

# Pollutant first flush identification and its implications for urban runoff pollution control: a roof and road runoff case study in Beijing, China

Wei Zhang, Juan Li, Huichao Sun and Wu Che

## ABSTRACT

First flush is a common phenomenon in urban runoff pollution. Typical cement roof and asphalt road runoff in Beijing, China were monitored for 2 years. Based on the  $M(v)$  curve, the suspended solids (SS), chemical oxygen demand (COD), total phosphorus (TP) and particulate phosphorus in cement roof runoff presented a stronger first flush than those in asphalt road runoff. The first flush volume ( $V_{FF}$ ) of SS, COD, total nitrogen (TN) and TP in asphalt road runoff differed slightly from the cement roof. There were also differences in the first flush assessment depending on which method was used. We proposed a new method based on the runoff depth versus pollutant cumulative mass curve. According to the national standards in China ( $V_{FF} = 3$  mm), various masses of different pollutants, such as  $91.42 \pm 9.80\%$  (cement roof) and  $78.49 \pm 19.41\%$  (asphalt road) of SS and  $86.85 \pm 13.54\%$  (cement roof) and  $72.80 \pm 25.79\%$  (asphalt road) of COD, can be effectively controlled, but our mass control efficiencies were 55.91%–66.65% when  $V_{FF} = 1$  mm. The new method proposed in this study provides an alternative approach for assessing runoff pollution control efficiency of different  $V_{FF}$ .

**Key words** | first flush, road runoff, roof runoff, runoff pollution, stormwater management

## HIGHLIGHTS

- SS, COD, TP and PP exhibited a stronger first flush in cement roof runoff than in asphalt road runoff by  $M(v)$  method.
- $V_{FF}$  of SS, COD, TN and TP in asphalt road runoff differed slightly from the cement roof.
- Differences were found in the first flush assessment results between  $M(v)$  and  $V_{FF}$  methods.
- A new method based on the runoff depth versus pollutant cumulative mass curve was proposed.

## INTRODUCTION

With rapid urbanization, the issue of urban runoff pollution has become increasingly serious because it has detrimental

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

doi: 10.2166/wst.2021.157

effects on urban water quality (Perera *et al.* 2019). Urban runoff contains conventional pollutants, such as suspended solids (SS), chemical oxygen demand (COD), total nitrogen (TN) and total phosphorus (TP), as well as trace pollutants, such as heavy metals and polycyclic aromatic hydrocarbons (PAHs) (Sansalone & Cristina 2004). Heavy metals and PAHs are mostly bound to particulate matter and exist in

**Wei Zhang** (corresponding author)

Juan Li

Huichao Sun

Wu Che

Key Laboratory of Urban Stormwater System and Water Environment, Ministry of Education, Beijing University of Civil Engineering and Architecture, 100044 Beijing, China

E-mail: [weizhang021@163.com](mailto:weizhang021@163.com)

**Wei Zhang**

Wu Che

Beijing Engineering Research Center of Sustainable Urban Sewage System Construction and Risk Control, Beijing University of Civil Engineering and Architecture, 100044 Beijing, China

**Wei Zhang**

Beijing Advanced Innovation Center for Future Urban Design, 100044 Beijing, China

particulate form in urban runoff; there is a significant correlation between these pollutants and particulate matter content in runoff (Winston & Hunt 2017). The build-up and wash-off characteristics of particulate matter in runoff are important in understanding overall pollution characteristics of heavy metals, PAHs and other trace pollutants in urban runoff.

The first flush is a common phenomenon in urban runoff pollution, and it was first introduced in the 1970s (Sartor & Boyd 1972). The  $M(v)$  curve, which is a curve of the dimensionless cumulative pollutant mass  $M$  as a function of the dimensionless cumulative runoff volume  $V$ , was used to assess first flush strength (Geiger 1984). Stahre & Urbonas (1990) proposed that first flush could be observed if at least 80% of the total pollutant mass is transported in the first 20% of runoff volume during a rainfall event. Subsequently, different proportions of total pollutant mass to runoff volume (e.g., 25%/50% and 30%/80% for volume/mass, respectively) were proposed along with other criteria (Wanielista & Yousef 1993; Bertrand-Krajewski *et al.* 1998). In addition, the mass first flush ratio (MFFR) based on the  $M(v)$  curve was proposed (e.g., MFFR<sub>20</sub>, MFFR<sub>30</sub>: the percentage of pollutant mass is contained in the first 20% or 30% of the runoff volume.) and has been widely used in research (Han *et al.* 2006; Barco *et al.* 2008; Jeung *et al.* 2019).

Although the first flush strength can be quantitatively described by the MFFR, there remains some limitations. First, the MFFR is dimensionless; therefore, it disregards the impact of storm volume and the total volume of a smaller event may only be as much as the first 20% of volume of a larger event. Thus, a smaller event could be contained within a first flush and mistakenly counted as such (Bach *et al.* 2010). Second, traditional methods characterize the first flush solely on changes in pollutant loading within an event and not within the overall background levels of a catchment area (Bach *et al.* 2010). The purpose of first flush assessment is to determine the first flush volume ( $V_{FF}$ ), which is usually expressed in millimeters of rainfall depth; this assessment can guide runoff pollution control facilities. The  $M(v)$  curve cannot provide this type of information and guidance. Thus, an innovative first flush assessment method was proposed to address the deficiencies of traditional methods (Bach *et al.* 2010). In the new method,  $V_{FF}$  is defined as the amount of runoff when the last background concentration reading is initiated. The difference between the higher runoff pollution concentrations during the suspected first flush and background pollutant concentrations are then analyzed to confirm and quantify the first flush. This method means that  $V_{FF}$  can be used in the

design of pollution control facilities. Todeschini *et al.* (2019) used this method to analyze monitoring data for Lombardy, Italy for the first flush. One disadvantage of this method is that it requires a high amount of rainfall monitoring data. Furthermore, whether the  $V_{FF}$  determined by this method is consistent with the first flush volume standards in different countries needs further study.

An accurate characterization of the first flush is a concern in many countries because runoff pollution control based on the first flush is required in many local or national standards or specifications. For example, in Italy, design criteria for wet-weather detention tanks impose storage capacities of 25–50 m<sup>3</sup> per contributive impervious hectare, corresponding to 2.5–5 mm of runoff (Todeschini *et al.* 2019). The UK Highways Agency recommends treating at least 10 mm of runoff on the basis that it is mostly first flush (UK Highways Agency 2006). Similarly, a minimum water quality treatment volume of 12.7 mm has been used in the United States (Sansalone & Cristina 2004).

