Department of Civil and
Structural Engineering.

Accounting for rainfall variability in sediment wash-off modelling using uncertainty propagation

By
Manoranjan Muthusamy

எப்பொருள் யார்யார்வாய்க் கேட்பினும் அப்பொருள்

மெய்ப்பொருள் காண்ப தறிவு

In whatever matter and from whomever heard,
wisdom will witness its true meaning.

Thirukkural, Verse 423 (Thiruvalluvar, 132 BC) Translated by GU Pope, 1886

Abstract

Urban surface sediment is a major source of pollution as it acts as a transport medium for many contaminants. Accurate modelling of sediment wash-off from urban surfaces requires an understanding of the effect of variability in the external drivers such as rainfall on the wash-off process. This study investigates the uncertainty created due to the urban-scale variability of rainfall, in sediment wash-off predictions. Firstly, a rigorous geostatistical method was developed that quantifies uncertainty due to spatial rainfall variability of rainfall at an urban scale. The new method was applied to a unique high resolution rainfall dataset collected with multiple paired gauges for a study designed to quantify rainfall uncertainty. Secondly, the correlation between calibration parameters and external drivers - rainfall intensity, surface slope and initial load- was established for a widely used exponential wash-off model using data obtained from new detailed laboratory experiments. Based on this, a new wash-off model where the calibration parameters are replaced with functions of these external drivers was derived. Finally, this new wash-off model was used to investigate the propagation of rainfall uncertainty in wash-off predictions. This work produced for the first time quantitative predictions of the variation in wash-off load that can be linked to the rainfall variability observed at an urban scale.

The results show that (1) the assumption of constant spatial rainfall variability across rainfall intensity ranges is invalid for small spatial and temporal scales, (2) wash-off load is sensitive to initial loads and using a constant initial load in wash-off modelling is not valid, (3) the level of uncertainty in predicted wash-off load due to rainfall uncertainty depends on the rainfall intensity range and the "first-flush" effect. The maximum uncertainty in the prediction of peak wash-off load due to rainfall uncertainty within an 8-ha catchment was found to be ~15%.

Keywords: sediment wash-off, sediment build up, model uncertainty, exponential wash-off model, rainfall variability

Acknowledgement

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D	2 Dimensional
AARI	Areal Average Rainfall Intensity
BUWO	Build Up and Wash Off
CV	Coefficient of Variation
CV_R	Coefficient of Variation associated with rainfall intensity prediction
EMC	Event Mean Concentration
IID	Independently Identically Distributed
LCM	Lump catchment models
MCMC	Markov Chain Monte Carlo
NEM	New Exponential Model
NST	Normal Score Transformation
OEM	Original Exponential Model
PICP	Prediction Interval Coverage Probability
QUICS	Quantifying Uncertainty in Integrated Catchment Studies
RMSE	Root Mean Square Error
$RMSE_{NEM}$	Root mean square error of Original Exponential Model
$RMSE_{OEM}$	Root mean square error of Original Exponential Model
TB	Tipping Bucket
VRF	Variance Reduction Factor
WaPUG	Wastewater Planning Users Group
CVW	Coefficient of Variation associated with instantaneous wash-off prediction

1. Introduction

1.1 The bigger picture

Urban surface sediment is a major source of pollutants in the urban environment which contributes to the degradation of urban water quality, mainly due to its ability to act as a transport medium for many contaminants (Guy, 1970; Collins and Ridgeway, 1980; Mitchell *et al.*, 2001; Lawler *et al.*, 2006). For example, Collins and Ridgeway (1980) found that in urban storm events, sediment smaller than < 63μm transported more than 50% of the total pollutant load. In addition, transported sediment can also contribute to urban flood risk by depositing in urban drainage systems and consequently reducing their hydraulic capacity (Delleur, 2001; Ivan, 2001). Sediment deposition also causes problems such as early and more frequent overflows, larger pollutant discharges and costly removal (Ashley *et al.*, 1992; Delleur, 2001; Heal *et al.*, 2006). The erosion of sediments in sewers can also release pollutants in high concentrations from combined sewer overflows that exceed the levels found in the various contributing sources of the sediments and pollutants (Ashley *et al.*, 1992). Hence, the importance of accurate modelling of sediment transport from urban surfaces (also known as sediment wash-off) is important in water quality and flood risk based decision making. But, modelling sediment wash-off is not a straightforward exercise as it requires the understanding of complex interactions between external drivers with a highly variable nature associated with rainfall, catchment surface and particle characteristics (Sartor and Boyd, 1972; Deletic *et al.*, 1997; Egodawatta and Goonetilleke, 2008). Due to difficulty in modelling this complex interactions using first principals, most widely used wash-off models that are empirically derived using limited laboratory and field experimental data (e.g. Sartor and Boyd, 1972; Egodawatta *et al.*, 2007; Francey *et al.*, 2011). The inherent shortcoming of these empirical methods is that the calibration parameters in these models do not have a strong physical meaning and therefore applicability and transferability of these parameters to other catchments is questionable. The absence of any commonly accepted look-up tables/chart of these calibration parameters in the literature and inconsistency in the previous estimations makes it even harder for the accurate prediction of wash-off using these models.

Among the external drivers, rainfall data is an essential input in the prediction of sediment wash-off (Sartor and Boyd, 1972; Egodawatta *et al.*, 2007). However, due to

the highly variable nature of rainfall over a wide range of scales, it is not always possible to measure rainfall with an appropriate temporal and spatial resolution required by hydrological modelling applications such as sediment wash-off modelling. Hence the effect of rainfall variability in such resolution is often neglected in these modelling applications. This is common in both rural and urban catchments. But urban areas are characterised by smaller catchment sizes with a higher proportion of impervious area than rural catchments. These factors result in faster catchment reaction times and higher surface run-off volumes than rural catchments. Hence, inadequate representation of any spatial and temporal variability of rainfall can be a source of uncertainty in urban runoff predictions and in any other urban hydrological predictions which are driven by rainfall and/or runoff such as sediment wash-off (Al and Elson, 2005; Segond *et al.*, 2007; Gires *et al.*, 2012; Ochoa-Rodriguez *et al.*, 2015). For example, Ochoa-Rodriguez et al. (2015) showed that the error in peak pipe discharge predictions due to inadequate representation of rainfall spatial variability is up to 250% for drainage areas of the order of 1 ha and up to 50% for drainage areas of ~800 ha.

Rainfall intensity, in addition to being the main input in existing widely used wash-off empirical models, also affect the calibration parameters of these models as the wash-off process is physically driven by rainfall and runoff (Sartor and Boyd, 1972; Deletic *et al.*, 1997; Egodawatta *et al.*, 2007). Hence, any unrepresented rainfall variability would have a direct effect on sediment wash-off prediction. For instance, when spatial and temporal aggregation is used and variability is neglected, the rainfall intensity peaks get smoothed out and these reduced peaks would result in an underestimation in the prediction of sediment wash-off. Hence, it is important to get a measure of this uncertainty in the prediction of wash-off due to this unaccounted variability of rainfall. However, sediment wash-off has not been explored much in terms of the effect of small scale rainfall variability.

The question of why the effect of rainfall variability has not been investigated in depth for sediment wash-off predictions may have multiple answers, but most of them can be brought under two major reasons. (1) The challenges in accurately representing the small-scale rainfall variability in lumped (a single spatial unit) sediment wash-off models which are the most widely used wash-off model type. (2) Current model structures of sediment wash-off models are not complex enough to adequately describe the key physical processes and the model structures do not have provision to study the

propagation of uncertainty due to variability of rainfall. The main challenge with representing the small-scale rainfall variability in lumped catchment models is that such representation requires a method such as geostatistics with demanding data requirements to account for spatial configuration of measurement locations and to correctly quantify uncertainty due to spatial variability. The major challenge with current sediment wash-off modelling is that it still needs an in-depth investigation on the calibration parameters so as to understand how these calibration parameters responds to variability in external drivers such as rainfall. This essentially means that the nature of this PhD requires individual and in-depth investigations of two different areas: small scale rainfall variability and sediment wash-off modelling; both areas pose their own challenges. Hence, the PhD is compartmentalised into three major parts. The first part investigates small scale rainfall variability and develops a rigorous stochastic method which can be used in lumped catchment models including sediment wash-off models to represent this small scale rainfall variability. The second part focuses on developing better understanding of the sediment wash-off process and then to improve current wash-off models by establishing the correlation between the calibration parameters and external drivers including rainfall. The final part uses a comprehensive uncertainty analyses method and the findings from first two parts to study propagation of different sources of uncertainty in the improved sediment wash-off model including uncertainty due to rainfall variability. The aims and objectives of this PhD, described in section 1.2 provide more details on these three major parts.

1.2 Aims and objectives

The overall aim of this PhD is to investigate the effect of uncertainty due to small-scale variability of rainfall in predicting urban sediment wash-off. Sub-aims and corresponding objectives are presented in Table. 1.1

Table 1.1: **Sub-aims (1-3) and corresponding objectives (a-c)**

Sub-aims	Corresponding objectives
• Develop a computational method to estimate the areal rainfall and corresponding uncertainty due to small-scale variability of rainfall that can be used in lumped urban hydrological models including sediment wash-off models	1a. Select suitable spatial and temporal scales and identify a suitable computational method and associated data requirements to quantify the uncertainty associated with areal estimation of rainfall due to its variability at the selected scales
	1b. Collect data from a catchment with spatial and temporal resolution required by the scale defined in (1a)
	1c. Develop a procedure to apply the computational method identified in (1a) to calculate areal rainfall estimates and associated uncertainty for the catchment selected in (1b)
• Improve understanding of sediment wash-off processes from urban surfaces and establish the correlation between model calibration parameters and external drivers in the current wash-off model	2a. Identify the limitations of current sediment wash-off modelling practice with regards to its applicability for different catchment conditions
	2b. Design and carry out a series of experiments to overcome the shortcomings identified from (2a)
	2c. Apply results from (2b) to establish the correlation between model calibration parameters and external drivers in wash-off modelling
• Investigate propagation of different sources of uncertainty including rainfall uncertainty in improved sediment wash-off modelling.	3a. Select a comprehensive uncertainty analyses method to study uncertainty propagation through sediment wash-off modelling
	3b. Improve/refine the sediment wash-off model developed in (2c) further to be able to investigate propagation of different sources of uncertainty
	3c. Use the improved sediment wash-off model from (3b) and uncertainty analyses method selected in (3a) to identify, quantify and separate different sources of uncertainty including model input uncertainty due to rainfall variability quantified in (1)

1.3 Thesis format and structure

1.3.1 Structure

This thesis consists of five other chapters in addition to the introduction chapter. Chapter 2 is a literature review followed by three core technical chapters (Chapter 3-5) each of which is clearly linked to one of the sub-aims mentioned in Table 1.1 in the same order. Table 1.2 provides short descriptions of Chapters 2-5. Chapter 6 is the final chapter which provides an over-arching summary and conclusion of the work described in the thesis.

1.3.2 Format

This thesis is prepared in a format where core chapters (Chapter 3- 5) consist of 2 journal publications and a manuscript partly prepared for publication as listed below. These papers are formatted to fit the format of this thesis and figure, table and equation numbers are changed to aid continuity within this thesis. Written permission from Faculty of Engineering to submit a thesis in this format is attached in Appendix

1. Chapter 3: Paper 1 (published): Muthusamy, M., Schellart, A., Tait, S. and Heuvelink, G. B. M. (2017) 'Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models', Hydrology and Earth System Sciences, 21(2), pp. 1077–1091. doi: 10.5194/hess-21-1077-2017.

2. Chapter 4: Paper 2 (published): Muthusamy, M., Tait, S., Schellart, A., Md, N. A. B., et al. (2018) 'Improving understanding of the underlying physical process of sediment wash-off from urban road surfaces', Journal of Hydrology, 557C, pp. 426–433. doi: https://doi.org/10.1016/j.jhydrol.2017.11.047.

3. Chapter 5: Part of this chapter is a pre-print version of the publication: Muthusamy, M., Wani, O., Schellart, A. and Tait, S.: Accounting for variation in rainfall intensity and surface slope in wash-off model calibration and prediction within the Bayesian framework, Water Res., doi:https://doi.org/10.1016/j.watres.2018.06.022, 2018.

Table 1.2:   Chapters and corresponding objectives

	Objectives								
	1a	1b	1c	2a	2b	2c	3a	3b	3c
Chapter 2 - Literature review This chapter presents a review of the literature on (a) small-scale variability of rainfall and its representation in urban hydrology and (b) Modelling and associated challenges of urban sediment transport including uncertainty analysis to understand the research gaps that needs to be addressed in this PhD	✓			✓			✓		
Chapter 3 – Geostatistical Upscaling of Rain Gauge Data to Support Uncertainty Analysis of Lumped Urban Hydrological Models This chapter presents the development and application of a stochastic computational method that rigorously accounts for spatial configuration of rainfall measurement locations and correctly quantifies uncertainty due to spatial variability of rainfall at a selected scale	✓	✓	✓						
Chapter 4 – Improving Understanding of the Underlying Physical Process of Sediment Wash-off from Urban Road Surfaces This chapter presents the experimental results and consequent mathematical interpretation of these results to report on the development of an improved form of wash off model. The chapter then develops the correlation between calibration parameters and external drivers found in the current wash-off model				✓	✓	✓			
Chapter 5 – Accounting for uncertainty propagation in enhanced sediment wash-off modelling within a Bayesian framework This chapter presents the comprehensive uncertainty propagation analysis of improved sediment wash-off model developed using results from Chapter 4 to identify, quantify and separate different sources of uncertainty including rainfall uncertainty obtained from chapter 3.							✓	✓	✓

1.3.2.1 Authorship

I hereby confirm that I am the primary contributor in the writing of each of listed paper including the design and conduct of the reported research in each paper.

1.3.2.2 Copyright

I hereby confirm that all the necessary permissions have been obtained from relevant publishers to use this published material in this thesis.

2. Literature review

2.1 Small-scale rainfall variability and its representation in urban hydrology

2.1.1 Background: The what and the why

Urban catchments, in comparison to rural catchments, are associated with smaller catchment sizes with higher proportion of impervious area and smaller catchment response time resulting in a larger proportion of rainfall being converted to surface runoff which then reaches the catchment outlet faster. Hence, inadequate representation of any spatial and temporal variability of rainfall is can be one of the main sources of uncertainty in urban runoff predictions and any other urban hydrological prediction which is driven by rainfall and/or runoff such as sediment wash-off (Al and Elson, 2005; Segond *et al.*, 2007; Gires *et al.*, 2012; Ochoa-Rodriguez *et al.*, 2015). For instance, for a completely impervious surface, any uncertainty in rainfall estimation will result in a similar level of uncertainty in runoff prediction peaks, according to the well-established rational formula (Viessman and Lewis, 1995) which is still widely used for estimating design discharge in small urban catchments.

