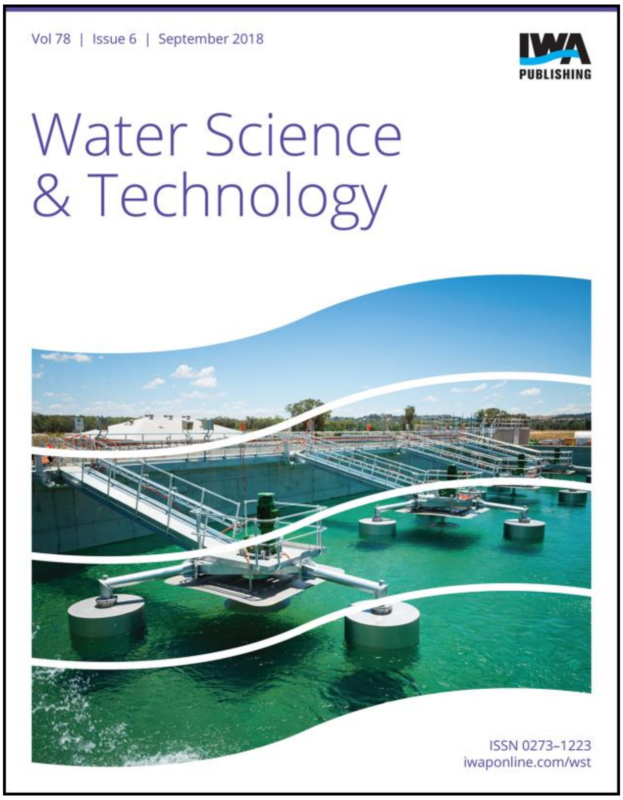
ELECTRONIC OFFPRINT
Use of this pdf is subject to the terms described below



This paper was originally published by IWA Publishing. The author's right to reuse and post their work published by IWA Publishing is defined by IWA Publishing's copyright policy.

If the copyright has been transferred to IWA Publishing, the publisher recognizes the retention of the right by the author(s) to photocopy or make single electronic copies of the paper for their own personal use, including for their own classroom use, or the personal use of colleagues, provided the copies are not offered for sale and are not distributed in a systematic way outside of their employing institution. **Please note that you are not permitted to post the IWA Publishing PDF version of your paper on your own website or your institution's website or repository.**

If the paper has been published "Open Access", the terms of its use and distribution are defined by the Creative Commons licence selected by the author.

Full details can be found here: http://iwaponline.com/content/rights-permissions

Please direct any queries regarding use or permissions to wst@iwap.co.uk

Performance and uncertainties of TSS stormwater sampling strategies from online time series

Santiago Sandoval, Jean-Luc Bertrand-Krajewski, Nicolas Caradot, Thomas Hofer and Günter Gruber

ABSTRACT

The event mean concentrations (EMCs) that would have been obtained by four different stormwater sampling strategies are simulated by using total suspended solids (TSS) and flowrate time series (about one minute time-step and one year of data). These EMCs are compared to the reference EMCs calculated by considering the complete time series. The sampling strategies are assessed with datasets from four catchments: (i) Berlin, Germany, combined sewer overflow (CSO); (ii) Graz, Austria, CSO; (iii) Chassieu, France, separate sewer system; and (iv) Ecully, France, CSO. A sampling strategy in which samples are collected at constant time intervals over the rainfall event and sampling volumes are pre-set as proportional to the runoff volume discharged between two consecutive sample leads to the most representative results. Recommended sampling time intervals are of 5 min for Berlin and Chassieu (resp. 100 and 185 ha area) and 10 min for Graz and Ecully (resp. 335 and 245 ha area), with relative sampling errors between 7% and 20% and uncertainties in sampling errors of about 5%. Uncertainties related to sampling volumes, TSS laboratory analyses and beginning/ending of rainstorm events are reported as the most influent sources in the uncertainties of sampling errors and EMCs.

Key words | automated sampler, event mean concentration, measurements representativeness, online monitoring, stormwater sampling

Santiago Sandoval (corresponding author)
Jean-Luc Bertrand-Krajewski
Université de Lyon, INSA Lyon, DEEP,
EA 7429, 34 avenue des Arts, F-69621
Villeurbanne cedex,
France
E-mail: *santiago.sandoval-arenas@insa-lyon.fr*

Nicolas Caradot
Kompetenzzentrum Wasser Berlin GmbH,
Cicerostraße 24, D-10709 Berlin,
Germany

Thomas Hofer
Günter Gruber
Graz University of Technology, Institute of Urban Water Management and Landscape Water Engineering,
Stremayrgasse 10/I, A-8010 Graz,
Austria