Urban runoff pollution research has been an important topic since the 1990s, and the first flush has received more attention with the continuous advancements of sponge city construction initiatives. Currently, first flush volume requirements have also been incorporated into national standards in China. The latest edition of ‘*Technical code for rainwater management and utilization of building and sub-district*’ (GB 50400-2016), which was released in 2016, required that ‘*the first flush volume should be determined by monitoring the pollutant concentration of urban runoff. When the conditions are not met, the first flush volume of roof runoff can be assumed to be 2–3 mm, and the first flush volume of road runoff can be assumed to be 3–5 mm.*’

Although there were obvious differences in first flush standards in different countries, the approaches to handle the first flush present similarities. The first flush volume storage by a stormwater or rainwater tank was the commonly used approach. Additionally, a first flush divider was also normally used to divide the first flush into sewage system, and then transported to wastewater treatment plant. However, there are open questions regarding these standards, such as whether runoff pollution is effectively controlled if the first flush volume is treated according to the stated requirements and whether a generalized first flush volume is applicable in different regions, different land use districts and different types of road and roof surfaces.

To address one aspect of these open questions, we monitored typical cement roof (CR) and asphalt road (AR) runoff in Beijing, China to assess the first flush and determine  $V_{FF}$ . The objectives of this study were to (a) investigate

differences in pollution characteristics between CR and AR runoff, (b) evaluate the first flush strength based on the  $M(v)$  curve and  $V_{FF}$  and (c) quantify runoff pollution control performance for different  $V_{FF}$  values.

## MATERIALS AND METHODS

### Study site

We used a CR and AR located at the Beijing University of Civil Engineering and Architecture campus in this study. The campus is located between the 5th and 6th ring roads (39.4456°N, 116.1639°E) in Beijing. The CR was built in 2014, and is a typical building roof type in Beijing. It was 7.5 m tall and it consists of two floors. Runoff samples were collected from the downspout, and the roof catchment area is about 60 m<sup>2</sup>. The AR was also built in 2014, and is a commonly used pavement type in Beijing. Runoff samples were collected from the gully along the curb; the road has a catchment area of about 64 m<sup>2</sup>. The AR is swept twice a day by an electric road sweeper (Mingnuo MN-S1800). The road sweeping is the daily sanitation work of this university campus, and the frequency of twice a day was normal practice in Beijing to meet the environmental protection requirement. Additionally, the electric road sweeper has been widely used in road sweeping operations in Beijing. The selected road with regular sweeping that is permanent represents the real case. The monitored road at the campus was a two-way road of two lanes with a width of 5 m. The average daily traffic of AR was approximately 700 vehicles/day taken in 2019 by two-way traffic count. The distance between the CR and the AR was less than

100 m, and the rainfall characteristics of the monitoring sites were assumed to be the same.

### Rainfall characteristics

A tipping bucket rain gauge with 0.2 mm accuracy (HOBO U30 Station, Onset) was installed on the roof of a nearby building. The average annual rainfall in Beijing is 644 mm, with 80% occurring in May–September. Rainfall runoff data were collected from July 2019 to September 2020 with some interrupted periods due to instrumentation failure. Several rainfall events were excluded because of equipment failure or insufficient precipitation that limited sampling. In total, we monitored eight rainfall events, which spanned a wide range of precipitation depths, duration, intensity and antecedent dry periods (ADP) (Table 1).

### Field sampling and testing

Runoff from the CR and AR was sampled manually, according to protocols outlined in Burton & Pitt (2002). The runoff samples were recovered immediately following a rainfall event and transported to the laboratory for analysis. Samples were tested for SS, COD, turbidity, TN, ammonia (NH<sub>4</sub><sup>+</sup>-N), nitrate (NO<sub>3</sub><sup>-</sup>-N), TP, dissolved phosphorus (DP) and particulate phosphorus (PP) using standard methods (APHA 2012). The detailed sampling and water quality analytical procedures are provided in the Supplementary Materials. A Malvern Mastersizer 3,000 (Malvern Instruments Ltd, UK) with a particle size resolution from 0.01 to 3,500 μm was used to determine the particle size distribution within 6 h of collection. Two velocity-area flowmeters (NIVUS PCM F, Eppingen, Germany) were installed at the

**Table 1** | Characteristics of the monitored rainfall events

No.	Date of rain event YY-MM-DD	Duration (min)	Rainfall (mm)	Maximum intensity (mm/min)	Average intensity (mm/min)	Antecedent dry period (d)
1	2019-07-05	96	8.2	0.4	0.085	18.4
2	2019-07-22	264	20.0	1.0	0.076	0.2
3	2019-07-28	191	19.2	0.8	0.101	6.0
4	2019-08-20	221	3.2	0.2	0.014	7.0
5	2020-08-12	22	2.4	0.4	0.109	2.6
6	2020-08-18	41	1.8	0.2	0.044	0.3
7	2020-08-23	645	63.8	0.8	0.099	5.1
8	2020-09-01	110	12.4	1.2	0.113	4.9
Average		198.8 ± 186.6	16.4 ± 19.2	0.6 ± 0.4	0.080 ± 0.032	5.6 ± 5.4

catchment outlets to monitor flow rate and runoff volume at 5-min intervals during rainfall events.

### First flush analysis

The  $M(v)$  curve and the MFFR were used to assess the first flush strength (Deletic 1998; Lee & Bang 2000; Jeung *et al.* 2019) and the analytical procedures are detailed in the Supplementary Materials. The first flush volume was determined using a method proposed in previous studies (Bach *et al.* 2010; Todeschini *et al.* 2019), which includes five steps. Briefly, Step 1 was the determination of runoff depth, and we used a slice size of 1 mm; Step 2 involved calculating the mean pollutant concentrations; Step 3 described the pollutant concentration distribution in each slice; Step 4 was a test for differences between two slices; Step 5 led to the quantification of the first flush volume. Additional details are provided in the Supplementary Materials.

### Data analysis

The event mean concentration (EMC) was used to analyze runoff pollution (Lee *et al.* 2002); the EMCs of pollutants in AR and CR are provided in the Supplementary Materials. One-way analysis of variance (ANOVA) was used to determine significant differences (accepted  $p > 0.05$ ).