Figure 2.1, extracted from Schilling (1991), presents the spatial and the temporal scales of interest in hydrology. The spatial scales of interest in urban hydrology can go from as small as 10 m^2 to the order of $10^8 \, m^2$. The variability of the rainfall at a scale in the order of > 10^6 m^2 (highlighted in green) is comparably well studied in the literature, thanks to the increased number of rain gauge networks and improved radar technology. But the spatial scale in the order of < 1 km^2 (highlighted in red with darker gradient indicating the decreasing number of literature) is something that has received attention only very recently. But nevertheless, representation of the rainfall variability at this scale (< 1 km^2) was found to be important in urban hydrology especially when the catchment size gets smaller. For example, Ochoa-Rodriguez et al. (2015) showed that the error in peak pipe discharge predictions due to inadequate representation of rainfall spatial variability is up to 250% for drainage areas in the order of 1 ha and up to 50% for drainage areas of ~800 ha. Gires et al. (2012) quantified the uncertainty on urban runoff associated with the unmeasured small-scale rainfall variability i.e. rainfall at a resolution finer than 1 km x 1 km x 5 min. They used downscaling of C-band radar

network data for the 900 ha urban catchment of Cranbrook in London, UK to derive the rainfall fields at scales up to 125 m × 125 m × 1.25 min and showed that the uncertainty due to this rainfall variability on the simulated peak flow in conduits is significant, reaching for some conduits +/-25% and +/-40% for frontal and convective events respectively. Muthusamy et al. (2015) analysed even finer scale of rainfall measurements from an urban catchment of 8 ha and showed that neglecting the rainfall variability at this scale can cause up to around 20% of the variability in the runoff peak prediction. These studies showed that the uncertainty due to the unknown small-scale rainfall variability (<1 km^2) should not be neglected in urban hydrology especially for small urban catchments.
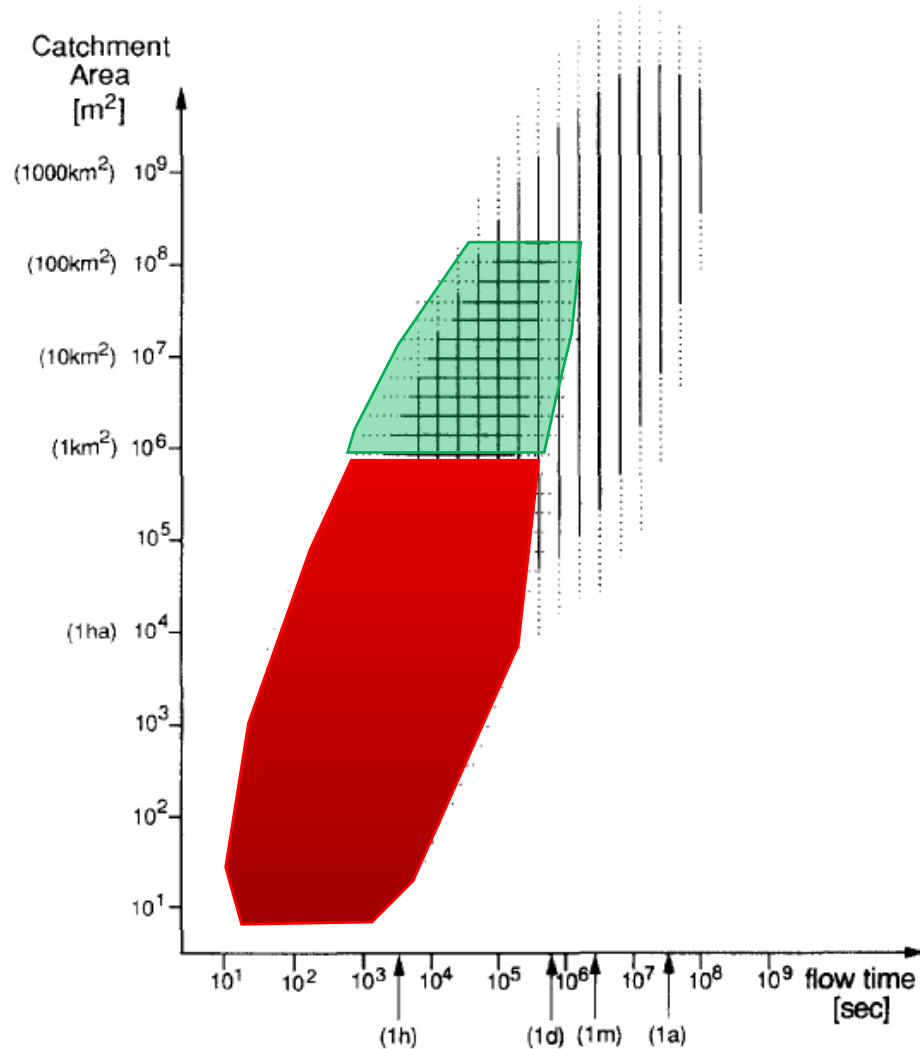
Figure 2.1: Typical catchment area and flow time for urban hydrology (horizontal lines) versus general hydrology (vertical lines) (Schilling, 1991)**.**

9

The main challenge to study the variability of rainfall at sub-kilometre scale is the lack of rainfall data. It is an uncommon practice to measure rainfall data at a sub-kilometre scale even for urban catchments, due to financial and practical (e.g. maintenance and data collection) difficulties. For instance, UK urban drainage guidelines (WaPUG, 2002) suggest a rain gauge resolution of 1 gauge/2km^2 for flat terrain and 1 gauge/km^2 for hilly terrain. Although fulfilling this guideline is already very challenging, this guideline apparently neglects any variability in rainfall at sub-kilometre scale. It is not possible to study this variability at every catchment financially and practically. One possible solution is to investigate this variability for a few selected catchments per region with similar climate conditions and to transfer this knowledge to other catchments in the region using measures such as a variance reduction factor (VRF) (Krajewski *et al.*, 2000; Villarini *et al.*, 2008; Peleg *et al.*, 2013a). In this regard, there have been a few studies, which focus on visualising and quantifying the sub-kilometre rainfall variability. These studies will be discussed further in section 2.1.3. The next section provides a critical overview of measurement methods used to measure rainfall at sub-kilometre scale.

2.1.2 Rainfall measurement

When it comes to measuring rainfall at the sub-kilometre scale, x band radar and rain gauge measurements are the common methods reported in the literature. Radar provides areal estimations of rainfall and rain gauges provide point measurements. The transformation from point estimation to areal estimation or vice versa is possible through upscaling or downscaling techniques respectively as explained in Fig. 2.2. In addition to point and areal measurements, commercial microwave links have also been used recently to estimate path integrated rainfall measurement (Brauer *et al.*, 2011; Overeem *et al.*, 2011), but this method has not been as widely tested compared to radar and point measurement methods.
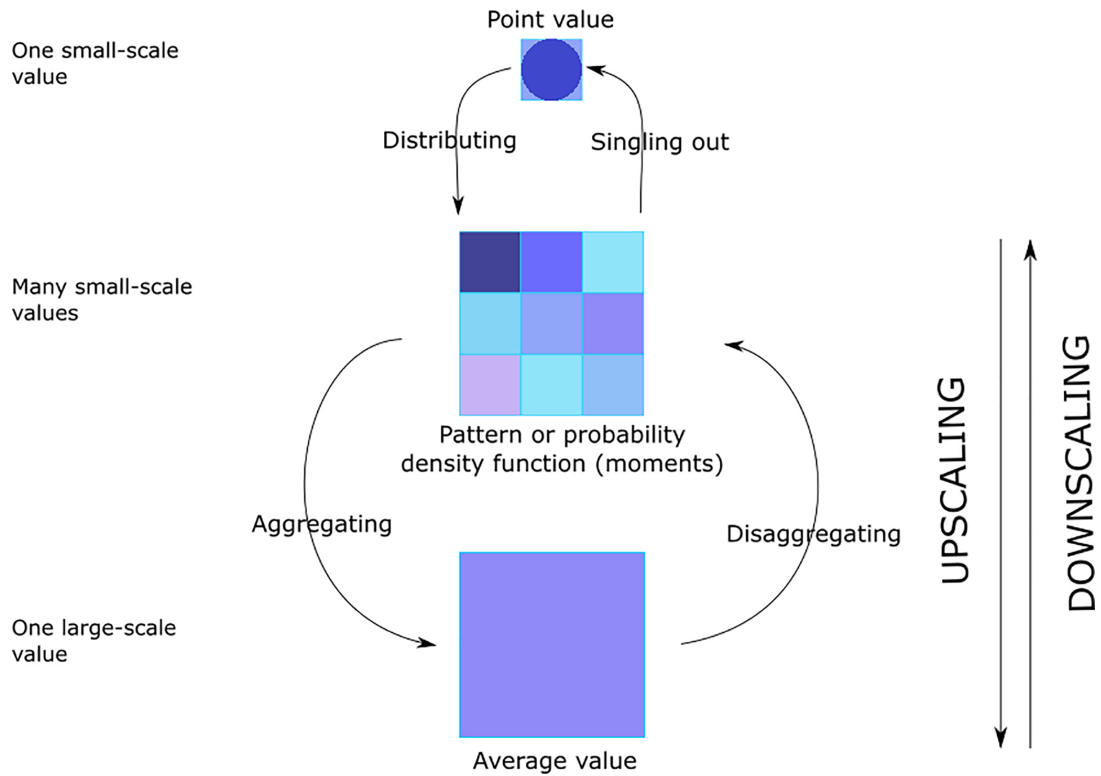
Figure 2.2: Downscaling and upscaling processes (from Cristiano et al. 2016)

Thanks to the advancement in radar technology, X-band radar is able to provide rainfall estimates up to 100 m grid resolution which can be put to use in investigating spatial variability given these predictions are accurate enough for this purpose. But the main problem with radar rainfall data is that the rainfall intensity is estimated based on an indirect relationship using the reflectivity of radar signals (Berne and Krajewski 2013; Cristiano et al. 2016). Hence the radar rainfall estimation needs to be calibrated using direct measurements such as rain gauge data. In this regard, various techniques have been proposed in the literature to calibrate and to merge radar data with direct point measurements (Wood et al. 2000; Gires et al. 2014; Cole and Moore 2008; Smith et al. 2012). Studies also showed that radar data generally underestimate rainfall intensity compared to rain gauges (Overeem *et al.*, 2009; van de Beek *et al.*, 2010; Smith *et al.*, 2012). In addition to this main weakness of radar data, various other sources of uncertainties such as the interaction between beam propagation and topography and hardware issues are discussed in detail in Berne and Krajewski (2013). Hence despite the possible availability of radar data at sub-kilometre scale, the accuracy of the corresponding rainfall estimation that are required by the nature of the study (i.e. investigation of sub-kilometre variability of rainfall) will not be as good as direct

measurements such as data obtained from rain gauges (Berne *et al.*, 2004; Jensen and Pedersen, 2005; Gires *et al.*, 2012; Peleg *et al.*, 2013b; Cristiano *et al.*, 2016). This prompted some researchers to use and recommend dense networks of point measurements to study the spatial variability at sub-kilometre scale (Jensen and Pedersen, 2005; Villarini *et al.*, 2008; Cristiano *et al.*, 2016). In this regard, Table 2.1 presents some of the important studies which analysed spatial variability of rainfall at sub kilometre scale using such dense network(s) of point rainfall measurements.

Although point measurement using rain gauge data are considered to be more accurate and even taken as "ground truth" when calibrating and merging radar data, rain gauge measurements also contain uncertainty due to measurement errors caused by various sources depending on the location and rain gauge types. The conventional vessel type rain gauges include tipping bucket gauges, pluviographs and weighing gauges (Lanza and Vuerich, 2009). Among these, the tipping bucket type is the most widely used rain gauge for a wide range of applications mainly due to its reliability, cost and widely accepted technology (Lanza and Vuerich, 2009). The conventional vessel type rain gauges are subject to errors due to wind, wetting, evaporation, and splashing (Sevruk and Hamon, 1984; Fankhauser, 1998; Dingman, 2015). Modern recording gauges such as optical gauges and disdrometers which do not use a collecting vessel have also been used recently to minimise these errors (Jaffrain and Berne, 2012; Dingman, 2015). In addition to measurement errors, counting based rain gauges such as tipping bucket gauges are also subject to errors due to its sampling mechanism which varies against accumulation time and rainfall intensity (Habib *et al.*, 2001). Habib et al. (2001) investigated the sampling error of tipping bucket type rain gauges and concluded that this error decreases with increasing accumulation time and increasing rainfall intensity. They showed that this sampling error can be as high as 100% for rainfall intensities measured at 1 min accumulation time using a bucket size of 0.254 mm.

Although it is impossible to completely remove the measurement and sampling errors, it is important to minimise them especially when higher accuracy data are required. For instance, Ciach and Krajewski (2006) used a paired rain gauge set-up as a quality control measure to filter out any unreliable measurements as their objective of analysing small-scale variability of rainfall data required high accuracy rainfall data.

2.1.3 Lessons learned from studies based on a dense network of point rainfall measurement

Table 2.1 provides a summary of findings from studies based on a dense network of point rainfall measurement to study sub-kilometre rainfall variability.

Table 2.1: **Studies on sub kilometre spatial variability of rainfall using a dense network of point measurements**

Network information	Reference	Relevant observation(s) and finding(s) to this thesis
25 paired stations over 3×3 km^2 (Tipping bucket type rain gauges) (EVAC PicoNet), Airport area in Oklahoma, USA	(Ciach and Krajewski, 2006)	(i) The larger the accumulation time the better the spatial correlation (ii) The type of dependence of spatial variability on rainfall intensity depends on the threshold value which separates weaker and stronger rainfall types.
16 stations at 1 × 1 km^2 (Disdrometers) Lausanne Campus, Switzerland	(Jaffrain and Berne, 2012)	(i) The larger the accumulation time the better the spatial correlation (ii) The error associated with the use of point measurements as areal estimates at larger scales increases with the size of the domain. (iii) At a domain of 1000 x 1000 m^2, this error corresponds to rainfall intensity estimates at 1 min accumulation time is on the order of 25%
2 networks of 8 stations at 2 × 2 km^2 (Tipping bucket type rain gauges) (Part of HYREX network) Brue basin, UK	(Villarini *et al.*, 2008)	(i) The larger the accumulation time the better the spatial correlation (ii) The measurement errors and the small-scale variability of rainfall substantially decrease for accumulation times larger than five minutes.

Network information	Reference	Relevant observation(s) and finding(s) to this thesis	
9 stations at 500 × 500 m^2 (Optical drop-counting and tipping bucket type rain gauges) Farm and Estuary in Aarhus, Denmark	(Jensen and Pedersen, 2005)	(i)	The measured event based accumulated rainfall indicates up to a 100% variation for some events between neighbouring rain gauges within 500 x 500 m^2
	(Pedersen *et al.*, 2010)	(i)	The coefficient of variation (CV) values decrease with increasing rainfall depths, indicating that the largest variability (up to 77%) is in events with a rainfall depth of less than 5 mm
14 paired stations at 4 km^2 (Tipping bucket type rain gauges) Coastal area in Kibbutz Galed, Isreal	(Peleg *et al.*, 2013a)	(i)	Measurement error increases with accumulation time
		(ii)	The variance reduction factor (VRF), representing the uncertainty from averaging a number of rain stations over 1000 x 1000 m^2, ranged from 1.6% for the 1 min timescale to 0.07% for the daily scale.

The above studies explore various aspects of the spatial variability of rainfall at sub-kilometre scales as well as measurement and sampling errors used for the analyses. In summary:

- The main findings centered on the pattern of spatial rainfall variability and the behaviour of measurement and sampling errors at sub-kilometre scales. The minimum spatial scale studied in the above studies is 100×100 m, thus the spatial variability of rainfall at < 100 m still needs to be investigated.