INTRODUCTION

Accurately assessing stormwater pollutant loads and their variability through monitoring programs is essential for scientific, operational, legal, political and environmental requirements (Larrarte 2008; Ackerman *et al.* 2010). Monitoring is frequently based on sampling, which itself depends on monitoring objectives, legal constraints, costs and logistics considerations (see Duncan 1999; Brombach *et al.* 2005; Ellis *et al.* 2006). The main goal is to collect the most accurate, reliable and representative data, given technical and resources limitations (Ackerman *et al.* 2010). Considering that various sources of error affect the monitoring results, especially the sampling technique and the heterogeneity in both space and time of the sampling target (Paakkunainen *et al.* 2007), generalizable and efficient strategies for stormwater quality sampling, adaptable to technical and operational restrictions, remain a relevant question.

Total suspended solids (TSS) concentration is one of the usual key variables to estimate stormwater pollutant loads, with high dynamics within storm events and a great variability between events (e.g. Métadier & Bertrand-Krajewski 2012). This paper therefore deals with the evaluation of different stormwater TSS sampling strategies. Sampling strategies are defined hereafter as operational rules for sampling TSS during rainfall events with an automated sampler or grab sampling, aimed at maximizing the representativeness of the collected samples.

One common indicator used to assess pollutant load is the event mean concentration (EMC) (USEPA 1983; Carleton *et al.* 2001; Kim *et al.* 2005; Lee *et al.* 2007), which can be estimated from different sampling schemes (e.g. grab, flow weighted, time weighted composite samples) (Lee *et al.* 2007). However, EMCs have shown to be very

doi: 10.2166/wst.2018.415

variable, depending on the sampling strategy applied for their calculation (Lee *et al.* 2007; Ki *et al.* 2011).

Previous studies have evaluated the performance of various sampling strategies for quantifying pollutant concentrations and loads during rainfall events (e.g. Izuno *et al.* 1998; Stone *et al.* 2000; Ma *et al.* 2009; Ackerman *et al.* 2010). They have focused on comparing the EMCs obtained by measurements with a reference EMC calculated from numerical estimations (Shih *et al.* 1994), Monte Carlo simulations (Richards & Holloway 1987) and statistical methods (King *et al.* 2005; Ma *et al.* 2009), including a limited number of data in the analyses (Madarang & Kang 2013; McCarthy *et al.* 2018). Indeed, the main challenge in these studies is that the differences between the EMCs obtained from sampling strategies and the reference EMCs are established by hardly verifiable assumptions, precisely due to the absence of sufficient experimental data (e.g. 2 min temporal resolution water quality measurements). On the other hand, further studies estimate the reference EMC using the outputs of stormwater quantity and quality high temporal resolution conceptual models, calibrated with flow rate and water quality measurements (Ackerman *et al.* 2010; Ki *et al.* 2011). However, conceptual stormwater quality models have highly uncertain outputs due to uncertainties in model structure, model parameters and data (input and output) (e.g. Benedetti *et al.* 2013; Bonhomme & Petrucci 2017), affecting therefore the evaluation of any given sampling strategy as well.

In the last years, measurement of stormwater quality with online sensors (e.g. turbidity, UV-VIS spectrometry) has been increasingly used for several purposes (e.g. control of urban water systems, modelling, real time control, etc.). Indeed, multiple authors have reported the ability of online monitoring to provide high resolution information on space and time variability of stormwater quality and on the complexity of the underlying processes (e.g. Gruber *et al.* 2005; Métadier & Bertrand-Krajewski 2012). Accordingly, online monitoring appears as a promising alternative for estimating reference EMCs and to make comparisons across a range of conditions (adapted from Ackerman *et al.* 2010).

This paper presents the simulation and evaluation of four different stormwater TSS sampling strategies proposed in the literature or applied by practitioners (e.g. Rossi 1998; Ackerman *et al.* 2010) by using high temporal resolution flowrate and TSS time series (approximately one minute time-step, with about one year of data). EMCs calculated from sampling strategies are simulated by sampling TSS time series and calculating a weighted average of the samples by their sampling volumes. These simulated EMCs are compared to the reference EMC calculated as a weighted average of the complete time series (flow rate and TSS). The work is done with datasets from four urban drainage systems, aimed at comparing the results among different catchments. The four cases are the following ones: (i) Berlin, Germany, combined sewer overflow (CSO); (ii) Graz, Austria, (CSO); (iii) Chassieu, France, (separate sewer system); and (iv) Ecully, France (combined sewer system).

To our best knowledge, uncertainties and sensitivity analysis in the evaluation of sampling strategies have not been extensively addressed in the literature (e.g. Rossi 1998; King *et al.* 2005). Therefore, uncertainties in (i) sampling volumes, (ii) laboratory analyses, (iii) online measurements (TSS and flow rate) and (iv) settings of beginning and ending of storm events are assessed and propagated in the results by Monte Carlo simulations. The effects of these independent uncertainty sources on the total uncertainties are estimated by Sobol's sensitivity index (Glen & Isaacs 2012).