## RESULTS AND DISCUSSION

### Event mean concentrations

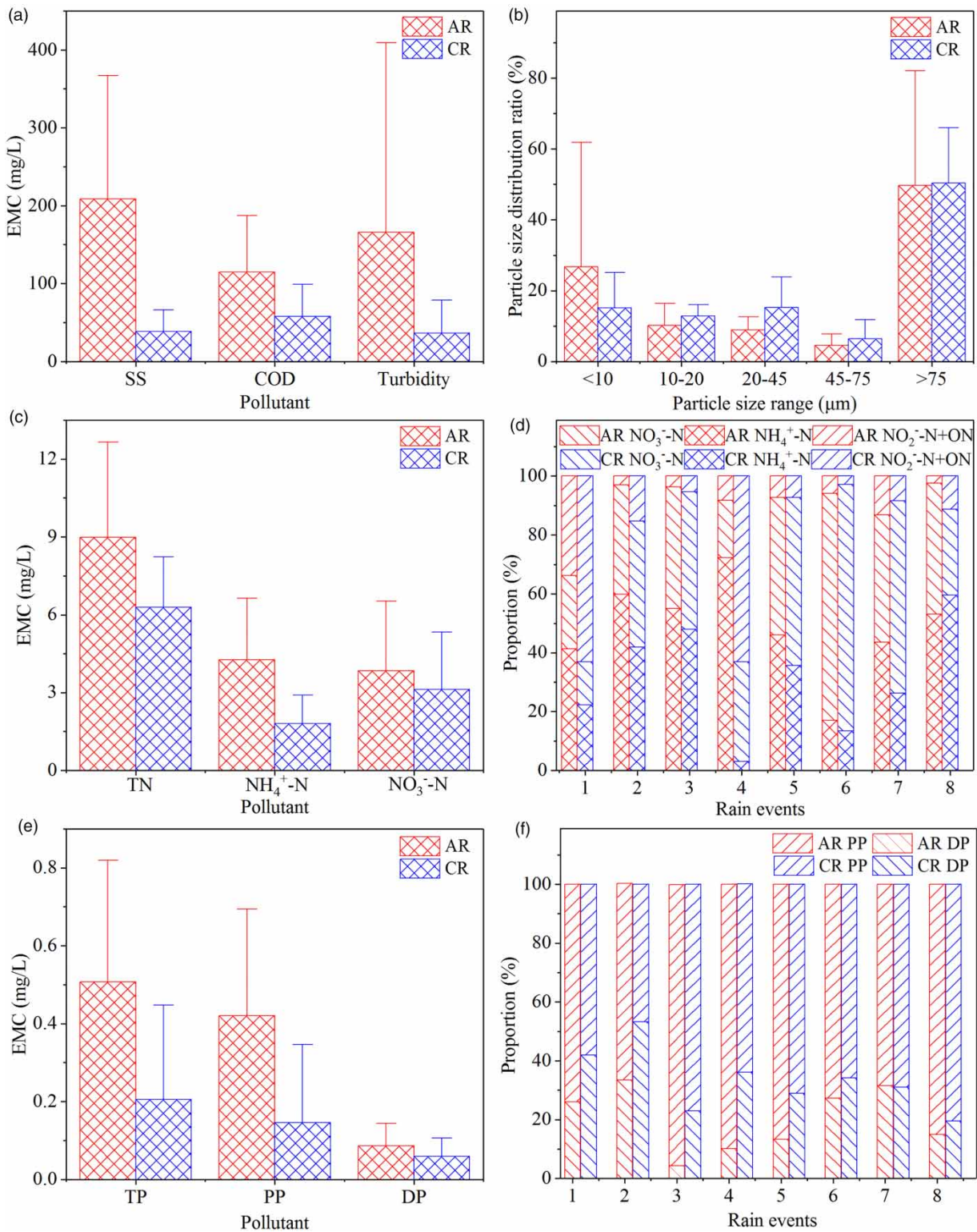
The EMCs of pollutants in CR and AR are illustrated in Figure 1. The EMCs of SS ( $208.00 \pm 148.02$  mg/L) and COD ( $114 \pm 68$  mg/L) in AR runoff were significantly higher than in CR runoff (SS:  $38.54 \pm 25.94$  mg/L; COD:  $58 \pm 38$  mg/L) ( $p < 0.05$ ). The EMCs of turbidity, nitrogen and phosphorus pollution were also significantly higher in AR runoff (Figure 1(a), 1(c), 1(e)) ( $p < 0.05$ ). Under the same rainfall characteristics, runoff pollution was related to the surface type, and the differences in runoff pollution sources accounted for the differences in runoff pollutant EMCs (Hou & Zhang 2014). The main pollution sources of the CR runoff were atmospheric deposition, wet deposition, roof material degradation and bird droppings (Egodawatta *et al.* 2009), and those for AR runoff included vehicle exhaust, tire wear, lubricants, road abrasion, rust, rubbish and solid waste (Hou & Zhang 2014). Based on

the EMCs of typical pollutants in both types of runoff, the runoff pollution level monitored in this study was consistent with previous results reported for Beijing (Zhao *et al.* 2010; Zhang *et al.* 2012; Ren *et al.* 2013; Chen *et al.* 2018). It was worth noting that the COD,  $\text{NH}_4^+\text{-N}$ , TN and TP in AR runoff and the COD,  $\text{NH}_4^+\text{-N}$ , TN in CR runoff much exceed the Class V surface water quality standard, 'Environmental quality standards for surface water (GB3838-2002)'. These roofs and roads runoff discharge will seriously threaten the environmental quality of surface water.

The particle size distribution of the CR and AR runoff is shown in Figure 1(b). Particles larger than  $75 \mu\text{m}$  were most abundant in CR and AR runoff and accounted for  $50.42 \pm 15.57\%$  and  $49.72 \pm 32.35\%$  of the total, respectively; there were no significant differences between CR and AR runoff ( $p > 0.05$ ). These results indicate that particles  $>75 \mu\text{m}$  are the main components in both CR and AR runoff. This finding is beneficial for runoff pollution control practices because if particles larger than  $75 \mu\text{m}$  can be effectively removed, then approximately half of all particles in runoff would be eliminated.

Nitrogen in runoff exists as  $\text{NH}_4^+\text{-N}$ ,  $\text{NO}_3^-\text{-N}$ , nitrite ( $\text{NO}_2^-\text{-N}$ ) and organic nitrogen (ON) with  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_3^-\text{-N}$  generally considered the main forms (Taylor *et al.* 2005; Lusk *et al.* 2020). For the CR and AR runoff monitored in this study, the sum of  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_3^-\text{-N}$  accounted for  $77.94 \pm 23.95\%$  (CR) and  $90.32 \pm 9.67\%$  (AR) of nitrogen forms (Figure 1(d)), consistent with previous studies. However, there were significant differences in the sum of  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_3^-\text{-N}$  between CR and AR runoff ( $p < 0.05$ ). The proportion of  $\text{NH}_4^+\text{-N}$  ( $31.27 \pm 17.44\%$ ) in CR runoff was significantly lower than that in AR runoff ( $48.53 \pm 15.15\%$ ) because of differences in  $\text{NH}_4^+\text{-N}$  pollution sources of the CR and AR runoff.

Proportions of PP in the CR and AR runoff were  $66.55 \pm 9.44\%$  and  $79.86 \pm 10.04\%$ , respectively, and the phosphorus overall was mainly particulate. As a result of differences in pollutant sources, the proportion of PP in AR runoff was higher than that in CR runoff (Hou & Zhang 2014). In addition to the pollution source, rainfall characteristics are also a factor affecting the forms of phosphorus found in runoff. For example, PP accounted for 47% of CR runoff during the second rainfall event because the ADP was only 0.2 d; thus, the PP had been fully washed out in the previous event. The PP proportion in AR runoff during the same rainfall event was also relatively minimal, which indicates that ADP has a strong influence on PP proportion.