- Tipping bucket type rain gauges are used in most of the cases, hence findings on measurement and sampling error are mostly relevant to tipping bucket type rain gauges.

  As previously stated the sampling error decreases with accumulation time and found to be reduced substantially for accumulation times more than 5 min. A few studies used paired rain gauges in their network to effectively filter out any unreliable measurements

- With regard to spatial variability of rainfall, one of the most common and obvious observation is that spatial variability decreases with increasing accumulation time. Although the actual quantification of this variability varies between studies, it is shown to be significant (25% - 100%) when the accumulation time is less than 5 min.

The above summary clearly indicates that attention should be paid to sub-kilometre spatial variability in urban hydrology due to two reasons. First, as discussed at the beginning of this chapter, for urban catchments any uncertainty in rainfall estimation will have significant effect on runoff predictions due to a high proportion of impervious area. Second, time steps used in urban hydrologic and hydrodynamic modelling can be as small as 1 min where the level of uncertainty in rainfall intensity can be as high as 100% as shown in the above studies. Since it is impossible to avoid this uncertainty completely, it is important to somehow represent this uncertainty in urban hydrologic and hydrodynamic modelling. This information will help to differentiate input uncertainty from total uncertainty thereby helping to understand other sources of uncertainty due to model parameter and model structure. This estimate of the relative importance of uncertainty sources can help to avoid false calibration and force fitting of model parameters (Vrugt et al., 2008). Although some of the above studies provide quantification of uncertainty in upscaling of rainfall data due to spatial variability and measurement errors, these measures are mostly derived by simply calculating the variance between the measurements which apparently ignores the effect of the spatial correlation structure of the rainfall data and the spatial configuration of the rain gauge locations. Further, the uncertainty in upscaling has not been thoroughly investigated against different time scales or/and different rainfall intensity ranges. Investigation of these research gaps requires a review of the literature on available upscaling methods of point rainfall data.

2.1.4 Rainfall upscaling

The main purpose of rainfall upscaling is to derive the areal estimate of the rainfall volume/ or intensity from point measurements as most of the widely used modelling software products (e.g MIKE, InfoWorks) for water quantity and quantity predictions, including prediction of urban sediment run-off and wash-off, use a lumped catchment model structure (LCM). Hence, upscaling of rainfall data is a standard practice in both academic studies and industrial projects. There are a number of interpolation methods

available and several different techniques have been used in software packages to upscale point rainfall data. The simplest interpolation method is to take the arithmetic average (Chow, 1964) of the point observations within the catchment. But this does not account for the spatial correlation structure of the rainfall data and the spatial configuration of the rain gauge locations. Another commonly used method in hydrological modelling is the nearest neighbour interpolation (Chow, 1964; Nalder and Wein, 1998) which leads to Thiessen polygons. In this method, the nearest observation is given a weight of one and other observations are given zero weights during interpolation. But again this method also assumes a homogeneity of rainfall for a certain spatial extent. There are also other interpolation methods of varying complex levels including inverse distance weighting (Dirks *et al.*, 1998), polynomial interpolation (Tabios III and Salas, 1985), and the Moving Window Regression (Lloyd, 2005). The performance of the methods was found to be catchment dependent and no single method has been shown to be optimal for all catchments and rainfall conditions (Ly *et al.*, 2013). But one common drawback with all the above methods is that they do not provide any information on certainty and accuracy of the predictions as they all are deterministic approaches. As discussed before the quantification of this uncertainty is important in urban hydrological and hydraulic modelling.

Geostatistical methods such as kriging present a solution to this problem by providing an estimate of a spatially representative value and a measure of prediction error. In addition to this capability, these methods also take into account the spatial dependency structure of the measured rainfall data (Mair and Fares, 2011; Ly *et al.*, 2013). This is the main reason why studies focussed on rainfall variability, including the studies listed in Table 2.1, use geo-statistical measures. In the following sections, an introduction to geostatistical methods is provided followed by a discussion on its challenges and potential solutions.

2.1.5 Geostatistical methods

2.1.5.1 Measures of rainfall variability used in geo-statistics

Measures such as the variogram (Berne *et al.*, 2004; Jaffrain and Berne, 2012; Bruni *et al.*, 2015), correlogram (Ciach and Krajewski, 2006; Peleg *et al.*, 2013a; Jewski *et al.*, 2016) and covariance function (Gebremichael and Witold F Krajewski, 2004) are used for visualisation of the spatial correlation of rainfall. In addition, they can also be used

to define the weights of individual measurements when geo-statistical methods such as kriging are used for spatial interpolation. Among these three measures, the variogram requires a less restrictive statistical assumption on the stationarity of the spatial property being sampled (Baecher and Christian, 2003). Hence, the variogram is preferred over the other measures described above.

2.1.5.1.1 Variogram

The semivariogram function, $\gamma(h)$, was originally defined by Matheron (1963) as half the average squared difference between points separated by distance h. It can be formulated as,

$$\gamma(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (z_i - z_j)^2 \qquad (2.1)$$

Where $N(h)$ is the set of all pairwise Euclidean distances $i - j = h$, $|N(h)|$ is the number of distinct pairs in $N(h)$, and z_i and z_j data values at spatial locations i and j respectively. The above formulation is for the omnidirectional semivariograms where h represents a distance measure with magnitude only and not direction. Sometimes, it might be desirable to consider directional semi-variograms where h will be a vector (**h**) with both magnitude and direction.

Once valid empirical estimates of the theoretical semivariance are given by Eq.(2.1), it is then necessary to choose a type of theoretical variogram[1] model based on that estimate. Commonly used theoretical variogram shapes rise monotonically as a function of distance and this shape is typically characterised in terms of three parameters namely nugget, sill and range (Isaaks and Srivastava, 1989). These parameters are depicted on the generic variogram model shown in Fig. 2.2. They are defined as follows:

- Nugget - The nugget is the value of the semi-variance at a near-zero distance. It is often greater than zero because of random measurement error and micro-scale spatial variation.
- Range: The range is the distance beyond which the data are no longer spatially correlated.

---

[1] Although by definition variogram is 2γ(h), terms variogram and semivariogram are often used interchangeably. For conciseness we refer γ(h) as variogram throughout this thesis.

- Sill: The sill is the maximum variogram value and equal to the variance of the variable of interest

Some of the most commonly used theoretical variogram types include exponential, Gaussian, power and Spherical (Journel and Huijbregts, 1978; Christakos, 1984; Cressie, 1993). The Nugget variogram is a specific case of the variogram model when the range is zero, i.e. there is no spatial correlation even at zero distance. Typically, a more suited model is selected to fit the empirical variogram in an automated manner using methods such as least square, maximum likelihood, and robust methods (Cressie and Hawkins, 1980). Alternatively, a suitable model can also be selected based on visual inspection of the empirical semi-variogram (Cressie and Hawkins 1980a).
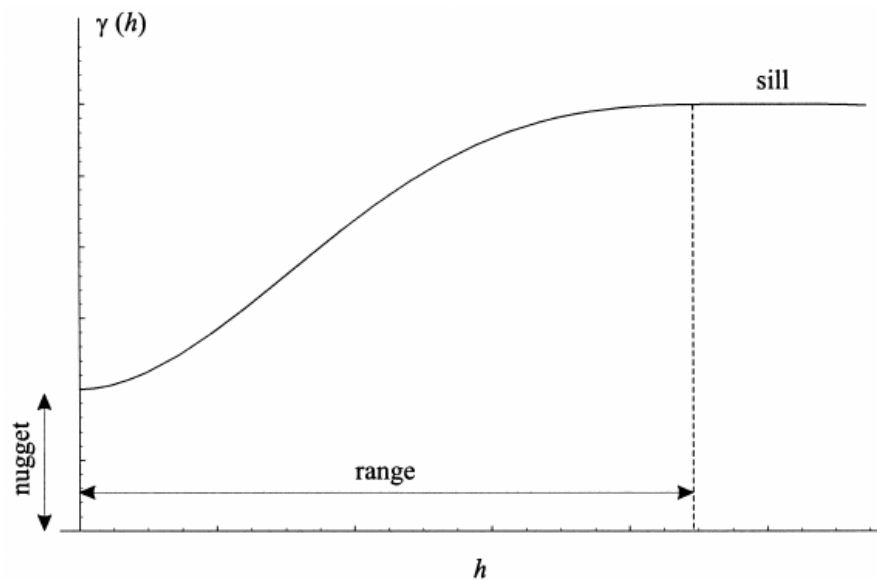
Figure 2.3: A generic variogram model

2.1.5.2 Geostatistical upscaling

A geostatistical model of rainfall intensity r at any location x can be written as (Isaaks and Srivastava, 1989):

$$r(x) = p(x) + \varepsilon(x) \qquad (2.2)$$

where $p(x)$ is the trend (explanatory part) and $\varepsilon(x)$ is the stochastic residual (unexplanatory part). The stochastic term ε is spatially correlated and characterised by a variogram model. Solving the above equation and obtaining the areal estimation of rainfall together with associated uncertainty using spatial stochastic simulation is

explained in detail in Chapter 3. The following section describes some of the common challenges in applying geostatistical upscaling methods to point rainfall data.

2.1.5.3 Challenges associated with Geostatistical method

Even with the advantages discussed previously, geostatistical methods such as block kriging are rarely used in areal rainfall estimation in LCM due to their complexity and demanding data requirements. Since these methods are heavily statistical based encompassing multiple parameters, the amount of spatial data required for model inference is much higher compared to the simpler deterministic methods (Dawson and Gerritsen, 2013). For instance, the precision of any variogram estimation to model spatial correlation strongly depends on the number of observations, i.e. the number of spatial data points of the study area. Webster and Oliver (2007) recommended around 100 measurement points to accurately estimate an anisotropic variogram. But catchments, especially those at small urban scales, do not contain as many measurement locations as the above recommendation.

In addition to the above requirement on the amount of data, geostatistical approaches especially kriging methods works better when the data is approximately normally distributed (Matheron, 1973). In particular, quantile and probability maps created using kriging assume that the data comes from a multivariate normal distribution. The assumption of normality of rainfall intensity data, like any other hydrological data, is not realistic and they are often positively skewed (Hirsch and Slack, 1984; Cecinati et al., 2017). If the empirical distribution of the data is skewed then the kriging estimators are sensitive to a few large data values (Matheron, 1973). Further, when data is skewed or has extremely high or low values, estimated variograms often exhibit erratic behaviours (Gringarten and Deutsch, 2001). Various alternatives to the traditional variogram, such as madograms, rodograms, general and pairwise relative variograms, have been proposed in the literature (Genton 1998; Cressie and Hawkins 1980b), but these alternative measures cannot serve as input for subsequent estimation or simulation algorithms (Gringarten and Deutsch, 2001). Instead, it is recommended to transform the data to the Gaussian domain before performing variogram calculations (Gringarten and Deutsch, 2001; Cecinati *et al.*, 2017). Among the transformation methods, Normal Quantile Transformation (NQT, also known as normal scaling transformation, Goovaerts (1997)) is a widely used method to map a variable to the Gaussian distribution. It has been applied in many hydrological applications (Bogner,

Pappenberger, and Cloke 2012, Montanari, A., & Brath 2004, Bogner, Pappenberger, and Cloke 2012). The concept of NST is to match the p quantile of the data distribution with the p quantile of the standard normal distribution as shown in Fig.2.4. Detailed description of NST including the steps involved can be found in Bogner et al. (2012); Van der Waerden (1953);and Weerts et al. (2011). The advantages of using such a transformation are (1) the difference between extreme values is reduced and (2) after the transformation the theoretical sill is known to be 1 making it easier for comparison of variograms (Gringarten and Deutsch, 2001).
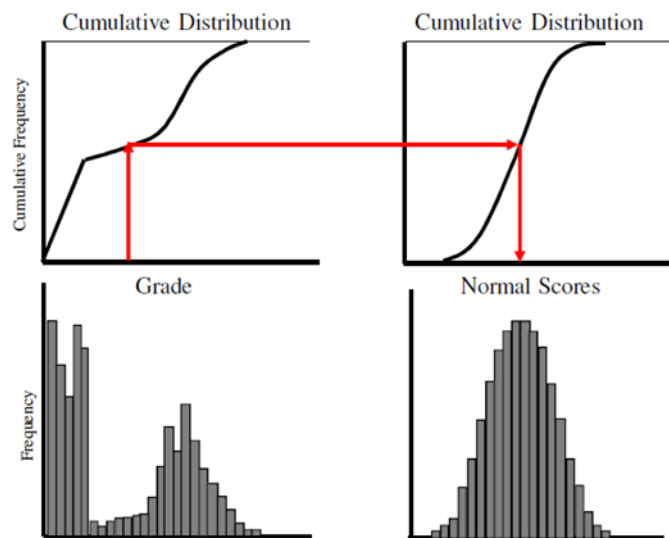
Figure 2.4: Sketch showing the transformation procedure for the normal quantile transformation (from Wilde 2010)

2.2 Urban sediment transport: Modelling and associated challenges

2.2.1 Background: The what and the why

Urban non-point sources are one of the major causes of pollution of surface water bodies (Guy, 1970; Collins and Ridgeway, 1980; Mitchell *et al.*, 2001; Lawler *et al.*, 2006). Such pollution consists of a complex mixture of materials including sediments, hydrocarbons, heavy metals, salts and nutrients. Among these pollutants, sediment washed off from the urban surfaces is one of the major causes of water quality issues of inland water bodies in urban areas mainly due to its ability act as a transport medium to many contaminants (Guy, 1970; Collins and Ridgeway, 1980; Chiew and Vaze, 2004). Collins and Ridgeway (1980) found that in urban storm events, sediment smaller than < 63µm itself transported more than 50% of the total pollutant load. Table 2.2

modified from Mitchell et al. (2001) discusses sources of urban sediment and effects in receiving waters in detail.

Table 2.2: Source, characteristics and effect of urban surface sediment. Modified from Mitchell et al. (2001)

Sources and characteristics of urban catchment surface sediment	Effect in receiving waters

• The major sources of sediment in the urban environment are atmospheric deposition, natural weathering, and construction sites.
• Atmospheric deposits range from large colloids such as windblown sand to small particulates such as PM10 (smaller than 10 μm) arising from vehicle emissions.
• Other sources include particulates deposited from vehicles (e.g. rust, rubber), abrasion of the road and building surfaces, permeable urban surfaces such as gardens and parks, application of de-icing salts, organic detritus, litter and a range of other wastes.
• Particle size highly varies between sites. e.g.
  o 200-1000 μm with a d_{50} of about 300-400 μm - based on a literature review on sediment size by Bertrand-Krajewski et al. (1993)
  o 63 μm – 6mm with a d_{50} of around 400 μm – based on sediment collected from six different road sites in London, UK by Butler and Clark (1995)
	• Sediments are an important mechanism for the transport of other pollutants which may cause problems related to toxicity, eutrophication, and suitability for recreational or portable use.
	• Particle size is important as half the pollutant load may be transported bound to particles of < 63μ, around 6% of the total sediment load.
	• Sediments can be detrimental to water quality even when chemically inert. They cause turbidity, inhibit visual feeders, blanket fish spawning sites and feeding areas, eliminate prey organisms, reduce light penetration and photosynthesis of aquatic plants, cause gill abrasion and fin rot, and scouring causes destruction of bed and bank habitat.