MATERIALS AND METHODS

Data sets

Datasets from four urban drainage systems are used to evaluate the sampling strategies. Table 1 gives a summary of the main characteristics of the catchments and of their time series.

The dry and wet weather periods for the combined systems in Graz and Ecully are defined based on flow rate values. Periods in which the inflows of the CSO-chamber are greater than 120 L/s (in Graz) and flow rates are greater than 30 L/s in Ecully are defined as the wet periods. For the case of Berlin, the monitoring station is installed in a major overflow sewer downstream of the main overflow structures. Therefore, the measured data can be directly considered as the wet period.

Sampling strategies

The four investigated sampling strategies are defined to be applicable with an automated sampler, jointly with a flowmeter in some cases, by using different criteria. For all sampling strategies, accounting for most commercially available automated samplers, it is assumed that each elementary sample cannot be larger than 1 L, and that all elementary

Table 1 | Main characteristics of the four urban catchments and their time series

Location	Year	Monitoring point	Area (km^2)	% imperv. (corresponding area)	Inhab. ~	# rainfall events	Time series time-step	Land use	Online TSS measurement	Source
Berlin, Germany	2010	CSO[a]	1	74 (0.74 km^2)	126,000	22	1 min	Residential	UV-VIS	Sandoval et al. (2013)
Graz, Austria	2009	CSO[a]	3.35	32 (1.08 km^2)	11,800	79	Dry weather: 3 min; wet weather: 1 min	Residential	UV-VIS	Gamerith (2011)
Chassieu, Lyon, France	2007	Separate sewer system outlet	1.85	75 (1.39 km^2)	NA	89	2 min	Industrial	Turbidity	Métadier (2011)
Ecully, Lyon, France	2007	CSO[a]	2.45	42 (1.03 km^2)	18,000	220	2 min	Residential	Turbidity	Métadier (2011)

[a]CSO: combined sewer overflow.

samples are mixed in a composite sample jar of maximum 20 L. It is assumed that the minimum volume for an elementary sample is 0.02 L; that is, the minimum sampling volume for which a TSS laboratory analysis can be conducted (guaranteeing TSS concentrations above the detection limits). As most sampling bottles have a capacity of 1 L, the maximum volume for an elementary sample is set to 0.9 L, in order to keep a security margin, in case of, for example, spilling some amount of the sample during the handling process.

The following paragraphs describe the four strategies to be evaluated, which are summarized in Table 2 (additional details can be found, for example in Rossi 1998).

cTcSV strategy: Time-paced composite sampling with constant sampling volume. One sample is collected based on equally spaced time intervals (cT) and sampling volumes are constant (cSV). The inputs for this strategy are the sampling interval during the rainfall event and the constant sampling volume. The sampling volume is predefined to any constant value between 0.02 and 0.9 L (set to 0.4 L for comparative purposes, as discussed below). Example: sampling each 10 min with a sampling volume of 0.4 L.

cTpQ strategy: Time-paced composite sampling with sampling volume proportional to instantaneous flowrate. One sample is collected based on equally spaced time intervals (cT) and sampling volumes are pre-set as proportional to the instantaneous flowrate at the sampling point measured at the sampling time (pQ). The sampling volumes are proportionally defined: the minimum flowrate value registered during a rainfall event then corresponds to the minimum sampling volume (0.02 L), and the maximum flowrate is sampled with a volume of 0.9 L. Therefore, in addition to the constant sampling interval, the minimum and maximum flowrate values during the event are required as inputs. Example: sampling each 10 min with a sampling volume of 0.02 L (resp. 0.1 L) if the instantaneous flowrate is 0.1 m^3/s (resp. 0.5 m^3/s).

cTpRV strategy: Time-paced composite sampling with sampling volume proportional to runoff volume discharged between two successive samples. One sample is collected based on equally spaced time intervals (cT) and sampling volumes are pre-set as proportional to the runoff volume at the sampling point cumulated since the previous sample

Table 2 | Sampling strategies description

Name	Sampling time interval	Sampling volume	Runoff volume between 2 samples	Input parameters
cTcSV	Constant	Constant	Variable	Constant sampling time interval
cTpQ	Constant	f(flow rate)	Variable	Constant sampling time interval; min(flow rate) and max(flow rate)
cTpRV	Constant	f(RV[a]) since last sample	Variable	Constant sampling interval; min(RV[a]) and max(RV[a]) between 2 samples
vTcRV	Variable	Constant	Constant	Sampling intervals based on RV[a]

[a]RV: runoff volume between successive samples.