**Figure 1** | The event mean concentrations (EMCs) of runoff pollutants (suspended solids (SS), chemical oxygen demand (COD), turbidity, total nitrogen (TN), ammonia ( $\text{NH}_4^+\text{-N}$ ), nitrate ( $\text{NO}_3^-\text{-N}$ ), total phosphorus (TP), particulate phosphorus (PP) and dissolved phosphorus (DP)), particle size distribution, nitrogen and phosphorus forms in asphalt road (AR) and cement roof (CR) runoff.

## *M(v)* curve

The *M(v)* curves of pollutants in CR and AR runoff under the eight monitored rainfall events are illustrated in Figure 2. The SS, COD, TP and PP presented a stronger first flush in CR runoff than in AR runoff. However, for COD, TN,  $\text{NH}_4^+\text{-N}$  and  $\text{NO}_3^-\text{-N}$ , there were no significant differences in first flush strength between CR and AR runoff.

The  $\text{MFFR}_{20}$  was used to describe the first flush strength of pollutants in AR and CR runoff (Figure 3). The mean  $\text{MFFR}_{20}$  values of SS in CR and AR runoff were  $46.30 \pm 19.64\%$  and  $37.20 \pm 9.44\%$ , respectively. SS was the pollutant with the most significant first flush. Particulate matter in runoff has been commonly observed to have a stronger first flush than other pollutants (Lee *et al.* 2002), and other pollutants that are mainly present in particulate form, such as PP and TP, also showed a pronounced first flush.

The mean  $\text{MFFR}_{20}$  values of SS, COD and PP in CR runoff were  $46.30 \pm 19.64\%$ ,  $36.09 \pm 17.00\%$  and  $40.17 \pm 18.93\%$ , respectively. All these values were significantly higher than those in AR runoff ( $p < 0.05$ ). The SS, COD and PP in runoff were mainly particulate. The explanation for the weaker first flush of particulate pollutants in AR runoff is that particulates were removed by twice-daily street sweeping, which eliminated the main source of AR particulate pollution (street dust).

## First flush volume

The first flush cannot be quantified as a specific runoff depth by the *M(v)* curve and  $\text{MFFR}_{20}$  value. Therefore, we determined the  $V_{\text{FF}}$  of SS, COD, TN and TP in CR and AR runoff (Figures 4 and 5) using the quantitative method proposed by Bach *et al.* (2010). The left columns on the graphs depict the concentration distribution in each 1-mm slide, while the right columns show the results after aggregation. For each box, the central mark indicates the median, and the bottom and top edges of the box the 25th and 75th percentiles, respectively. The whiskers indicate the minimum and maximum values without considering outliers; the ‘+’ symbol denotes outliers.

The calculated  $V_{\text{FF}}$  values of SS, COD, TN and TP in CR runoff were 1 mm, 1, 7 and 9 mm, respectively, while the  $V_{\text{FF}}$  values of SS, COD, TN and TP in AR runoff were 1 mm, 1, 8 and 5 mm, respectively. The SS and COD  $V_{\text{FF}}$  values were the same in CR and AR runoff, indicating the first flush strength was identical. The TN and TP  $V_{\text{FF}}$  values were both higher than SS and COD in CR and AR runoff, which suggests that the first flush of SS and COD was stronger than those

of TN and TP. The  $V_{\text{FF}}$  of SS and COD was 6 mm in the study by Todeschini *et al.* (2019), and SS  $V_{\text{FF}}$  was 2–17 mm in that by Bach *et al.* (2010). The explanation for lower  $V_{\text{FF}}$  values in this study compared with previous research is the small catchment areas for CR and AR (60–64 m<sup>2</sup>). Thus, it appears that the catchment area is an important factor for first flush strength, so that smaller catchment areas experience stronger first flushes (Zhang *et al.* 2012). Furthermore, because the catchment areas were small, the first flush was pronounced in this study. But it is worth noting that when a runoff pollution control facility was used to retain the first flush, such as the stormwater tank, its catchment may not be as small as in this study. There will be several factors that should be considered in stormwater tank designing, including the variability of the rainfall over time and space and the overlaying runoff waves.

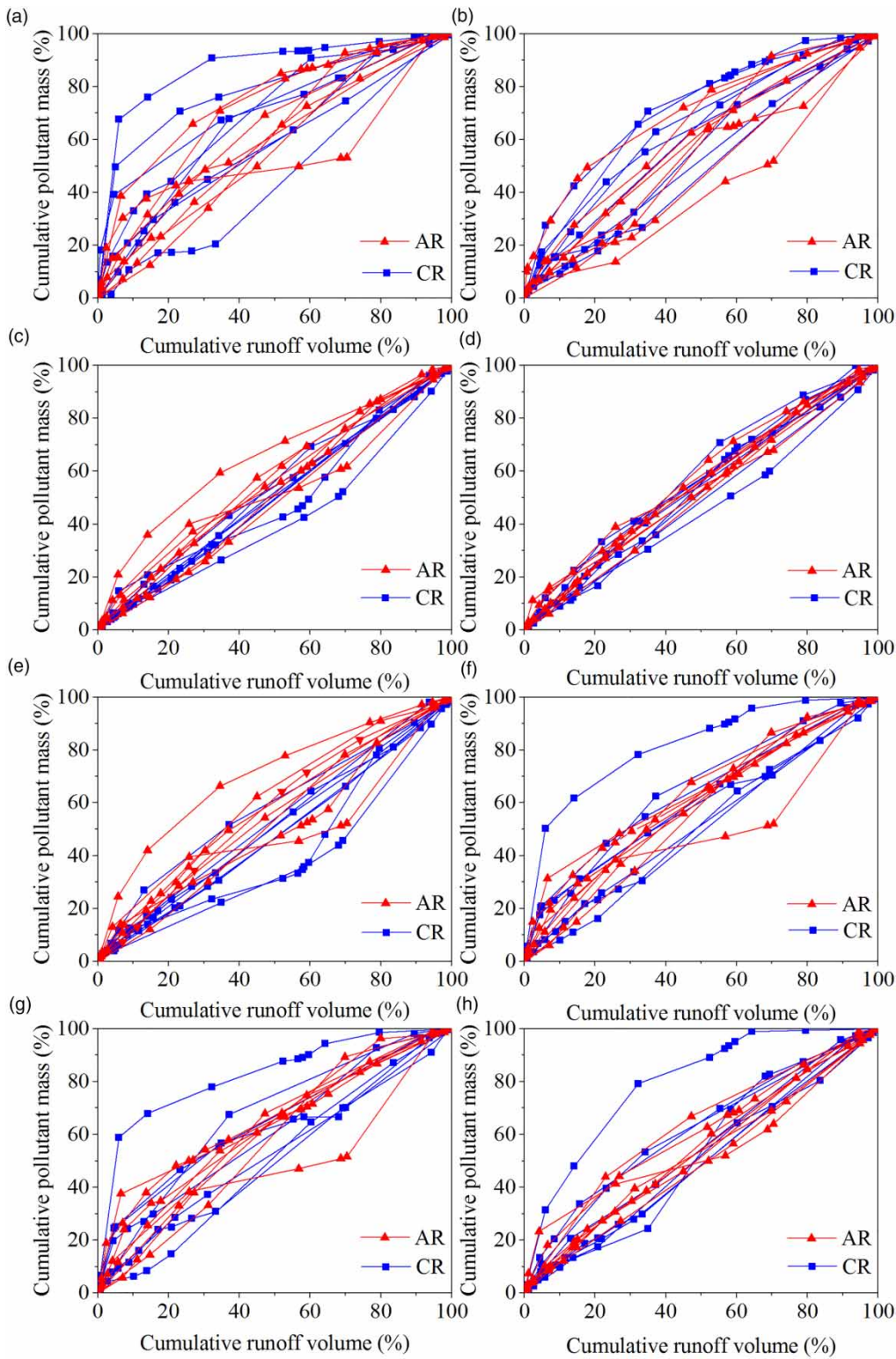
The  $V_{\text{FF}}$  of TN in AR runoff was higher than in CR runoff, whereas the TP  $V_{\text{FF}}$  in AR runoff was lower than that in CR runoff. These results indicate that the first flush strength of TN in CR and TP in AR runoff was stronger than that in any other type of runoff. There were significant differences between the first flush results of TN and TP obtained by  $\text{MFFR}_{20}$  and  $V_{\text{FF}}$ . Furthermore, there were significant differences in the first flush strength of SS and COD between CR and AR runoff when  $\text{MFFR}_{20}$  was used, but this was not the case with  $V_{\text{FF}}$ . In addition to the influence of factors such as the number of data included in analyses and analytical accuracy, reasons for the dissimilar results obtained via  $\text{MFFR}_{20}$  and  $V_{\text{FF}}$  require further study.