In addition to the effect on receiving water bodies, large amounts of sediment wash-off can also contribute to urban floods by depositing in drainage systems and reducing the hydraulic capacity of these systems that are designed to rapidly carry water away from roads and properties (Ivan, 2001). Sediment deposition will also cause problems such

as pipe surcharge, early overflows, large pollutant discharges and costly removal (Ashley *et al.*, 1992; Delleur, 2001; Heal *et al.*, 2006). The erosion of sediments in sewers can also release pollutants in concentrations that exceed the levels found in the various contributing sources of the sediments and pollutants (Ashley *et al.*, 1992). Hence accurate modelling of sediment wash-off from urban surfaces is important in water quality and flood risk based decision making for an urban catchment.

2.2.2 Modelling of urban sediment transport processes

Modelling sediment transport from urban catchment surfaces is not a straightforward exercise as it needs an understanding of complex underlying physical processes which involves parameters with highly variable nature associated with characteristics of rainfall, catchment surface and sediment particles. Hence most of the models which are in practice to predict sediment transport from urban catchment surfaces are developed from the results from experimental data with empirical parameters. With regards to the complexity, the modelling approach ranges from simple event mean concentration calculations (EMC models, Debo and Reese 1995) to more sophisticated BUWO (Build-Up and Wash-Off, Sartor and Boyd 1972) models.

2.2.2.1 Event mean concentration (EMC)

The EMC is the total mass of pollutant discharged for a given runoff event, divided by the runoff volume for that event (Debo and Reese, 1995). Assessments made using the EMC method have a low cost, and often can address a reasonable range of pollutants, and are sufficiently flexible to address a number of development scenarios (Mitchell *et al.*, 2001). Many EMC values are also provided with measures of variance (e.g. coefficient of variation), hence the associated uncertainty can be calculated (Novotny and Olem, 1994). But the major drawback of EMC models is that the intra-event dynamics in pollutant load cannot be captured.

2.2.2.2 Build-Up and Wash-Off (BUWO) models

There are mainly two stages involved in the process of sediment transport from urban surfaces (Sartor and Boyd, 1972). They are:

1. Build up – is the stage in which sediment accumulates on impervious surfaces
2. Wash-off – is the stage by which accumulated sediment is removed from the impervious surfaces and incorporated in the flow by the actions of rainfall and runoff.

A more sophisticated modelling approach of modelling sediment transport from the urban surfaces is by using BUWO models where sediment build-up is modelled first in order to subsequently model wash-off. Hence two separate equations, often empirical, are employed in this approach. This approach was first suggested by Sartor and Boyd (1972) based on wash off experiments conducted during the summer of 1970 in Bakerfield, California, USA which is considered to be the earliest controlled street experiments on wash-off. The tests used artificial rainfall and were conducted on concrete and old and new asphalt. Amongst their most important conclusion was that the build-up and the wash-off processes should be treated separately. In this regard, the following sections review the most commonly used build-up and wash-off models.

2.2.3 Modelling of build-up process

Sartor and Boyd (1972) found that the majority of sediments accumulated within a metre of the curb, and they estimated curb loadings for residential, commercial and industrial land-uses. According to their findings, the accumulated load of sediment (commonly known as build-up) exponentially varies against antecedent dry days and at one point it reaches its maximum where an equilibrium between the supply of pollutant and their removal by wind effects, traffic, street sweeping, and the natural degradation of solids. This relationship is written as (Alley and Smith, 1981):

$$\frac{dw}{dt} = a_1 - a_2 w \qquad (2.3)$$

Where, w is the amount of sediment on impervious areas (g), a_1 is a constant rate of sediment deposition (g/day), a_2 is a rate constant for sediment removal (day $^{-1}$), and t is time in days. The parameter a_2 can account for losses due to wind and vehicles as well as the biological and chemical decay.

Integration of Eq. (2.3) yields

$$w = A[1 - \exp(-a_2 T)] \qquad (2.4)$$

Where $A\ (=a_1/a_2)$ is the maximum amount of sediment on the impervious areas (g) and T is time since last period of street sweeping or storm runoff (accumulation time, days). a_1 and a_2 need to be calibrated for each catchment.

Although this approach of modelling build-up process as a function of the number of antecedent dry days has been used in some models (e.g. Bertrand-Krajewski et al.,

1993), it has also been criticised, especially in recent studies (Charbeneau and Barrett, 1998; He et al., 2010; Shaw et al., 2010). Among these studies, Shaw et al. (2010) found that the mass of washed-off particulate matter during a storm event is not strongly related to the time between storm events. This was confirmed by He et al. (2010) as they also could not find any relationship between the event mean values of total suspended solids and the length of the antecedent dry weather period for a semi-arid, urban residential catchment in Calgary, Alberta. Shaw et al. (2010) also pointed out that the build-up process is highly dynamic due to unpredicted occurrences of activities like construction work or the input of wind-blown debris from storms and argued that assuming that the build-up process was continuous and a simple function of antecedent dry days will not capture the dynamic nature of the sediment build up process. Shaw et al. (2010) even questioned the necessity of modelling build-up separately as he claimed a constant (initial) load model produced very similar predictions of wash-off loads. In summary, the criticisms of modelling sediment build-up processes as a function of antecedent dry days have been provided with supporting evidence from a range of catchments. But, on the other hand, to answer the question of whether there is a need to model build-up at all, raised by the above studies, particularly Shaw et al. (2010), a more detailed investigation is required. It is because one would expect a variation in wash-off load corresponding to a variation in initial load (modelled using a build-up model) and it is conceptually hard to accept that a constant (initial) load model would produce acceptable wash-off predictions under varying catchment and rainfall conditions. In fact, according to the most widely used exponential model to predict wash-off, discussed in the next section, the wash-off load varies proportionally against the initial load.

2.2.4 Modelling of wash-off process

Wash-off loads have been estimated by using variables such as:

1. Total runoff volume (e.g. Characklis et al. 1979),
2. Total event rainfall (e.g. Reinertsen 1981)
3. Runoff rate (e.g. Wischmeier 1969)
4. Rainfall intensity (e.g. Coleman 1993)

   Or a combination of these variables

But out of these variables aggregate measures such as total runoff volume, and total event rainfall will not be able to predict intra-event load dynamics and will not be suited

24

to study the effect of spatial and temporal variability of rainfall on sediment loads. Between rainfall intensity and runoff rate, modelling approaches using rainfall intensity has been found to produce a better prediction of wash-off (Vaze and Chiew, 2003). It can partly be due to rainfall intensity data's added ability to capture the effect of the impact energy of rainfall drops on sediment mobilisation (Shaw *et al.*, 2010), although not necessarily explicitly included in the wash-off equations. Another advantage of using rainfall data is that it provides a practical means of predicting pollutant loads as it is one of the most readily available data (Francey *et al.*, 2011).

Unlike the build-up process, the wash-off process has been investigated relatively extensively in the literature due to its direct link to urban water quality modelling. Hence, there have been more than one model recommended in the literature. They are discussed in detail in the following section.

2.2.4.1   The exponential model

The exponential wash-off model, originally proposed by (Sartor and Boyd, 1972), is the most widely used method to predict the sediment wash-off. The original exponential wash-off equation proposed by Sartor and Boyd (1972) is given below

$$w_t = w_0\big(1 - e^{-kit}\big) \qquad (2.5)$$

where w_t is transported sediment load after time t, w_0 is initial load of the sediment on the surface; i is rainfall intensity; and k is the wash-off coefficient. Several authors have refined and/or adapted these equations for a range of software in order to get better calibration and simulation of experimental data. Table.2.3 summarises selected studies focused on the improvement of original exponential wash-off model. However, most of the these refinements are very site-specific and not easily generalised. Also, most of these studies paid attention to one single parameter in isolation, thereby ignoring the effect and interactions of the other parameters. For instance, although the multiplication of a capacity factor suggested by Egodawatta et al. (2007), has been shown to be a meaningful modification, it has only been investigated against rainfall intensity. Hence the effect of multiple parameters and their interaction in the underlying processes of sediment wash-off still need to be investigated in a systematic and an integrated way. Considering the influencing parameters, another interesting observation is the lack of attention given to the surface slope in the above studies. Two processes that drive

sediment mobilisation are impact energy from rainfall drops (Coleman, 1993) and shear stress from the overland flow (Akan, 1987; Deletic *et al.*, 1997) both of which are sensitive to surface slope, especially the latter. With the exception of Nakamura, (1984) none of the above studies paid attention to the effect of the slope. Nakamura (1984) results show that *k* increases with surface slope, but this study was based only on two randomly selected slopes and was not extensive enough to be used in subsequent studies or in practical applications.

Table 2.3:   Summary of the selected studies focused on the improvement of original exponential wash-off model

Reference	Modification/Suggestion	Effect/Reason
(Ammon, 1979; Sonnen, 1980; Nakamura, 1984)	*k* needs to be calibrated for surface characteristics, rainfall and runoff characteristics and particle size	-
(Huber and Dickinson, 1992)	Suggestion of *k* value of 4.6 in^{-1} (0.18 mm^{-1})	To replicate that the first one-half inch of total run-off in one hour would wash-off 0.9 fraction of the initial load
(Huber and Dickinson, 1992)	Introduction of an exponent ω for rainfall intensity, ranging from 0.8 to 2 and with mean values about 1.4 to 1.8	Better prediction of peak values
(Zug *et al.*, 1999)	Introduction of two terms in the expression of *k* with different exponent values	Better calibration in general
(Alley, 1981; Egodawatta *et al.*, 2007)	multiplicative capacity factor on the right side of the Eq.(2.5) which varies with rainfall intensity	To redefine the maximum available sediment to be washed off

2.2.4.2   Other models

Apart from the most widely used exponential wash-off, there are also some other modelling equations proposed by a few authors. Among these models, the power model originally proposed by Duncan (1995) has been used in a number of studies. The original form of this model is given below

$$w_t = \sum_{i=1}^{n} b_1 I^{b_2} \qquad (2.6)$$

Where I = rainfall intensity as recorded in each of the n time steps over a period t; and b_1 and b_2 = calibration coefficients; n = number of time steps over a period t. I is calculated by assuming that total rainfall depth, recorded in one-time step, occurred just within that very time step. Therefore, I strongly depends on the resolution of rainfall records. The above model was formulated based on the regression modelling approach and the cumulative nature of this model represents the ongoing input of energy produced by raindrop impact (Francey *et al.*, 2011). The applicability of this model or a very similar approach (e.g.: Σ runoff rate models, square of the rainfall intensity) has been tested in a few studies (Vaze and Chiew, 2003; Brodie, 2007; Francey *et al.*, 2011). But, in comparison to the exponential model this model has not been tested under a wide range of catchment conditions. Similarly, a few other modelling approaches such as MOSQITO (Moys *et al.*, 1988) and P/r Model (Zhao *et al.*, 2015) have also not been widely tested and proven better than the exponential model for a wide range of catchment conditions in comparison to the exponential wash-off model. Among these models, MOSQITO, although explicitly includes effects of both rainfall and runoff in the wash-off process, the results, however, did not show much of an improvement compared to the exponential model mainly due to the lack of knowledge about the calibration parameters (Bertrand-Krajewski, 2006).

2.2.4.3 Effect of spatial variability of rainfall in urban sediment wash-off modelling

To study the effect of the spatial variability of rainfall on urban sediment wash-off there could possibly be two ways. A more sophisticated and comprehensive way is to develop a 2D sediment transport model of the catchment/surface of interest where areal rainfall can be used as an input. This way the spatial variability of rainfall over the catchment/surface can be explicitly captured (Shaw *et al.*, 2010). But the problem with this method is the application of the most commonly used sediment transport models discussed in sections 2.2.4.1 and 2.2.4.2 have not been widely tested with 2D grid type models, hence their performance is still unknown. Further, a 2D model would need a significantly higher computational power compared to a lumped catchment models without necessarily adding any significant improvement in the predictions. These issues lead to the other way to study the effect of spatial variability of rainfall in

sediment wash-off models which is capturing the spatial variability in a stochastic way and use that information in lumped catchment models. The use of the geo-statistical methods to capture the uncertainty due to spatial variability in an aerial estimation of rainfall has already been discussed in section 2.1. Once such estimation is obtained an uncertainty propagation method can be employed to investigate the effect on the wash-off prediction. This approach needs a comprehensive and holistic uncertainty analysis where different sources of uncertainty in wash-off modelling can be identified, investigated and separated. In this regard, previous literature on uncertainty analyses of sediment wash-off is discussed in the following section.

2.2.5 Uncertainty analyses in modelling of urban sediment transport

Urban sediment wash-off models, like any other mathematical models, represent only a part of reality which leads to uncertain results. Since the complete elimination of this uncertainty is not possible (Beck, 1987) it is important at least to quantify it in order to measure the reliability of the model results. This information can be used in risk-based decision making in water quality control measures. In this regard, adequate treatment of uncertainties in model calibration and prediction is currently a heavily researched issue in hydrology. However, there are limited studies on uncertainty specifically focusing on wash-off modelling. In this regard, Dotto et al. (2012) compared a number of uncertainty analysis techniques in urban stormwater quality modelling and found the Bayesian approach (using a sampling technique, such as Markov Chain Monte Carlo (MCMC) sampling), although computationally demanding, to be one of the preferable techniques. The following section provides an overview of the application of Bayesian inference in urban sediment wash-off modelling in literature.

2.2.5.1   Bayesian inference using Markov Chain Monte Carlo (MCMC) sampling

Bayesian inference is a method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available (Halls-Moore, no date). During a Bayesian calibration, the joint probability density of parameter and model results, the product of the prior of the parameters and the likelihood, is conditioned on the data (Del Giudice et al. 2013).

The analytical evaluation of Bayesian inference becomes more challenging when the models require a large number of parameters, as is often the case in hydrology (Dotto *et al.*, 2012; Del Giudice *et al.*, 2013). To cope with this, numerical techniques such as

Markov Chain Monte Carlo (MCMC) simulations have widely been used (Dotto *et al.*, 2012; Del Giudice *et al.*, 2013; Bonhomme and Petrucci, 2017). Further, Bayesian inference requires the definition of the likelihood function and the prior distribution of the parameters. The prior distribution of parameters is usually derived from relevant literature and by using practical constraints. Likelihood functions (also known as the sampling model) often formulated as a combination of a deterministic model (e.g. Eq (2.7)) and probabilistic error models (Del Giudice *et al.*, 2013). Identically independent distributed (IID) error is the most commonly used form of error model in urban hydrology (Freni *et al.*, 2009; Dotto *et al.*, 2011; Breinholt *et al.*, 2012; Bonhomme and Petrucci, 2017) mainly because of its simplicity. When the error model is assumed to be IID the observed system output Y (e.g. measured wash-off) can be formulated as

$$Y(x, \omega, sd.E) = y(x, \omega) + E(sd.E) \qquad (2.7)$$

Where x is external driving forces (e.g. rainfall intensity), ω is deterministic model parameters (e.g. wash-off coefficient, k), y is the deterministic model output (e.g. predicted from Eq. (2.5)). $sd.E$ is standard deviation of IID error to account for measurement noise of the system response, E. The main drawback of IID model is that it requires an absence of the serial correlation in error distribution, which can lead to underestimation of uncertainty and biased parameter estimates (Del Giudice *et al.*, 2013). This makes IID less robust for different urban hydrological applications. Hence, recent studies show an autoregressive bias error model, as shown in Eq (2.8), is a better representation of the error process observed in many hydrological processes.

$$Y(x, \omega, sd.B, l, sd.E) = y(x, \omega) + B(sd.B, l) + E(sd.E) \qquad (2.8)$$

Where $sd.B, l$ are the standard deviation and correlation length which characterise the autoregressive stationary random process which accounts for errors due to model structural deficit (also known as model bias, B) in addition to measurement noise, E modelled as IID. An autoregressive error model represents model structural deficit better than IID as it accounts for the "memory" in error time series (Del Giudice *et al.*, 2013). This autoregressive bias error model was originally suggested in generic statistical applications (Craig *et al.*, 2001; Kennedy and O'Hagan, 2001; Higdon *et al.*, 2004; Bayarri *et al.*, 2007) and later adapted for environmental engineering applications (Reichert and Schuwirth, 2012).

