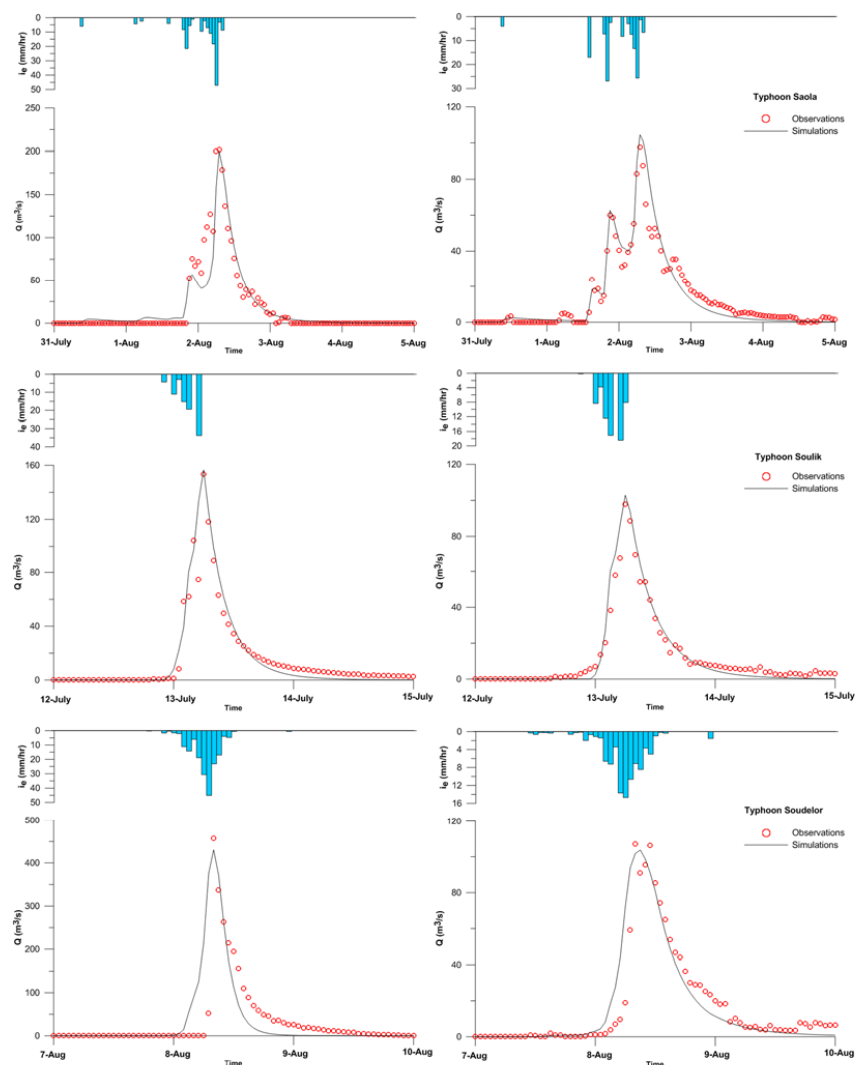
608   **Figure 5** Comparison of simulated discharges (red circles) and recorded discharges

609   (solid lines) for model calibration (Typhoons Saola and Soulik) and validation

610   (Typhoon Soudelor) experiments at Hsinsheng (left) and Yuanshen (right). The blue

611   bars are the hourly spatial-average rainfall intensities measured in the watershed

612   upstream of Hsinsheng and Yuanshen.