## Runoff pollution control efficiency

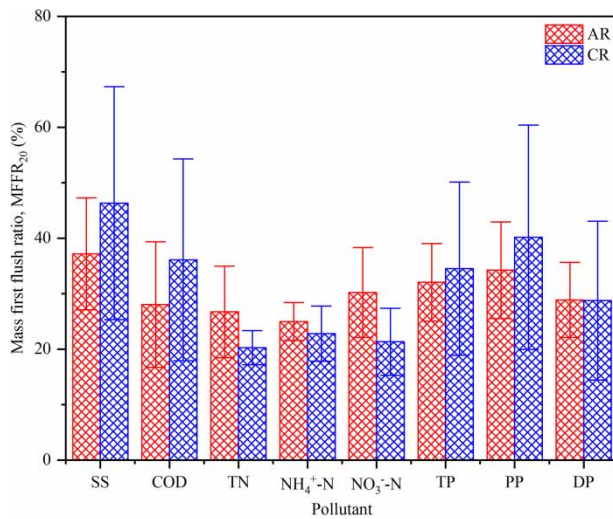
The purpose of a first flush analysis is to determine the  $V_{\text{FF}}$  and then use that information to guide urban runoff pollution control. Therefore, it is important to consider whether requirements ( $V_{\text{FF}} = 3$  mm) in the national standards of China (GB 50400-2016) and the first flush volume determined in this study ( $V_{\text{FF}} = 1$  mm) can effectively control runoff pollution.

Based on the conventional *M(v)* curve, we proposed a new method that represents the relationship between runoff depth and pollutant cumulative mass (Figure 6(a), 6(c), 6(e), 6(g)). The new curve can be used to quantitatively evaluate whether runoff pollution control corresponds to different  $V_{\text{FF}}$  values. The relationship between runoff depth and the runoff pollutant concentration was illustrated in Figure 6(b), 6(d), 6(f), 6(h).

If  $V_{\text{FF}}$  is treated according to the 3-mm requirement in Chinese national standards, the mean SS cumulative mass that would be effectively controlled in CR and AR runoff



**Figure 2** |  $M/V$  curves of pollutants in asphalt road (AR) and cement roof (CR) runoff. (a) Suspended solids (SS), (b) chemical oxygen demand (COD), (c) total nitrogen (TN), (d) ammonia ( $\text{NH}_4\text{-N}$ ), (e) nitrate ( $\text{NO}_3\text{-N}$ ), (f) total phosphorus (TP), (g) particulate phosphorus (PP) and (h) dissolved phosphorus (DP).



**Figure 3** | The mean MFFR<sub>20</sub> values of pollutants (suspended solids (SS), chemical oxygen demand (COD), total nitrogen (TN), ammonia (NH<sub>4</sub><sup>+</sup>-N), nitrate (NO<sub>3</sub><sup>-</sup>-N), total phosphorus (TP), particulate phosphorus (PP) and dissolved phosphorus (DP)) in cement roof (CR) and asphalt road (AR) runoff.

is  $91.42 \pm 9.80\%$  and  $78.49 \pm 19.41\%$ , respectively, and the mean COD cumulative mass effectively controlled is  $86.85 \pm 13.54\%$  (CR) and  $72.80 \pm 25.79\%$  (AR). This is a satisfactory result. However, if  $V_{FF}$  is treated according to the results of this study (1 mm), only 55.91%–66.65% of SS and COD cumulative mass can be effectively controlled. We assessed SS and COD cumulative mass from eight rainfall events in this study and the pollutant cumulative mass control efficiency may be decreased with heavy rain events.

In the Chinese national standards, SS and COD concentrations after 3 mm of  $V_{FF}$  in runoff should be 40 and 100 mg/L, respectively. In our study, SS and COD concentrations did not completely meet these requirements (Figure 6(b), 6(d), 6(f), 6(h)). If the  $V_{FF}$  was 3 mm, COD concentrations in CR runoff samples met the requirement, but approximately 18.18% of CR runoff samples exceed 40 mg/L of SS. Therefore, even if  $V_{FF}$  was treated according to the national recommendations, there is no guarantee that runoff after the first flush would meet water quality standards. It was worth noting that the analysis results were obtained based on the eight rainfall events monitored in this study, and the results may be related to the sampling interval, sampling frequency and the water samples number. The relationship between runoff depth and runoff pollutant cumulative mass would benefit from more monitoring data for different types of rainfall events, including more heavy rainfall events to verify our findings. Additionally, different land use with various pollution characteristics, such as industrial zone, residential zone, cultural and educational zone, and commercial zone need be comprehensively considered in future monitoring tasks.