A few previous studies applied Bayesian inference in sediment wash-off modelling to assess the build-up/wash-off model performance by using long-term continuous road runoff turbidity measurements (Bonhomme and Petrucci, 2017), to compare different error model and sampling technique in the calibration of wash-off modelling (Egodawatta *et al.*, 2014; Sage *et al.*, 2016) and to assess build-up/wash-off model performance at the scale of the urban catchment (Bonhomme and Petrucci, 2017). The most striking feature of the Bayesian approach which has been used in these studies is that it helps to identify different sources of uncertainty such as parameter uncertainty, model bias and measurement error and consequently, helps to separately analyse them, though it requires knowledge about the error process (Dotto *et al.*, 2012). This estimate of the relative importance of uncertainty sources can help to avoid false calibration and forced fitting of model parameters (Vrugt *et al.*, 2008). Furthermore, it helps to better understand the propagation of uncertainty in external drivers (e.g. rainfall uncertainty) through the model which is also an area that has not been explored in depth with regards to wash-off modelling.

2.3 Summary of findings

Inadequate representation of the spatiotemporal variability of rainfall has been found to be one of the main sources of uncertainty in urban runoff predictions and other urban hydrological predictions which are driven by rainfall and/or runoff due to higher impervious area and smaller catchment response time associated with urban catchments(Al and Elson, 2005; Segond *et al.*, 2007; Gires *et al.*, 2012; Ochoa-Rodriguez *et al.*, 2015). The variability of the rainfall at a scale in the order of $> 10^6$ m^2 is comparably well studied in the literature, thanks to the increased number of rain gauge networks and improved radar technology. But the spatiotemporal variability of rainfall at the sub-kilometre scale has not been investigated in detail in comparison to larger scales mainly due to lack of data, but there have been number of studies which show that this variability should not be neglected in urban hydrology especially for small urban catchments in the order of < 10 km^2. In this regard, it is shown in literature that despite the possible availability of radar data at sub-kilometre scale, the accuracy of the corresponding rainfall estimation that are required for the investigation of sub-kilometre variability of rainfall will not be as good as direct measurements such as rain gauges (Berne *et al.*, 2004; Jensen and Pedersen, 2005; Gires *et al.*, 2012; Peleg *et al.*,

2013b; Cristiano *et al.*, 2016). Consequently, high-density point measurement networks at sub kilometre scale have been recommended in the recent literature to study spatial variability of rainfall. There have been a limited number of studies in which sub-kilometre rainfall variability is analysed based on point measurement networks. From these studies it can be concluded,

- o The spatial variability of rainfall at < 100 m still needs to be investigated.
- o Geostatistical measures such as the variogram have been the most preferred way to present the spatial correlation structure of rainfall.
- o The spatial variability needs to be investigated at shorter temporal averaging interval (< 5 min) as the variability increases with decreasing temporal averaging interval
- o Since the investigation of sub-kilometre rainfall variability requires high accuracy data it is important to give attention to measurement and sampling errors.
- o Tipping bucket type rain gauges have been the most widely used rain gauges in many studies and the corresponding sampling error decreases with accumulation time. Since this error is as significant as spatial variability at shorter temporal averaging interval (< 5 min), this error need to be taken into account at these at such temporal averaging intervals.
- o Although the actual quantification of the uncertainty due to spatial variability and measurement error varies between studies, it has been shown to be significant (25% - 100% CV in intensity predictions) at shorter temporal averaging interval (< 5 min).

Since it is impossible to avoid the uncertainty due to sub-kilometre variability and measurement and sampling error completely, it is important to represent this uncertainty in urban hydrological predictions which are driven by rainfall and/or runoff. Urban sediment transport is also a process that is mainly driven by rainfall impact energy (Coleman, 1993) and runoff shear stress (Akan, 1987b; Deletic et al., 1997) and it is one of the major source of pollutants in an urban environment as sediment act as a transport medium to many contaminants. Urban sediment deposition also causes problems such as pipe surcharge, early overflows, large pollutant discharges and costly removal (Ashley et al., 1992; Delleur, 2001; Heal et al., 2006). Since rainfall is the main physical driver of sediment transport any uncertainty in

rainfall estimation is very likely to have a direct impact on sediment transport from urban surfaces. However, the effect of rainfall uncertainty has not been investigated in detail in the literature. One of the main reasons being that the quantification of the uncertainty in upscaling rainfall in a stochastic way that is representable in urban sediment transport modelling has not been investigated in depth in the literature. In this regard, geo-statistics is an appropriate computational tool mainly due to its ability to estimate the uncertainty associated with upscaling of rainfall data by taking into account the spatial correlation structure of the rainfall data. But the application of geo-statistical methods such as kriging to rainfall data is challenging mainly due to the non-normality of rainfall data and data scarcity. A transformation of rainfall data to the Gaussian domain is necessary as quantile and probability maps created using kriging assume that the data comes from a multivariate normal distribution. In this regard, Normal Quantile Transformation was shown to be a suitable method to transform rainfall data to the Gaussian domain to overcome the problem of non-normality.

Another reason why the effect of rainfall variability has not been investigated in urban sediment transport is the poor understanding of how the calibration parameters in the widely used empirical models to predict urban sediment transport are linked with the underlying physical processes. Hence, before investigating the effect of small scale variability of rainfall sediment transport predictions, it is important to understand the limitations of current modelling approach of sediment transport. It has been shown in the literature, that there are mainly two processes involved in sediment transport from an impervious surface, Build-up in which sediment accumulates and Wash-off where the accumulated sediment gets removed by rainfall and runoff. Traditionally build-up which is essentially the initial load for wash-off prediction has been modelled mainly using antecedent dry days, but it has also been criticised recently as recent literature shows that wash-off loads are relatively insensitive to the time between storm events. Despite these criticisms, the effect of build-up on wash-off modelling has not been investigated in detail for various catchment conditions. Hence the question of whether there is a need to model build-up at all, raised by the recent studies, still remains unanswered and so needs further investigation.

Among the modelling approaches of wash-off, the exponential model has been the most widely used and tested model for a wide range of catchment conditions associated with

characteristics of sediment, surface and rainfall intensity. Although numerous studies focused on improving/redefining the original exponential model (Eq. (2.5)) one of the major drawbacks in most of these studies is the lack of an integrated and systematic approach. Such approaches are vital because of the complex interaction between the influencing parameters associated with the characteristics of rainfall, sediment and catchment surface. Considering the influence of surface slope on the main underlying processes of sediment wash-off which are impact energy (from rainfall) and shear stress (from runoff), there have only been one study which focuses on the surface slope in sediment wash-off modelling. Hence, the effect of surface slope on sediment wash-off needs further investigation. Due to difficulty in modelling this complex interactions using first principals, most widely used wash-off models are empirically derived using laboratory and field experimental data (e.g. Sartor and Boyd, 1972; Egodawatta et al., 2007; Francey et al., 2011). The inherent shortcoming of these empirical methods is that the calibration parameters in these models (e.g. k in Eq 2.5) does not have much physical meaning and therefor applicability and transferability of these parameters to other catchments is limited. In general, these parameters are estimated for different catchments by using a combination of empirical look-up tables/charts and interpolation/extrapolation of existing data. However, with the absence of such commonly accepted look-up tables/chart in the literature and inconsistency in the previous estimations of the calibration parameters, the modellers are forced to use a constant value for parameters regardless of catchment conditions. This calls for an alternative and a more transparent way of estimating the calibration parameters

Commonly used wash-off models have not widely been tested in 2D models. Hence, as discussed before, the other potential way of studying the effect of spatial variability on sediment wash-off is capturing the spatial variability of rainfall in a stochastic way and then use that information in a lumped catchment sediment wash-off model together with an error propagation method. This approach needs

  a. A stochastic method developed to capture the rainfall variability for a spatial scale of interest
  b. A comprehensive uncertainty analysis where different sources of uncertainty in wash-off modelling can be identified, investigated and separated.

Among the uncertainty methods used in urban water quality models, the Bayesian approach although computationally demanding has been shown to be comprehensive

as it helps to identify different sources of uncertainty such as parameter uncertainty, model bias and measurement error and consequently helps to separately analyse them. The Bayesian approach requires, in addition to prior knowledge, the definition of the likelihood function. Likelihood functions often formulated as a combination of a deterministic model and a probabilistic error model. Although identically independent distributed (IID) is the most used form of error model in hydrologic application due to its simplicity, recent studies show that the autoregressive bias error model, although more complicated, is a better representation of the error process observed in hydrological modelling.

3. Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models

Manoranjan Muthusamy[1], Alma Schellart[1], Simon Tait[1], Gerard B.M. Heuvelink [2]

[1] Department of Civil and Structural Engineering, University of Sheffield, Sheffield, S1 3JD, UK
[2] Soil Geography and Landscape group, Wageningen University, Wageningen, 6700, The Netherlands

Correspondence to: Manoranjan Muthusamy (m.muthusamy@sheffield.ac.uk)

Abstract

In this study, we develop a method to estimate the spatially averaged rainfall intensity together with associated level of uncertainty using geostatistical upscaling. Rainfall data collected from a cluster of eight paired rain gauges in a 400×200 m^2 urban catchment are used in combination with spatial stochastic simulation to obtain optimal predictions of the spatially averaged rainfall intensity at any point in time within the urban catchment. The uncertainty in the prediction of catchment average rainfall intensity is obtained for multiple combinations of intensity ranges and temporal averaging intervals. The two main challenges addressed in this study are scarcity of rainfall measurement locations and non-normality of rainfall data, both of which need to be considered when adopting a geostatistical approach. Scarcity of measurement points is dealt with by pooling sample variograms of repeated rainfall measurements with similar characteristics. Normality of rainfall data is achieved through the use of Normal Score Transformation. Geostatistical models in the form of variograms are derived for transformed rainfall intensity. Next spatial stochastic simulation which is robust to nonlinear data transformation is applied to produce realisations of rainfall fields. These realisations in transformed space are first back-transformed and next spatially aggregated to derive a random sample of the spatially averaged rainfall intensity. Results show that the prediction uncertainty comes mainly from two sources: spatial variability of rainfall and measurement error. At smaller temporal averaging intervals both these effects are high, resulting in a relatively high uncertainty in prediction. With longer temporal averaging intervals the uncertainty becomes lower due to stronger spatial correlation of rainfall data and relatively smaller measurement error. Results also show that the measurement error increases with decreasing rainfall intensity resulting in a higher uncertainty at lower intensities. Results from this study can be used for uncertainty analyses of hydrologic and

hydrodynamic modelling of similar sized urban catchments as it provides information on uncertainty associated with rainfall estimation, which is arguably the most important input in these models. This will help to better interpret model results and avoid false calibration and force-fitting of model parameters.

Keywords: Geostatistical upscaling, spatial stochastic simulation, areal average rainfall intensity, hydrological modelling, uncertainty

3.1 Introduction

Being the process driving runoff, rainfall is arguably the most important input parameter in any hydrological modelling study. But it is a challenging task to accurately measure rainfall due to its highly variable nature over time and space, especially in small urban catchments. Despite recent advances in radar technologies rain gauge measurements are still considered to be the most accurate way of measuring rainfall, especially at short temporal averaging intervals (< 1 h), which are of most interest in urban hydrology studies (Ochoa-Rodriguez *et al.*, 2015). However, many commonly used urban hydrological models (e.g. SWMM, HBV) are lump catchment models (LCM) where time series of areal average rainfall intensity (AARI) are needed as model input. Therefore, point observations of rainfall need to be scaled up using spatial aggregation in order to be fed in to a LCM. There are a number of interpolation methods available for spatial aggregation and used in the various LCM to scale up point rainfall data. The simplest method is to take the arithmetic average (Chow, 1964) of the point observations within the catchment. But this method does not account for the spatial correlation structure of the rainfall and the spatial organisation of the rain gauge locations. Another commonly used method in hydrological modelling is the nearest neighbour interpolation (Chow, 1964; Nalder and Wein, 1998) which leads to Thiessen polygons. In this method the nearest observation is given a weight of one and other observations are given zero weights during interpolation, thereby ignoring spatial variability of rainfall to a certain extent. There are also other methods, with varying complexity levels, including inverse distance weighting (Dirks *et al.*, 1998), polynomial interpolation (Tabios III and Salas, 1985) and moving window regression (Lloyd, 2005). The predictive performance of the above methods are found to be case-dependent and no single method has been shown to be optimal for all catchments and rainfall conditions (Ly *et al.*, 2013). One common drawback with all the above methods

36

is that they do not provide any information on the uncertainty of the predictions of AARI as all the methods are deterministic. The uncertainty in prediction of AARI mainly comes from two sources; uncertainty due to measurement errors and uncertainty associated with spatial variability of rainfall. The characteristics of measurement errors can vary depending on the rain gauge type. For example, errors associated with commonly used tipping bucket rain gauges range from errors due to wind, wetting, evaporation, and splashing (Sevruk and Hamon, 1984; Fankhauser, 1998) to errors due to its sampling mechanism (Habib *et al.*, 2001). In addition to measurement errors and since rainfall can vary over space significantly, any spatial aggregation method for scaling up the point rainfall measurements incorporates more uncertainty (Villarini *et al.*, 2008). The magnitude of the uncertainty depends on many factors including rain gauge density and location, rainfall variability, catchment size, topography and the spatial interpolation technique used. Quantification of the level of uncertainty is essential for robust interpretation of hydrological model outputs. For instance, the absence of information on uncertainty can lead to force fitting of hydrological model parameters to compensate for the uncertainty in rainfall input data (Schuurmans and Bierkens, 2006).

Geostatistical methods such as kriging present a solution to this problem by providing a measure of prediction error. In addition to this capability, these statistical methods also take into account the spatial dependence structure of the measured rainfall data (Mair and Fares, 2011; Ly *et al.*, 2013). Although these features make geostatistical methods more attractive than deterministic methods, they are rarely used in LCM due to their inherent complexity and heavy data requirements. Since they are statistical methods encompassing multiple parameters the amount of spatial data required for model inference is higher compared to deterministic methods. In addition the underlying assumption of geostatistical approaches typically requires data to be normally distributed (Isaaks and Srivastava, 1989). In general, catchments, especially those at small urban scales, do not contain as many measurement locations as required by geostatistical methods. Furthermore, rainfall intensity data are almost never normally distributed, especially at smaller averaging intervals (< 1 h) (Glasbey and Nevison, 1997). Despite these challenges geostatistical methods can provide information on uncertainty associated with predicted AARI. This capability can be utilised in uncertainty propagation analysis in hydrological models. In literature, geostatistical methods have been used to analyse the spatial correlation structure of

rainfall at various spatial scales (Berne *et al.*, 2004; Ciach and Krajewski, 2006; Emmanuel *et al.*, 2012; Jaffrain and Berne, 2012), however its application to support uncertainty analyses of upscaling rainfall data has not been explored.