33

613
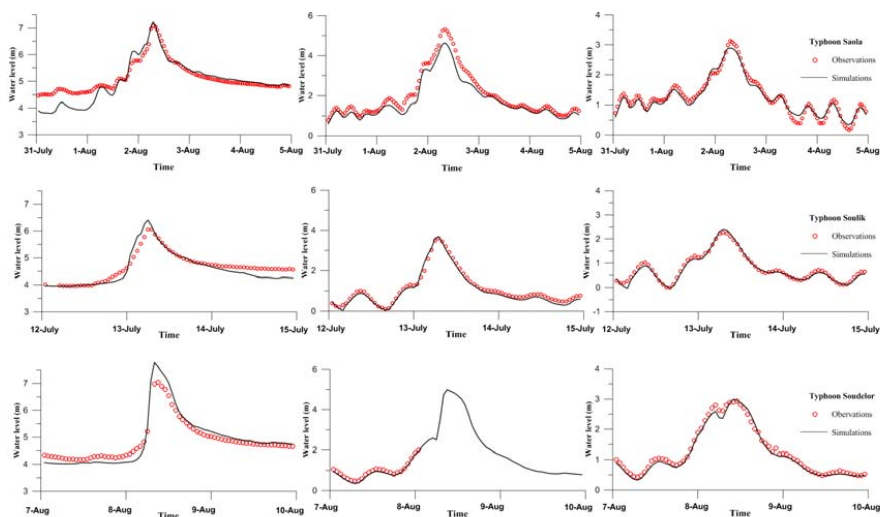
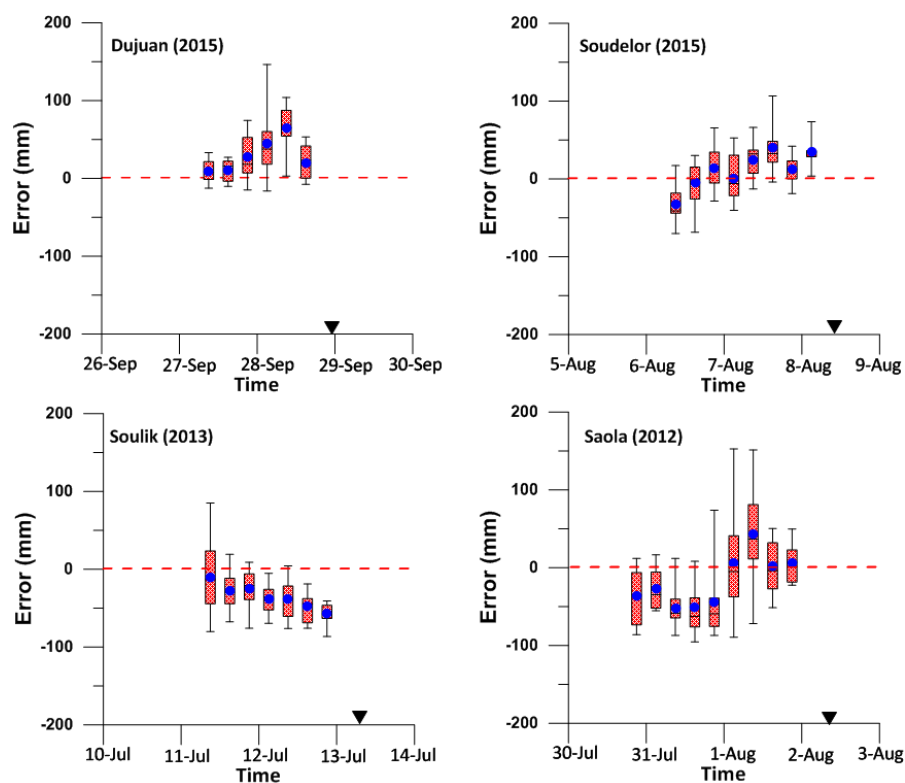**Figure 6** Comparison of simulated (red circles) and recorded (solid lines) water levels

for model calibration (Typhoons Saola and Soulik) and validation (Typhoon Soudelor)

experiments at Zhongshan (left), Leawood (central) and Jhungwei (right).

Hydrology and
Earth System
Sciences
Discussions

Open Access

(a) Magnitude error of maximum 4-hour rainfall

(b) Timing error of maximum 4-hour rainfall
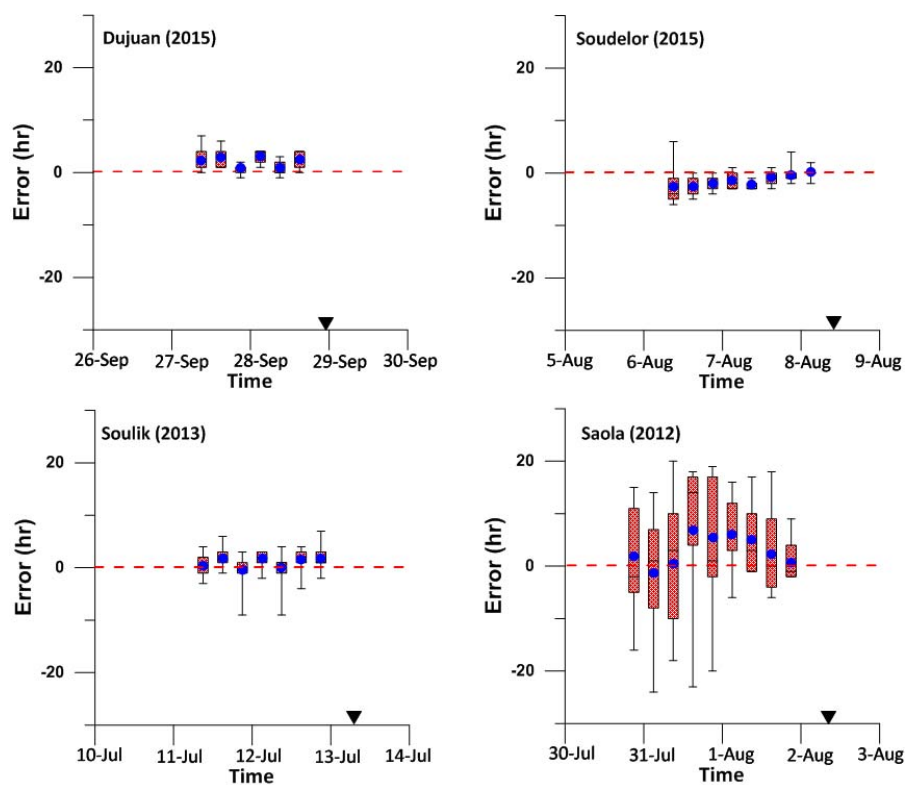
617 **Figure 7** Box-and-whisker plot at the watershed upstream of the Zhongshan Bridge