(pRV). The sampling volumes are proportionally defined: the minimum runoff volume discharged between two samples during a given rainfall event then corresponds to the minimum sampling volume (0.02 L), and the maximum runoff volume between two samples is sampled with a volume of 0.9 L. Therefore, in addition to the constant sampling interval, the minimum and maximum volumes between two samples during the event are required as inputs. Example: sampling each 10 min with a sampling volume of 0.1 L (resp. 0.5 L) if the runoff volume discharged since the previous sample is 2 m^3 (resp. 10 m^3).

vTcRV strategy: Volume-paced composite sampling. For this strategy, the automated sampler takes a sample for a pre-set runoff volume (cRV) during a given rainfall event. This leads to non-equally spaced time intervals (vT) between samples. The sampling volume is predefined to any constant value between 0.02 L and 0.9 L (set to 0.4 L for comparative purposes, as discussed below) as for the cTcSV strategy. Thus, the inputs for this strategy are the constant sampling volume and the pre-set runoff volume. Example: a constant volume of 0.4 L is sampled for each 10 m^3 of runoff volume discharged between two successive samples.

cTcSV strategy does not require any flowmeter. cTpQ strategy requires a flowmeter. The cTpRV and vTcRV strategies require a flowmeter and the integration of discharge to calculate runoff volumes between two samples.

Methodology

Given that comparing the results of simultaneous campaigns with several automated samplers (one for each strategy with different inputs) is unfeasible in both economic and operational terms, measured TSS and flow rate continuous time series are used to simulate the EMCs_{sim} that would have been obtained by different sampling strategies. The TSS time series is sampled virtually assigning a TSS_i value (i equal to the sampling time-step) and a sampling volume SV_i to each sampling bottle, as a function of the sampling strategy.

The EMC_{sim} is calculated as a weighted average of the TSS_i values by the sampling volumes SV_i. This calculation is equivalent to grabbing multiple elementary samples and mixing them in a composite 20 L sample jar. The EMCs_{sim} are simulated for each sampling strategy with a range of inputs (Table 2). The reference EMC_{true} is calculated by using the complete continuous time series of Q and TSS. $\text{EMC}_{\text{true}}(j)$ and $\text{EMC}_{\text{sim}}(j)$ are calculated for each rainfall event j (from 22 to 220 rainfall events depending on the catchment and dataset, see Table 1).

A range of sampling time intervals (from 1 to 60 min) is evaluated for the three strategies with constant sampling intervals (cTcSV, cTpQ and cTpRV strategies). Based on recommendations from Ackerman *et al.* (2010) and field experience, the recommended minimum sampling time interval is 5 min. Nevertheless, lower values have been tested for comparison purposes. For applying proportional sampling volume strategies (cTpQ and cTpRV), the minimum and maximum Q values of each event should be known in advance. In practice, these values can be estimated only after the end of the rainfall event. The simple solution adopted for this work assumes that all bottles are filled with the highest possible sampling volume (0.9 L). Afterwards, the corresponding sampling volume SV to be taken from each bottle is corrected by using the then known minimum and maximum flow rate values, before mixing the elementary samples into the 20 L composite jar.

The performance of vTcRV strategy (variable sampling time intervals) is evaluated for several runoff volumes RV, from the 5th percentile to the 95th percentile of the total runoff volumes during rainfall events. The mean sampling time interval is then calculated for comparison with constant sampling time interval strategies (cTcSV, cTpQ and cTpRV).

The following performance indicators are used to evaluate each sampling strategy with different inputs: (i) the *residuals_vector* for each rainfall event (Equation (1)) (bias estimation), (ii) the Mean Squared Relative Error (*MSRE*) for all rainfall events (accuracy estimation) (Equation (2)) and (iii) the uncertainty or confidence interval CI of *MSRE*, designated by $CI(MSRE)_{95\%}$.

$$Residuals\_vector(j) = EMC_{true}(j) - EMC_{sim}(j) \tag{1}$$

$$MSRE = \frac{\sqrt{\sum_{j=1}^{N} residuals(j)^2}}{N \cdot mean(EMC_{true})} \tag{2}$$

The *residuals_vector* has a length equal to the number N of rainfall events in each data set and can be considered as a bias estimation (over or under estimation of the reference EMC_{true}), as the mean of this vector should be significantly close to zero for guaranteeing a non-biased estimation over all j rainfall events (adapted from Gy 1998).

The expanded uncertainty $CI(MSRE)_{95\%}$ is calculated by Monte Carlo simulations as the minimum interval covering 95% of the distribution of *MSRE* values (ISO 2009). Six uncertainty sources are considered, which are listed in Table 3 (adapted from Rossi 1998).