In this paper we present a geostatistical approach to derive AARI and the level of uncertainty associated with it from observations obtained from multiple "paired" rain gauges located in a small urban catchment. The proposed approach presents solutions to the above described challenges of geostatistical methods. First, it uses pooling of sample variograms of rainfall measurements at different times but with similar characteristics to increase the number of paired observations used to fit variogram models. Second, a data transformation method is employed to transform the rainfall data to obtain a normally distributed data set. The level of uncertainty in the prediction of AARI is then quantified for different combinations of temporal averaging intervals and intensity ranges for the studied urban catchment. We focused on a small urban catchment with a spatial extent of less than a kilometre given the findings of recent research on the high significance of unmeasured spatial rainfall variability at such spatial scales, especially for urban hydrological and hydrodynamic modelling applications (Gires *et al.*, 2012; Ochoa-Rodriguez *et al.*, 2015)

3.2 Data collection

3.2.1 Location and rain gauge network design

The study area is located in Bradford, a city in West Yorkshire, England. Bradford has a maritime climate, with an average yearly rainfall of 873 mm recorded from 1981-2010 (MetOffice, UK). The rain gauge network, used in this study was located at the premises of Bradford University (Fig. 3.1) and rainfall data were collected from paired tipping bucket rain gauges placed at eight locations covering an area of 400×200 m^2. Data used in this study were collected from April, 2012 to August, 2012 and from April, 2013 to August, 2013. These stations were located on selected roofs of the university buildings, thereby providing controlled, secure and obstruction-free measurement locations. Each station consists of two tipping bucket type rain gauges mounted 1 m apart. On each roof the paired gauges were placed such that the height of the nearest obstruction is less than two times the distance between the gauges and the obstruction The rim of each rain gauge was set up around 0.5 m above the surrounding ground level following UK standard practice (MetOffice, UK). An example of the measurement

setup (Station 6) is also shown in Fig. 3.1. A histogram of the inter-station distances of the rain gauge network is presented in Fig. 3.2. Lag distances covered in this network are distributed between 21 m (St. 4-St. 5) and 399 m (St. 1-St. 3).
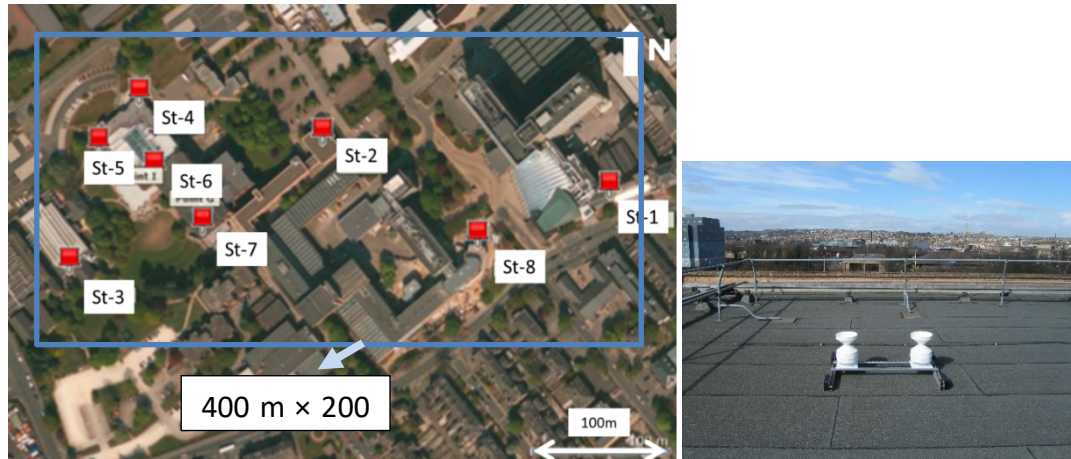
Figure 3.1: (Left) A aerial view of - rain gauge network covering an area of 400×200 m^2 at Bradford University, UK. (Right) A photograph of paired rain gauges at station 6.
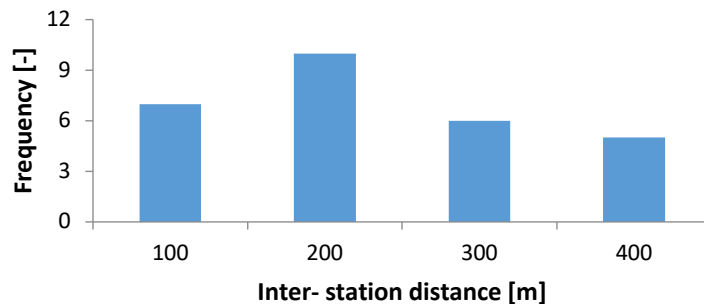
Figure 3.2: Histogram with class interval width of 100 m showing frequency distribution of inter-station distances (m)

All rain gauges are ARG100 tipping bucket type with an orifice diameter of 254 mm and a resolution of 0.2 mm. Dynamic calibration was carried out for each individual gauge before deployment and visual checks were carried out every 4-5 weeks during the measurement period to ensure that the instruments were free of dirt and debris. Data loggers were reset every 4-5 weeks during data collection to avoid any significant time drift. Measurements (number of tips) were taken every minute and recorded on TinyTag data loggers mounted in each rain gauge.

39

Quality control procedures were performed prior to statistical analysis, taking advantage of the paired gauge setup to detect gross measurement errors. The paired gauge design provides efficient quality control of the rain gauge data records as it helps to identify the instances when one of the gauges fails, and to clearly identify periods of missing or incorrect data (Ciach and Krajewski, 2006). During the dynamic calibration of all rain gauges in the laboratory before deployment, it was identified that the highest and lowest values of the calibration factors for the tipping bucket size are 0.196 mm and 0.204 mm. The gauges were recalibrated in the laboratory after the first period of measurement and it was found that the largest change in calibration factor for any gauge was a maximum of 4% of the original calibration factor. Therefore a maximum difference of 4% in volume per tip was assumed to be caused by inherent instrument error. It was therefore decided that this is the maximum acceptable difference between any pair of gauges. Sets of cumulative rainfall data corresponding to specific events from the paired gauges were checked against each other and if the (absolute) difference in cumulative rainfall was greater than 4%, that complete set was identified as unreliable and removed from further analysis.

3.2.2 Characteristics of the data

The total average network rainfall depth for the summer seasons of 2012 and 2013 are 538 mm and 207 mm, respectively. Figure 3.3 shows time series of daily rainfall averaged over the network for 2012 and 2013. There is a significant difference in cumulative rainfall between 2012 and 2013. This is because 2012 was the wettest year recorded in 100 years in the UK (MetOffice UK, 2016) and 558 mm of rainfall during 2012 summer was unusually high. An average rainfall of only 360 mm was recorded during April to September over the 1981 - 2010 period at the nearest operational rain gauge station at Bingley, which is around 8 km from the study site with a similar ground elevation (MetOffice UK, 2016).
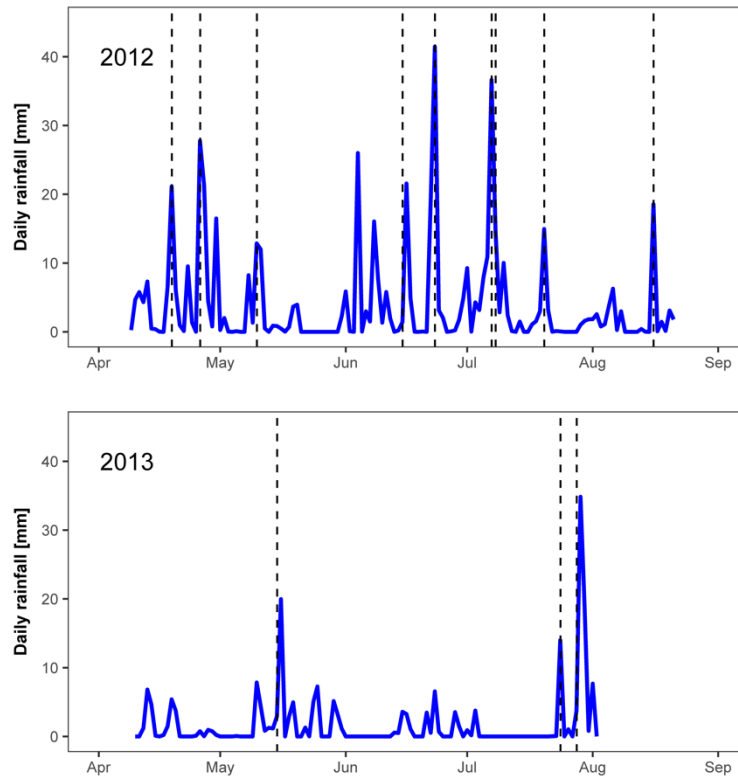
Figure 3.3: Time series of network average daily rainfall in the two seasons of 2012 and 2013 with vertical dashed lines indicating the events presented in Table 3.1

The data set for 2012 and 2013 contains 13 events yielding more than 10 mm network average rainfall depth each and lasting for more than 20 min. A summary of these events is presented in Table 3.1. Note that this event separation is only used for the presentation of results in chapter 4. Hence it does not leave out any data from the development and calibration of the geostatistical model as presented in chapter 3. Table 3.1 shows that the total event duration ranges from 1.5 h to 11.4 h while the event network average rainfall intensity varies from 1.79 mm/h to 7.96 mm/h. Table 3.1 also includes summary statistics of peaks of events (temporal averaging interval of 5 min) for the eight stations within the network. Although the spatial extent of the area is only 400×200 m^2, it is clear that there is a considerable difference in rainfall intensity measurements indicated by the standard deviation and range of peaks observed in the individual events. The maximum standard deviation between peaks of individual events is 9.27 mm/h for event 8, which is around 12.5% of the mean network peak intensity of 74.4 mm/h. This variation provides evidence of the potential importance of analysing uncertainty in the estimation of AARI even in such a small urban catchment.

Table 3.1:   Summary of events which yielded more than 10 mm rainfall and lasted for more than 20 min with summary statistics of event peaks (derived at 5 min temporal averaging interval) from all stations.

Event ID.	Date	Network average duration (h)	Network average intensity (mm/h)	Network average rainfall (mm)	Summary statistics of peaks between different stations (mm/h)			
					Mean	Std. Dev	Max	Min
1	18/04/2012	6.33	2.20	13.9	5.10	0.550	6.02	4.74
2	25/04/2012	6.42	2.55	16.3	7.05	0.751	8.32	5.92
3	09/05/2012	8.92	1.79	16.0	5.10	0.537	5.97	4.74
4	14/06/2012	6.83	1.99	13.6	5.25	0.636	6.04	4.74
5	22/06/2012	11.4	2.39	27.3	12.7	1.72	15.4	9.67
6	06/07/2012	4.42	5.31	23.4	38.5	4.52	42.9	30.5
7	06/07/2012	3.25	3.23	10.5	7.20	0.679	8.46	5.93
8	07/07/2012	1.50	7.84	11.8	74.4	9.27	86.5	61.9
9	19/07/2012	3.08	3.35	10.3	12.7	2.01	14.5	9.74
10	15/08/2012	2.00	7.96	15.9	43.0	3.69	47.8	37.5
11	14/05/2013	7.92	2.14	17.0	8.08	1.20	9.55	6.09
12	23/07/2013	1.75	6.51	11.4	37.7	2.09	42.6	35.7
13	27/07/2013	8.17	4.34	35.5	26.6	1.23	27.5	23.8

3.3   Methodology

Figure 3.4 summarises the procedure of geostatistical upscaling of the rainfall data adapted in this study in a step-by-step instruction followed by the detail descriptions of each step. This complete procedure was repeated for temporal averaging intervals of 2 min, 5 min, 15 min and 30 min in order to investigate the effect of temporal aggregation on the prediction of AARI. The entire ten months of collected data were used for the development and calibration of the geostatistical model.
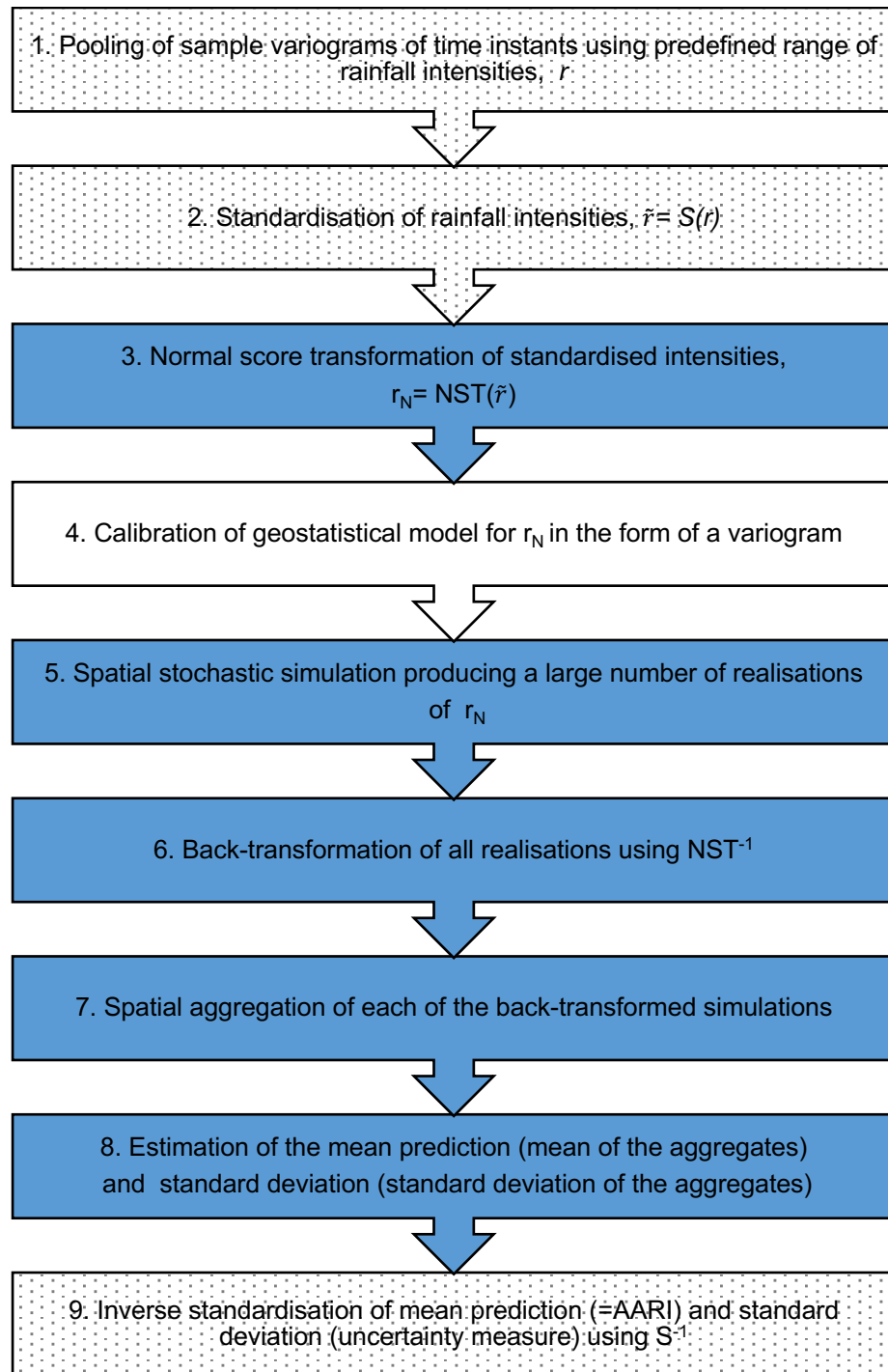
Figure 3.4: Step by step procedure developed in this study to predict AARI and associated level of uncertainty. Boxes highlighted in dots indicate the steps to resolve the problem of scarcity in measurement locations, blue boxes show the steps introduced to address non-normality of rainfall data.