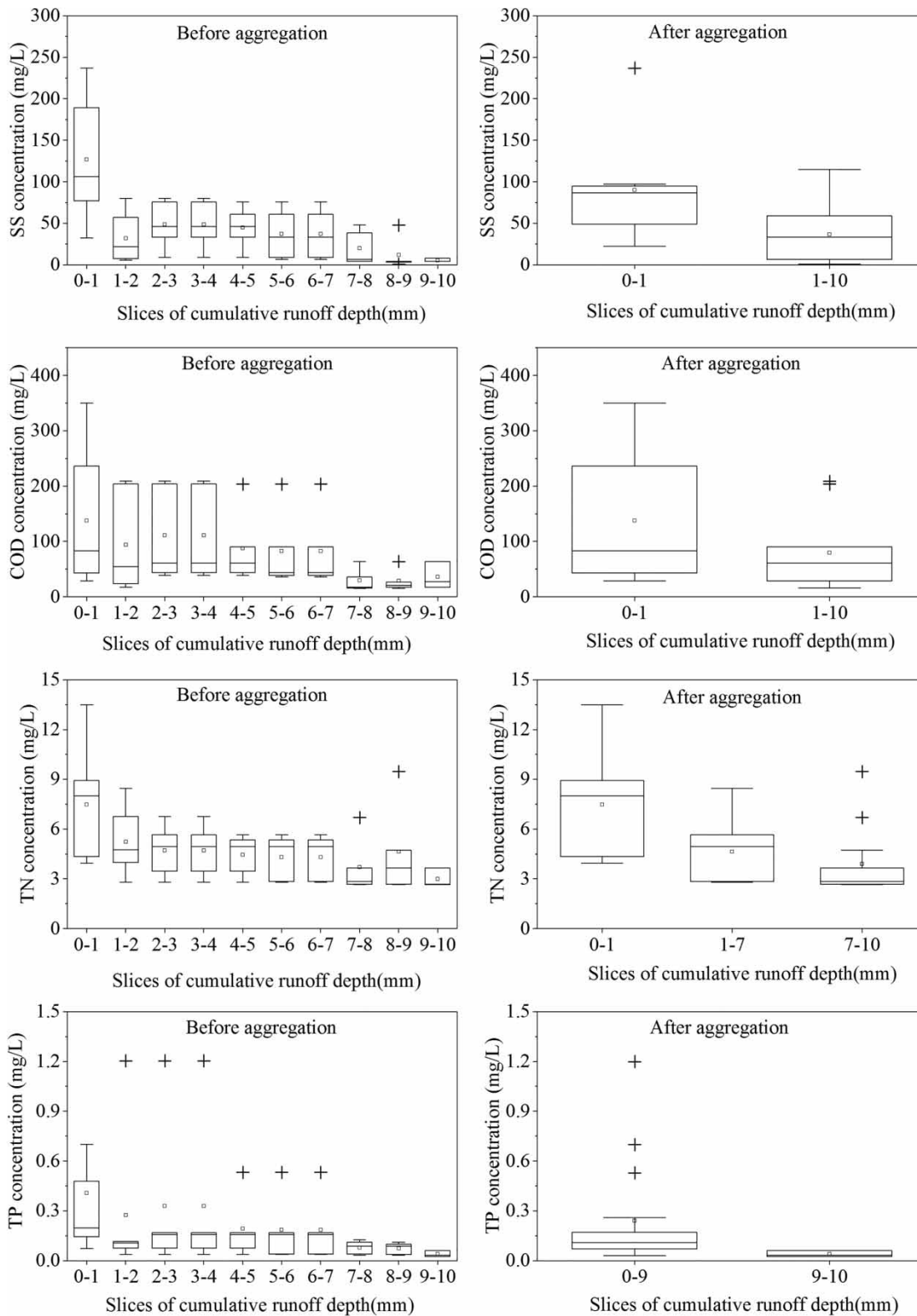
## CONCLUSIONS

In this study, the pollution characteristics and first flush strength in CR and AR runoff were analyzed and runoff pollution control efficiency was assessed. Our main findings were as follows.

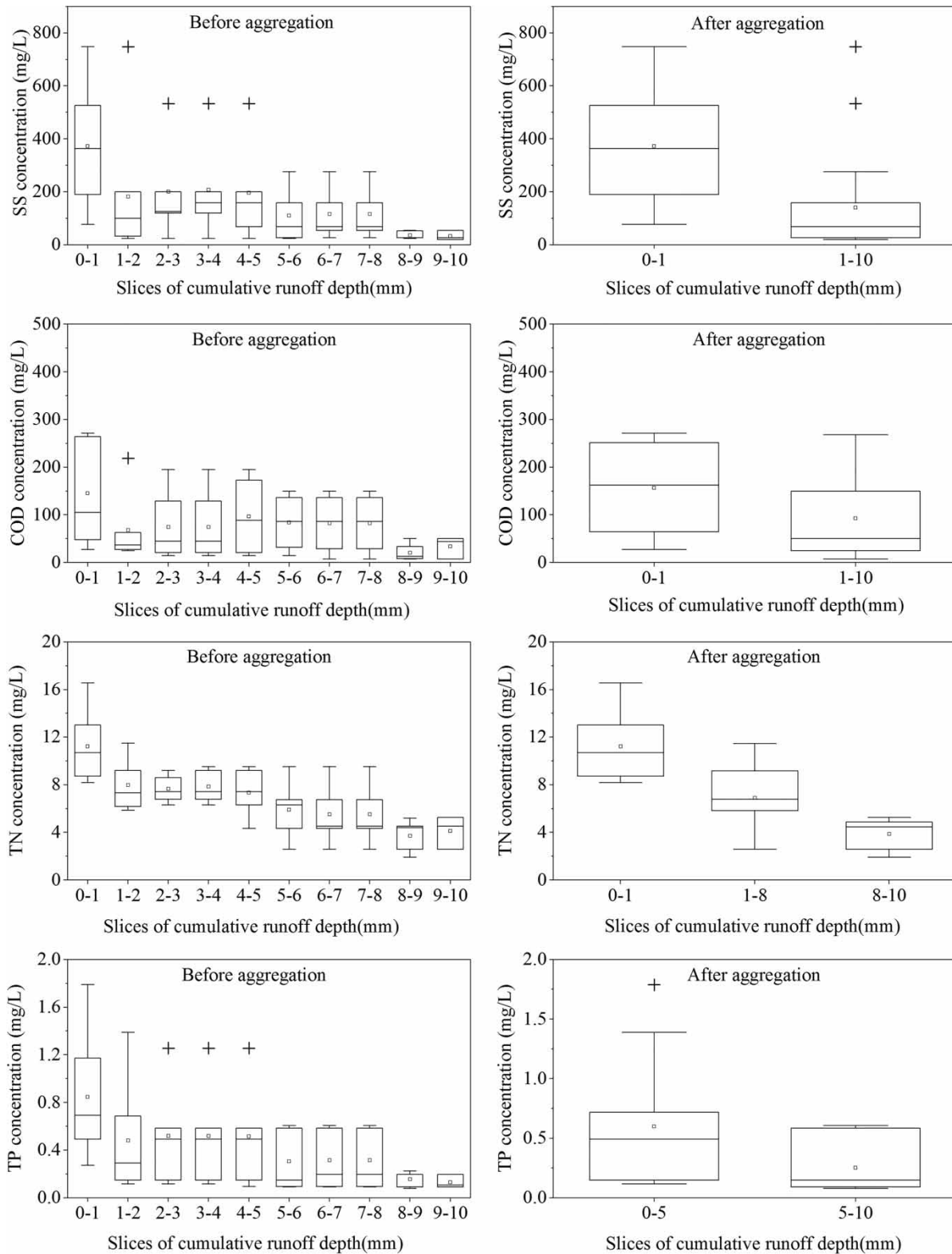
- (1) The EMCs of SS, COD, TN, and TP in AR runoff were significantly higher than those in CR runoff ( $p < 0.05$ ). Approximately  $66.55 \pm 9.44\%$  of phosphorus existed in particulate form, and  $77.94 \pm 23.95\%$  of nitrogen was present as NH<sub>4</sub><sup>+</sup>-N and NO<sub>3</sub><sup>-</sup>-N. The proportion of particles larger than 75 μm in CR and AR runoff was  $50.42 \pm 15.57\%$  and  $49.72 \pm 32.35\%$ , respectively, and there was no significant difference between CR and AR runoff ( $p > 0.05$ ).
- (2) Based on the  $M(v)$  curve, the SS, COD, TP and PP, which were mainly present as particulates, exhibited a stronger first flush in CR runoff than in AR runoff, but the first flush strength of TN, NH<sub>4</sub><sup>+</sup>-N, NO<sub>3</sub><sup>-</sup>-N and DP did not show significant differences between CR and AR runoff ( $p > 0.05$ ). The  $V_{FF}$  values of SS, COD, TN and TP were similar between CR runoff (1 mm, 1, 7 and 9 mm, respectively) and AR runoff (1 mm, 1, 8 and 5 mm, respectively). However, there were differences in the first flush results between the traditional  $M(v)$  and  $V_{FF}$ .
- (3) A runoff depth against pollutant cumulative mass curve was proposed to assess the runoff pollution control performance of different  $V_{FF}$  values. If the  $V_{FF}$  was treated according to the national standards (3 mm), the mean SS cumulative mass that is effectively controlled in CR and AR runoff would be  $91.42 \pm 9.80\%$  and  $78.49 \pm 19.41\%$ , respectively, and the effectively controlled mean COD cumulative mass would be  $86.85 \pm 13.54\%$  and  $72.80 \pm 25.79\%$ . If the  $V_{FF}$  was treated according to the results of this study (1 mm), only 55.91%–66.65% of SS and COD cumulative mass would be effectively controlled.