618 during the four selected typhoon events. The blue dots indicate the ensemble means.

619 The inverted triangles indicate the time of occurrence of the maximum 4-hour rainfall.

620 The results show that there is no obvious trend in lead time for the errors in either the

621 stage or timing.

(a) Scores for peak-stage forecasts
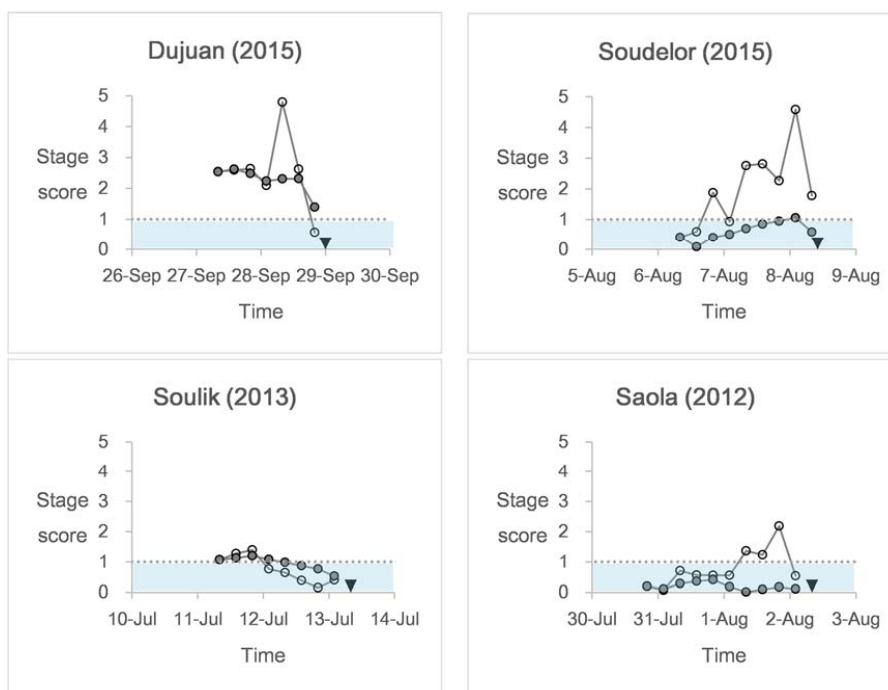
(b) Scores in peak-timing forecasts
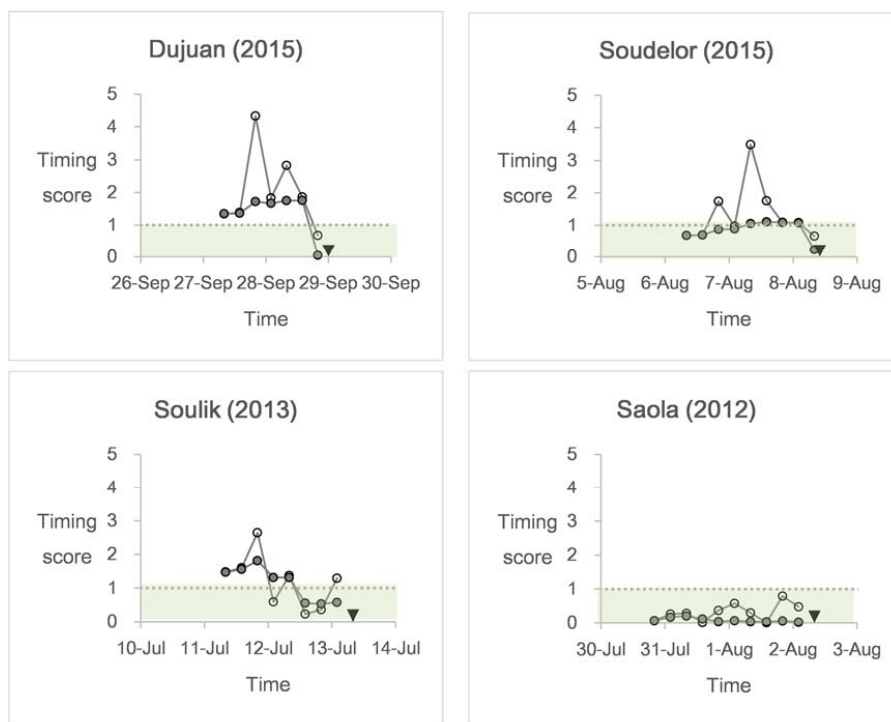
622   **Figure 8** The scores of the single ('SD-Box Single') and accumulating ('SD-Box All')