3.3.1 Step 1: Pooling of sample variograms

The rain gauge network contains eight measurement locations. These eight measurement locations give 28 spatial pairs at a given time instant which yields too few spatial lags than would normally be used in geostatistical modelling. For example, Webster and Oliver (2007) recommends around 100 measurement points to calibrate a geostatistical model. The procedure adapted in this study increases the number of pairs by pooling sample variograms for time instants with similar rainfall characteristics. With n measurement locations and measurements taken at t time instants, the pooling over t time instants creates $t \times \frac{1}{2} \times n \times (n-1)$ spatial pairs. Although this procedure increases the number of spatial pairs by a factor t, the spatial separation distances for which information is available will be limited to the original configuration of the n measurement locations.

The underlying assumption of this pooling procedure is that the spatial variability over the pooled time instants is the same. Therefore, it is important to pool sample variograms of rainfall measurements with similar rainfall characteristics. Since the spatial rainfall variability is often intensity-dependent (Ciach and Krajewski, 2006), the characteristics of a less intense rainfall event may not be the same as that of a high intensity rainfall event. Hence to make the assumption of consistency of spatial variability, the range of rainfall intensity over the pooled time instants should be reasonably small. On the other hand, one should also make sure that there are enough time instants within a pooled subset to meet the data requirement to calibrate the geostatistical model. Based on the above two criteria, three rainfall intensity classes were selected. The maximum threshold value was limited to 10mm/hr to have enough time instants for the highest range (i.e. > 10 mm/hr) in order to produce stable variograms even at 30 min temporal averaging interval. It was then decided to divide the 0 – 10 mm/hr class to two equal subclasses (i.e. < 5mm/hr and 5-10 mm/hr). This resulted in three subclasses, which is a reasonable number given the size of the data set and computational demand. The number of time instants (t) within each rainfall intensity class is presented for three temporal averaging intervals in Fig. 3.5. The natural characteristic of rainfall data results in the dominance of lower intensity rainfall (0.1-5.0 mm/h) over the recording period. In addition, the number of time instants t obviously reduces with increasing temporal averaging intervals due to the aggregation process. As a consequence there are only seven time instants for the intensity range >

10 mm/h at the 30 min temporal averaging interval. This limits the maximum temporal averaging interval to 30 min for our analyses. For a catchment of this size (400 × 200 m^2) it is very unlikely to have a response time of more than 30 min. Hence, from a hydrological point of view consideration of temporal averaging intervals longer than 30 min would not be sensible. Note that although there are only seven time instants, the pooling procedure will produce 196 (=7×28) points to calculate and calibrate the geostatistical model for that intensity class.
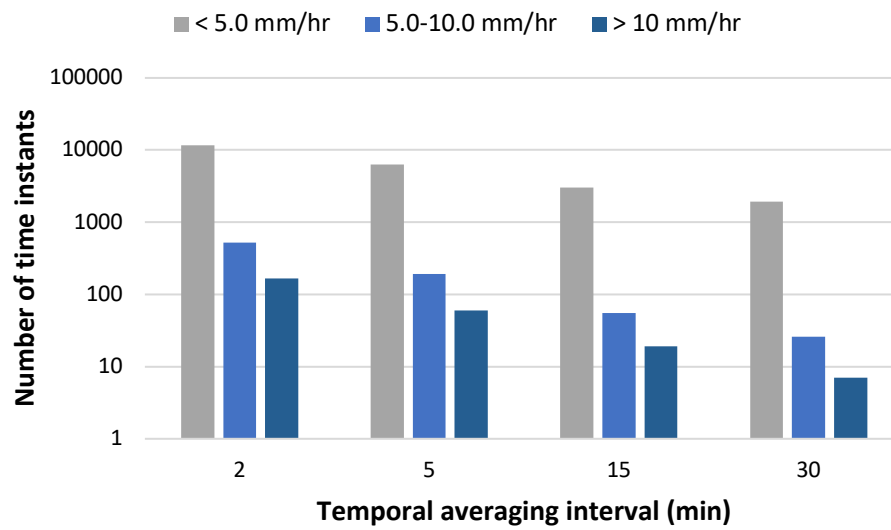
Figure 3.5: Number of time instants for each temporal averaging interval and rainfall intensity class combination.

3.3.2 Step 2: Standardisation of rainfall intensities

Having chosen the rainfall intensity classes to create pooled time instants, there can still be inconsistency in spatial variability between time instants within a class and therefore assuming a single geostatistical model for the whole subset may not be realistic. To reduce this effect to a certain extent, all observations within an intensity class were standardised using the mean and standard deviation of each time instant as follows:

$$\tilde{r}_{ix} = \frac{r_{ix} - m_i}{sd_i} \qquad (3.1)$$

45

where $i=1 \ldots t$, $x=1 \ldots n$; $\tilde{r}_{ix}$ is standardised rainfall intensity at a time instant i and location x; r_{ix} is rainfall intensity at time instant i and location x and; m_i, sd_i are mean and standard deviation of rainfall intensities at time instant i, respectively. Further steps were carried out on the standardised rainfall intensity.

3.3.3 Step 3: Normal transformation of data

The upper part of Fig. 3.6 shows the distribution of standardised rainfall intensity for a temporal averaging interval of 5 min derived using Eq. (3.1). From the figure it is clear that the data are not normally distributed. Distributions for other temporal averaging intervals (i.e. 2 min, 15 min and 30 min) show a similar behaviour. But the geostatistical upscaling method to be used is based on the normal distribution. This requires the rainfall data to be normally distributed prior to the calibration of the geostatistical model. The Normal Score Transformation (NST, also known as normal quantile transformation (Van der Waerden, 1953)) is a widely used method to transform a variable distribution to the Gaussian distribution. It has widely been applied in many hydrological applications (Bogner et al., 2012; Montanari andBrath, 2004; Todini, 2008; Weerts et al., 2011). The concept of NST is to match the p-quantile of the data distribution with the p-quantile of the standard normal distribution. Consider a standardised rainfall intensity $\tilde{r}$ with cumulative distribution $F_{\tilde{R}}(\tilde{r})$. It is transformed to a r_N value with a Gaussian cumulative distribution $F_{R_N}(r_N)$ as follows:

$$r_N = F_{R_N}^{-1}\big(F_{\tilde{R}}(\tilde{r})\big) \qquad (3.2)$$

Detailed description of NST including the steps involved can be found in Bogner et al. (2012), Van der Waerden (1953) and Weerts et al. (2011). The lower part of Fig. 3.6 shows the transformed standardised intensity for the temporal averaging interval of 5 min.
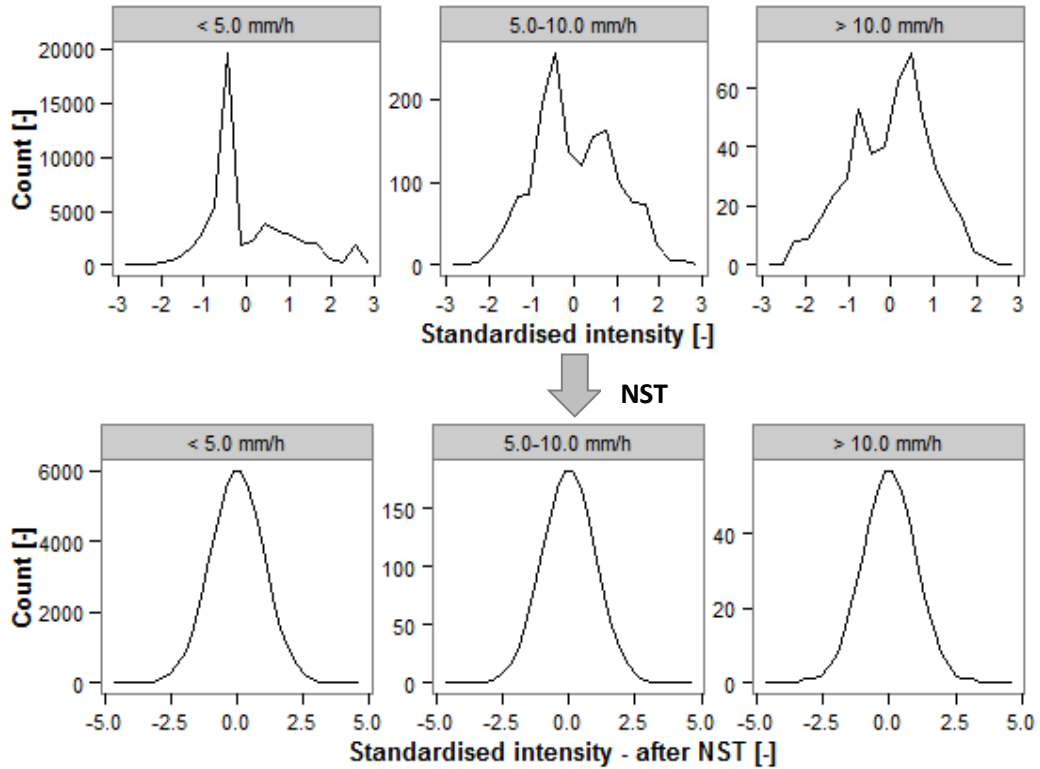
Figure 3.6: Distribution of standardised rainfall intensity for different rainfall intensity classes at a temporal averaging interval of 5 min before (upper part) and after (lower part) normal score transformation (NST).

3.3.4 Step 4: Calibration of Geostatistical model

A geostatistical model of (normalised) rainfall intensity r_N (derived from Section 3.3.3) at any location x can be written as:

$$r_N(x) = p(x) + \varepsilon(x) \qquad (3.3)$$

where $p(x)$ is the trend (explanatory part) and $\varepsilon(x)$ is the stochastic residual (unexplanatory part). Considering the availability of data, small catchment size and scope of this study, it was assumed that the trend is constant and does not depend on explanatory variables (e.g. topography of the area, wind direction). The stochastic term ε is spatially correlated and characterised by a variogram model. A variogram model typically consists of three parameters; nugget, sill and range (Isaaks and Srivastava, 1989). The nugget is the value of the semi-variance at near-zero distance. It is often greater than zero because of random measurement error and micro-scale spatial variation. The range is the distance beyond which the data are no longer spatially

47

correlated. The sill is the maximum variogram value and equal to the variance of the variable of interest (Isaaks and Srivastava, 1989)

3.3.5 Step 5: Spatial stochastic simulation

The assumption of a constant trend makes that the spatial interpolation can be solved using an ordinary kriging system (Isaaks and Srivastava, 1989):

$$\sum_{y=1}^{n} w_y \, \gamma_{xy} - \mu = \gamma_{xz} \quad \forall \, x = 1, \dots, n \qquad (3.4)$$

$$\sum_{x=1}^{n} w_x = 1 \qquad (3.5)$$

where $w_x, x = 1, \dots, n$ are ordinary kriging weights, γ_{xy} is the semivariance between rainfall intensities at locations x and y , γ_{xz} is the semivariance between rainfall intensities at location x and prediction location z, and μ is a Lagrange parameter.

Once the ordinary kriging weights are calculated using Eq. (3.4) and Eq. (3.5), point rainfall intensities can be predicted using point kriging at any given point by taking the weighted average of the observed rainfall intensities, using the w_x as weights. In this case we need a change of support from point to block as our intention is to predict the average rainfall intensity over the catchment. This is usually done by predicting at all points inside the catchment and integrating these over the catchment. This procedure is known as block kriging (Isaaks and Srivastava, 1989), which also has provisions for calculating the prediction error variance of the catchment average. But the procedure of NST as explained in Section 3.3.3 also involves back-transformation of kriging predictions to the original domain at the end (Step 6). Since this transformation is typically non-linear, the back-transform of the spatial average of the transformed variable that is obtained from block kriging is not the same as the spatial average of the back-transformed variable; we need the latter and not the former. In principle, we could predict at all points within the block, back-transform all and next calculate the spatial average, but standard block kriging software implementations do not support this and neither is it possible to compute the associated prediction error variance. Hence block kriging cannot be applied. The alternative used in this study is to apply a computationally more demanding spatial stochastic simulation approach, which involves generation of a larger number of realisations and spatial averaging of these

realisations. Unlike kriging, spatial stochastic simulation does not aim to minimize the prediction error variance but focuses on the reproduction of the statistics such as the histogram and variogram model (Goovaerts, 2000). The output from spatial stochastic simulation is a set of alternative rainfall realisations ('possible realities'). The mean of a large set of realisations approximates the kriging prediction, while their standard deviation approximates the kriging standard deviation. We used the sequential Gaussian simulation algorithm which involves the following steps (Goovaerts, 2000):

i. Define a prediction grid (a 25m × 25m regular grid in this case);

ii. Visit a randomly selected grid cell that has not been visited before and predict the transformed rainfall intensity at the grid cell centre using ordinary kriging, this yields a kriging prediction and a kriging standard deviation;

iii. Use a pseudo-random number generator to sample from a normal distribution mean equal to the kriging prediction and standard deviation equal to the kriging standard deviation and assign this value to the grid cell centre;

iv. Add the simulated value to the conditioning data set, in other words treat the simulated value as if it were another observation;

v. Go back to step ii and repeat the procedure until there are no more unvisited grid cells left.

The five steps above produce a single realisation. This must be repeated as many times as the number of realisations required (500 in this study). It must also be repeated for each time instant, which explains that the computational burden can be high. Implementation of these steps with the gstat package in R (Pebesma, 2004) is straightforward.

The grid size and number of simulations (i.e., the sample size) were selected considering the spatial resolution of available measurements and computational demand. It was observed that neither a finer grid nor more simulations improved the results significantly. Increasing the resolution to 10 m × 10 m only reduces the standard deviation of the prediction by less than 5% in most cases while making the computational time six times higher (a summary on computation power is presented in Appendix 3B).

3.3.6 Step 6-9: Calculation of AARI and associated uncertainty

Once the realisations have been prepared these are back-transformed by applying the inverse of Eq. (3.2) to all grid cells (step 6). Some values derived from spatial stochastic simulation were outside the transformed data range. Hence during back transformation (step 6) of these values linear extrapolation was used. These linear models were derived using a selected number of head and tail portion of normal Q-Q plot. This is one of the simplest and most commonly used solutions for NST back-transformation (Weerts *et al.*, 2011; Bogner *et al.*, 2012). Considering the scope of this study and the relatively small number of data which had to be extrapolated, other extrapolation methods were not explored. After step 6, the back transformed realisations are spatially averaged one by one (step 7). This yields as many spatially averages as the number of realisations that had been generated in step 5. This set of values is a simple random sample from the probability distribution of the catchment average rainfall. Thus, the sample mean and standard deviation provide estimates of the mean and standard deviation of the distribution, respectively (step 8). Finally, by doing the inverse standardisation of the mean and standard deviation of the distribution to account for step 2, the AARI and associated uncertainty measure (standard deviation) were derived (Step 9).