Statistically significant results were obtained from monitoring eight rainfall events, and the results are affected by the sampling interval, sampling frequency, and number of water samples. Although our new curve presents the relationship between runoff depth and runoff pollutant cumulative mass, more monitoring data included different land use and various rainfall types needed to verify our results. If confirmed, the new method will provide an alternative approach to assess runoff pollution control efficiency.

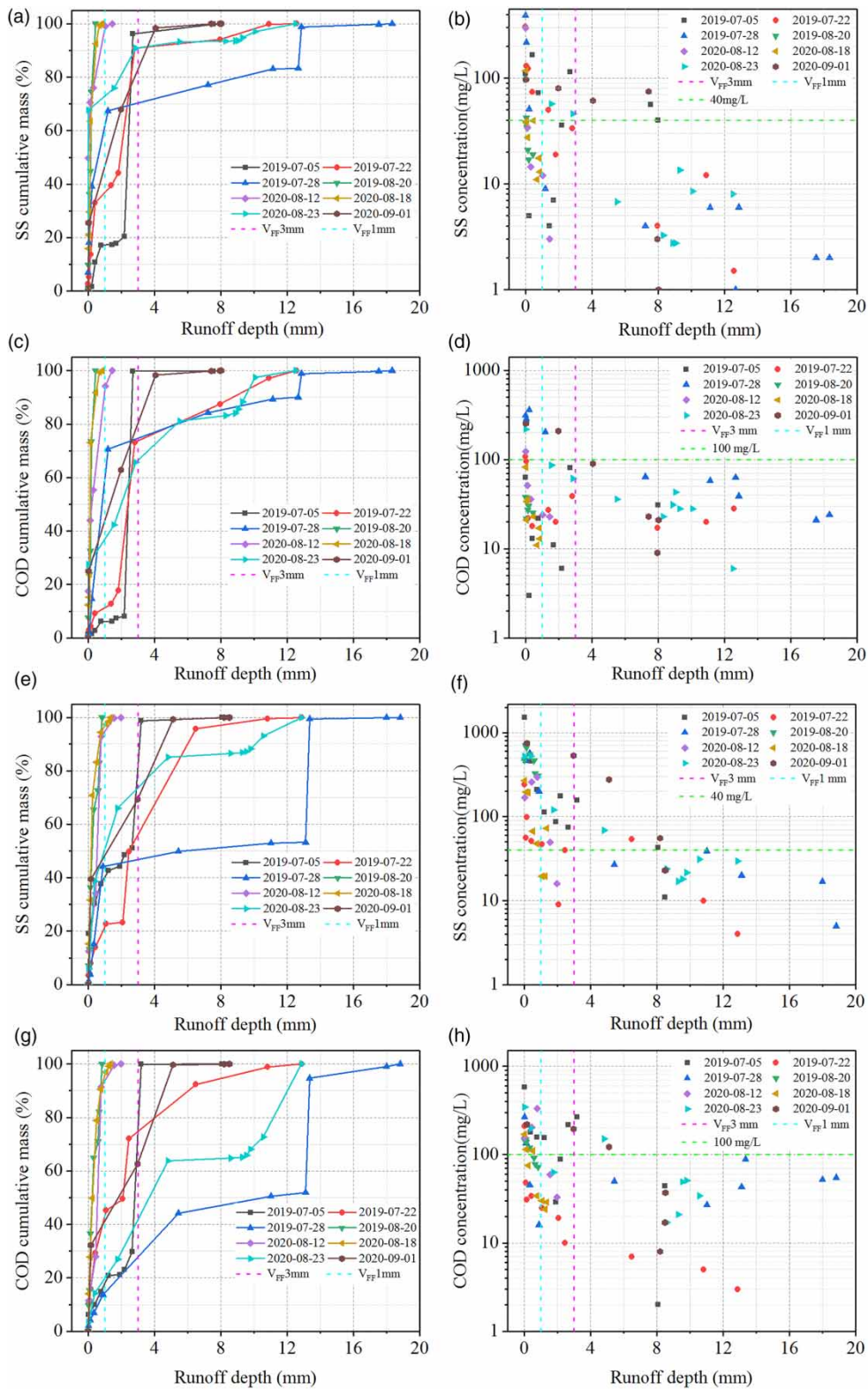




**Figure 4** | Ungrouped (slice size = 1 mm) and grouped (5% significance level) box and whisker plots for suspended solids (SS), chemical oxygen demand (COD), total nitrogen (TN) and total phosphorus (TP) concentrations in asphalt road (CR) runoff.