623   methods at the Zhongshan Bridge during the four selected typhoon events. The inverted

624   triangles indicate the time of occurrence of the observed peak stage. The results show

625   that the performance of the 'SD-Box All' method (solid circles) was more stable than

626   that of the 'SD-Box Single' method (open circles) in terms of both stage and timing.

627

628
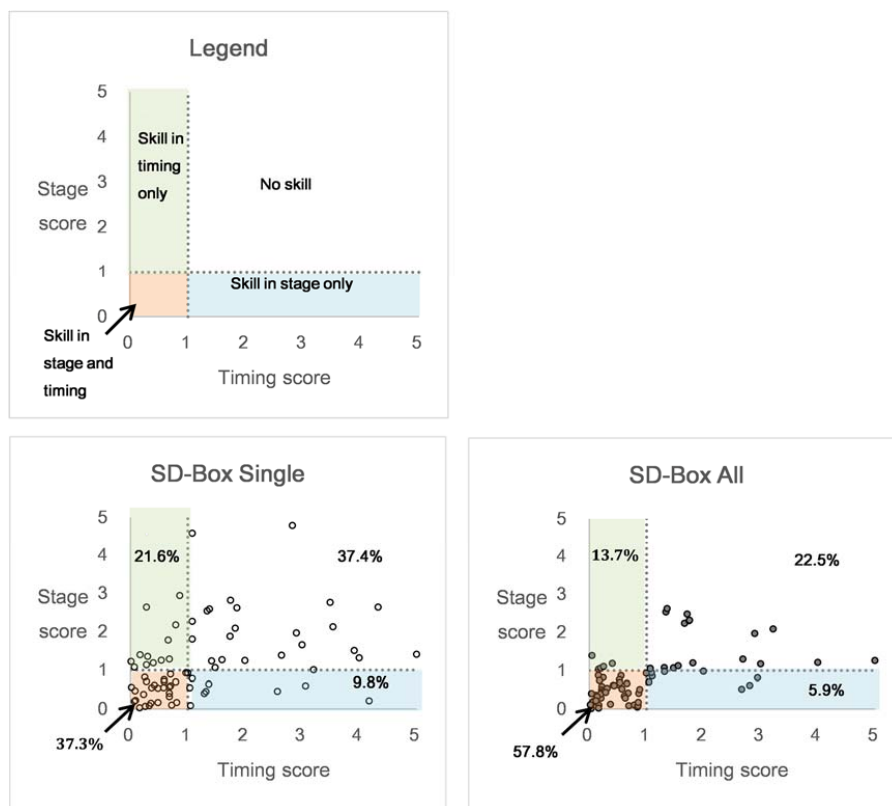
Figure 9 Comparison of scores obtained for 'SD-Box Single' and 'SD-Box All'. The

results show that the 'SD-Box All' approach significantly improves the performance

compared with the results obtained using the 'SD-Box Single' method.

Hydrology and
Earth System
Sciences
Open Access
Discussions

EGU

632   TALBLES

633   **Table 1**   All typhoons that invaded Taiwan during 2012 through 2015. A total of four

634   typhoons of Type-2 and Type-3, namely, Saola in 2012, Soulik in 2013, Soudelor in

635   2015, and Dujuan in 2015, were selected to calibrate the system and test the

636   performance in this study. Typhoon Matmo, a Type-3 typhoon that occurred in 2014,

637   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

638   (Source: Central Weather Bureau, Taiwan)

40

639 **Table 2** Comparisons of scores in peak stage and peak time between the 'IQR-Box'

640 and 'SD-Box' approaches. Scores less than one (highlighted) indicate that the

641 enveloping rectangle did contain the observed peak.

642 **(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

643

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