**Figure 5** | Ungrouped (slice size = 1 mm) and grouped (5% significance level) box and whisker plots for suspended solids (SS), chemical oxygen demand (COD), total nitrogen (TN) and total phosphorus (TP) concentrations in cement roof (AR) runoff.



**Figure 6** | Relationships of (a–d) cement roof (CR) and (e–h) asphalt road (AR) runoff with runoff depth and (b, d, f, h) suspended solids (SS) and chemical oxygen demand (COD) concentrations, and runoff depth and (a, c, e, g) SS and COD cumulative mass. The vertical dashed lines are the  $V_{FF}$  values required in Chinese national standards (3 mm) (pink) and determined in this study (1 mm) (light blue). The horizontal green dashed line is the SS (40 mg/L) and COD (100 mg/L) concentrations required in the national standards.

## ACKNOWLEDGEMENTS

This study was financially supported by the National Natural Science Foundation of China (No. 51608026) and the Scientific Research Project of Beijing Educational Committee (No. KM201910016012). The authors thank the research team members for their enthusiastic support. We thank Kara Bogus, PhD, from Liwen Bianji, Edanz Editing China ([www.liwenbianji.cn/ac](http://www.liwenbianji.cn/ac)), for editing the English text of a draft of this manuscript.

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## REFERENCES

- APHA, AWWA, WEF 2012 *Standard Methods for the Examination of Water and Wastewater*, 22nd edn. American Public Health Association, Washington, DC, USA.
- Bach, P. M., McCarthy, D. T. & Deletic, A. 2010 [Redefining the stormwater first flush phenomenon](#). *Water Research* **44**, 2487–2498.
- Barco, J., Papiri, S. & Stenstrom, M. K. 2008 [First flush in a combined sewer system](#). *Chemosphere* **71**, 827–833.
- Bertrand-Krajewski, J.-L., Chebbo, G. & Saget, A. 1998 [Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon](#). *Water Research* **32** (8), 2341–2356.
- Burton, G. A. & Pitt, R. E. 2002 *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers*. CRC Press, Lewis Publishers, Washington, DC, USA.
- Chen, L., Zhi, X., Shen, Z., Dai, Y. & Aini, G. 2018 [Comparison between snowmelt-runoff and rainfall-runoff nonpoint source pollution in a typical urban catchment in Beijing, China](#). *Environmental Science and Pollution Research* **25** (3), 2377–2388.
- Deletic, A. 1998 [The first flush load of urban surface runoff](#). *Water Research* **32** (8), 2462–2470.
- Egodawatta, P., Thomas, E. & Goonetilleke, A. 2009 [Understanding the physical processes of pollutant build-up and wash-off on roof surfaces](#). *Science of the Total Environment* **407** (6), 1834–1841.
- Geiger, W. F. 1984 [Characteristics of combined sewer runoff](#). In: *Proceedings of the 3rd International Conference on Urban Storm Drainage*, Goteborg, Sweden, pp. 851–860.
- Han, Y. H., Lau, S. L., Kayhanian, M. & Stenstrom, M. K. 2006 [Correlation analysis among highway stormwater pollutants and characteristics](#). *Water Science and Technology* **53** (2), 235–243.
- Hou, L. & Zhang, X. 2014 [Comparing urban runoff quality of a felt roof and an asphalt road in Beijing](#). *Hydrology Research* **45** (2), 282–291.
- Jeung, M., Baek, S., Beom, J., Cho, K. H., Her, Y. & Yoon, K. 2019 [Evaluation of random forest and regression tree methods for estimation of mass first flush ratio in urban catchments](#). *Journal of Hydrology* **575**, 1099–1110.
- Lee, J. H. & Bang, K. W. 2000 [Characterization of urban stormwater runoff](#). *Water Research* **34** (6), 1773–1780.
- Lee, J. H., Bang, K. W., Ketchum, L. H., Choe, J. S. & Yu, M. J. 2002 [First flush analysis of urban storm runoff](#). *Science of the Total Environment* **293** (1), 163–175.
- Lusk, M. G., Toor, G. S. & Inglett, P. W. 2020 [Organic nitrogen in residential stormwater runoff: implications for stormwater management in urban watersheds](#). *Science of the Total Environment* **707**, 135962.
- Perera, T., McGree, J., Egodawatta, P., Jinadasa, K. B. S. N. & Goonetilleke, A. 2019 [Taxonomy of influential factors for predicting pollutant first flush in urban stormwater runoff](#). *Water Research* **166**, 115075.
- Ren, Y.-F., Wang, X.-K., Ouyang, Z.-Y. & Hou, P.-Q. 2013 [Analysis of first flush effect of typical underlying surface runoff in Beijing urban city](#). *Environmental Sciences* **34** (1), 373.
- Sansalone, J. J. & Cristina, C. M. 2004 [First flush concepts for suspended and dissolved solids in small impervious watersheds](#). *Journal of Environmental Engineering* **130** (11), 1301–1314.
- Sartor, J. D. & Boyd, G. B. 1972 *Water Pollution Aspects of Street Surface Contaminants*, EPA-R2/72-081. U.S. Environmental Protection Agency, Washington, DC, USA.
- Stahre, P. & Urbonas, B. 1990 *Stormwater Detention*. Prentice Hall, Inc., Englewood Cliffs, NJ, USA, p. 338.
- Taylor, G. D., Fletcher, T., Wong, T. H. F., Breen, P. F. & Duncan, H. P. 2005 [Nitrogen composition in urban runoff—implications for stormwater management](#). *Water Research* **39** (10), 1982–1989.
- Todeschini, S., Manenti, S. & Creaco, E. 2019 [Testing an innovative first flush identification methodology against field data from an Italian catchment](#). *Journal of Environmental Management* **246**, 418–425.
- UK Highways Agency 2006 *Vegetated Drainage Systems for Highway Runoff*. HA 103/06, Vol. 4. Design Manual for Roads and Bridges, The Stationery Office, London, UK.
- Wanielista, M. & Yousef, Y. 1993 *Stormwater Management*. John Wiley and Sons Inc, New York, NY, USA, p. 579.
- Winston, R. J. & Hunt, W. F. 2017 [Characterizing runoff from roads: particle size distributions, nutrients, and gross solids](#). *Journal of Environmental Engineering* **143** (1), 4016074.
- Zhang, W., Che, W., Liu, D. K., Gan, Y. P. & Lv, F. F. 2012 [Characterization of runoff from various urban catchments at different spatial scales in Beijing, China](#). *Water Science and Technology* **66** (1), 21–27.
- Zhao, H., Li, X., Wang, X. & Tian, D. 2010 [Grain size distribution of road-deposited sediment and its contribution to heavy metal pollution in urban runoff in Beijing, China](#). *Journal of Hazardous Materials* **183** (3), 203–210.