3.4 Results and Discussion

3.4.1 Calibration of the geostatistical model of rainfall

As explained in Section 3.4, the geostatistical model of transformed rainfall data were calibrated using variograms for three different intensity ranges. This procedure was repeated for temporal averaging intervals of 2 min, 5 min, 15 min and 30 min. Exponential models were fitted to empirical variograms. The resulting variograms are presented in Fig. 3.7.
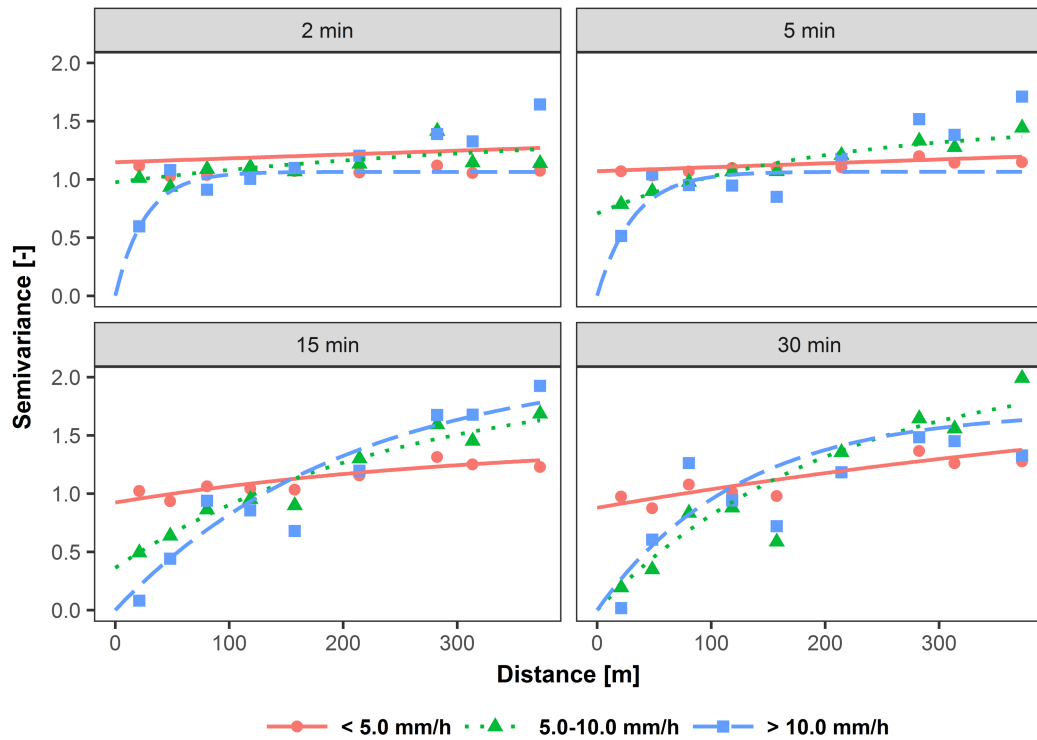
Figure 3.7: Calculated variograms for each intensity class within each temporal averaging interval

The variograms illustrate two properties of the collected rainfall measurements; spatial variability of rainfall, and measurement error. One of the main parameters which characterises these properties is the nugget. Theoretically at zero lag distance the variance should be zero. However most of the variograms exhibit a positive nugget effect (generally presented as nugget-to-sill ratio) at zero lag distance. This nugget effect can be due to two reasons; random measurement error and microscale spatial variability of rainfall. Unfortunately we cannot quantify these causes individually using the variograms. But there is a consistent pattern of nugget against both rainfall intensity class and temporal averaging interval which helps to interpret the variograms.

Considering the behaviour of nugget-to-sill ratio against rainfall intensity class, it can be observed that the smaller the intensity the higher the nugget-to-sill ratio, regardless of temporal averaging interval. For example, at 2 min averaging interval the nugget-to-sill ratio increases from zero to almost one (nugget variogram) as the rainfall intensity class changes from > 10 mm/h to < 5 mm/h. The pure nugget variogram at < 5 mm/h means that either there is no spatial correlation at the regarded distance, or the spatial correlation of the field cannot be detected by the measurements because of the

51

measurement error. Looking at the behaviour of nugget-to-sill ratio against temporal averaging interval, Fig. 3.7 shows that the smaller the averaging interval the higher the nugget-to-sill ratio, regardless of rainfall intensity class. For example, for rainfall intensity class 5.0–10.0 mm/h the nugget-to-sill ratio decreases from almost one to zero as the temporal averaging interval increases from 2 min to 30 min. Overall these observations show that the combined effect of random measurement error and microscale special variability of rainfall characterised by nugget-to-sill ratio decreases with increasing (a) rainfall intensity class and (b) averaging interval.

Regarding the behaviour of the nugget-to-sill ratio against averaging interval, it is expected that with the averaging interval the (microscale) spatial correlation of rainfall would increase, which partly explains the observed pattern. The increase in spatial correlation of rainfall intensity with increasing temporal averaging interval agrees with other similar studies (e.g. Ciach and Krajewski, 2006; Fiener and Auerswald, 2009; Krajewski et al., 2003; Peleg et al., 2013; Villarini et al., 2008). For example, Krajewski et al. (2003) observed in their study on analysis of spatial correlation structure of small-scale rainfall in central Oklahoma a similar behaviour using correlogram functions for different temporal averaging intervals. But commenting on the decreasing trend of the nugget-to-sill ratio against intensity class, it cannot be attributed to improvement in microscale spatial correlation as it is neither natural nor proven. In fact, in Fig. 3.7 the behaviour of spatial correlation against rainfall intensity class does not show a distinctive trend except at the origin, i.e. the nugget effect. The absence of any consistent trend of spatial variability against intensity class was also observed in Ciach and Krajewski, (2006). Meanwhile this decreasing trend of nugget-to-sill ratio against rainfall intensity corresponds well with measurement errors of tipping bucket type rain gauges caused by its sampling mechanism (hereafter referred as TB error). This is due to the rain gauges' inability to capture small temporal variability of the rainfall time series. The behaviour of TB error against rainfall intensity as seen from Fig. 3.7 complements results from previous studies (Habib *et al.*, 2001; Villarini *et al.*, 2008). These studies also show that the TB error decreases with temporal averaging interval. Habib et al. (2001) found similar behaviour of TB error with increasing intensity (0-100 mm/h) and also with increasing averaging interval (1 min, 5 min and 15 min). Although the bucket size used in their study (0.254 mm) is slightly different from our rain gauge bucket size of 0.2 mm, the characteristic of the TB error against rainfall intensity for different averaging interval is consistent in both cases. In summary, the

behaviour of nugget-to-sill ratio of the variograms against temporal averaging interval can be explained by the combined effect of microscale spatial variability of rainfall and TB error, while the behaviour of nugget-to-sill ratio against intensity range can mainly be attributed to the latter.

In addition to the nugget-to-sill ratio, another parameter that characterises the variograms is the range, i.e. the distance up to which there is spatial correlation. At lower temporal averaging intervals ($\leq$ 5 min) the variograms for all rainfall intensity classes reach the variogram range very quickly (< 100 m). But at averaging intervals $\geq$ 15 min, the range has not been reached even at a maximum separation distance, showing the improvement in spatial correlation. High spatial variability of rainfall at shorter temporal averaging interval ($\leq$ 5 min) is an important observation in the context of urban drainage runoff modelling, as the time step used in such models is generally around 2 min for small catchments.

The fact that the data set covers only 10 months of data from two years with varying climatology is something that need to be acknowledged. However, for previous studies using such a dense network the duration of data collection is similar (e.g.: 15 months - Ciach and Krajewski, 2006; 16 months - Jaffrain and Berne, 2012). These time periods are reflection of the practical and funding issues to maintain such dense networks operating accurately for extended periods. The characteristics of our data are comparable with (Ciach and Krajewski, 2006; Fiener and Auerswald, 2009) as these studies also used rainfall data from warm months to investigate the spatial correlation structure. Despite the fact that the data cover only 10 months all derived variogram models are stable and reliable. Webster and Oliver (2007) suggested around 100 samples to reliably estimate a variogram model. Even in the case of 30 min temporal averaging interval and > 10 mm/hr (where we had the fewest number of observations) we had a total of 196 spatial lags to calculate the variogram. Furthermore, we demonstrated that all derived variogram models are stable and reliable by examining sub-sets of the data. We randomly selected 80% of the data from each intensity class and reproduced the variograms to compare them with the variograms presented in Fig. 3.7. We had to limit the subclass percentage to 80% to give enough time instants to reproduce variograms for all subclasses. We repeated this procedure a few times. Comparing these variograms with Fig. 3.7 shows that these variograms are very similar. One set of the variograms computed from 80% of the data are presented in Appendix

3A. This analysis supports our claim that the variograms shown in Fig. 3.7 are stable and an adequate representation of the rainfall spatial variation for each intensity class and temporal averaging interval.

One of the assumptions we made during the pooling procedure is that the spatial variability is reasonably consistent within a pooled intensity class. We acknowledge that with narrower intervals the assumption of consistency in spatial variability would be more realistic. But with the available data we had to find a compromise with the number of time instants. We believe that using three intensity subclasses is a reasonable compromise. Further we also introduced step 2 (section 3.2) that standardises the rainfall for each time instant within a subset. Although variograms are derived only for the whole subset, step 2 (before geostatistical upscaling) and step 9 (after geostatistical upscaling) ensure that the probabilistic model is adjusted for each time instant separately. Effectively, we assume the same correlogram for time instants of the same subclass, not the same variogram. Although this does not justify the assumption of similar spatial correlation structure within the pooled classes, it at least relaxes the assumption of the same variogram within subclasses. To compare the behaviour of variogram models for a narrower intensity interval, we produced variograms for narrower intensity classes ranging from 0 to 14 mm/hr for the 5 min averaging interval. The highest intensity class is limited to ≥ 12 - <14 mm/hr as for further narrower ranges (i.e ≥ 14 - <16 mm/hr and so on) there are not enough sample points to produce a meaningful variogram. Narrower intensity classes means that the assumption of similar spatial variability within a pooled subset is more realistic. Comparing Fig. 3.8 with Fig. 3.7, we conclude that the variograms shown in Fig. 3.7 are accurate representations of the average spatial variability conditions for corresponding intensity classes.
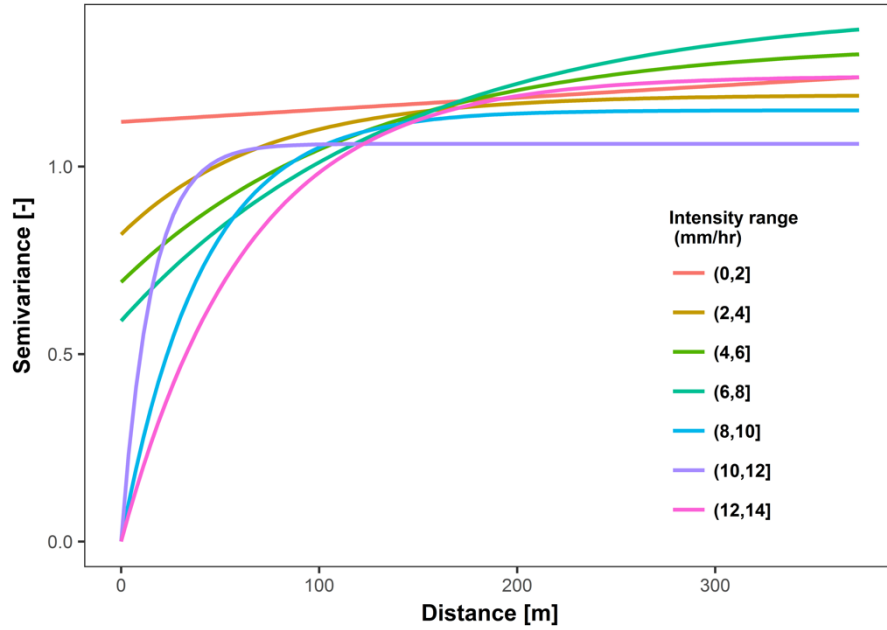
Figure 3.8: Calculated variograms for a narrower range of intensity at 5 min averaging interval.

3.4.2 Geostatistical upscaling of rainfall data

Having calculated all variograms, the next step is to apply spatial stochastic simulation for the time instants of interest followed by steps 6 to 9 in Fig. 3.4 to calculate the AARI together with associated uncertainty. This procedure was carried out for all events presented in Table 3.1. The following sections present and discuss the predicted AARI and associated uncertainty levels derived from step 9.

3.4.2.1 Prediction error vs AARI

The scatter plot in Fig. 3.9 shows the coefficient of variation of the prediction error (CV, refer Eq. (3.6)) plotted against predicted AARI at 5 min averaging interval for all time instants of all events presented in Table 3.1.

$$CV = \frac{AARI\ prediction\ error\ standard\ deviation}{Predicted\ AARI} \times 100\% \qquad (3.6)$$
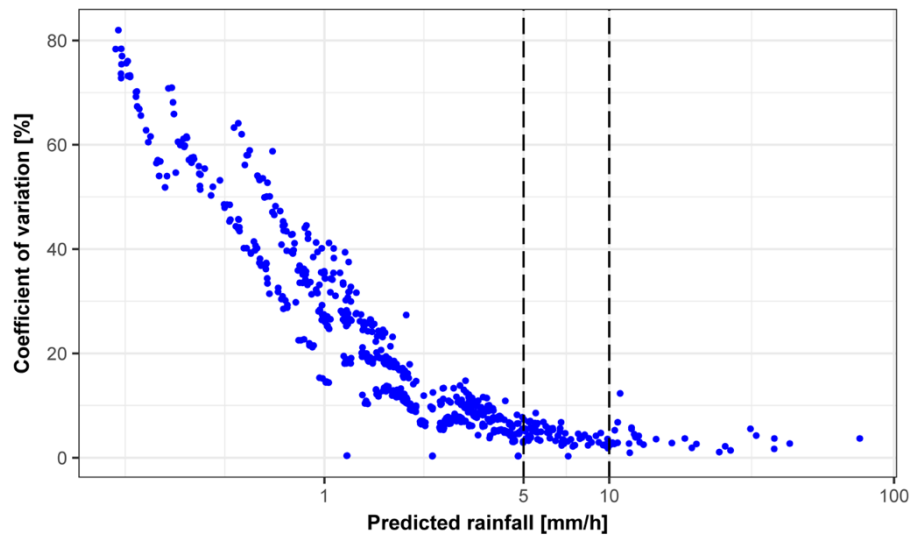
Figure 3.9: AARI prediction error CV (%) values against predicted AARI for averaging interval of 5 min.

The uncertainty level in the prediction of AARI represented by the CV is due to the combined effect of both spatial variability of rainfall and TB error in the rainfall data. It can be seen here that there is a clear trend of decreasing CV with increasing AARI. The CV values are as high as 80% when the AARI is smaller than 1 mm/hr and they get reduced to less than 10% when AARI is larger than 10 mm/hr. In a previous study by Pedersen et al. (2010) using rainfall measurements from similar tipping bucket type rain gauges, they also found that the uncertainty in prediction of mean rainfall depth decreases with increasing mean rainfall depth, but due to the limited information in their results they could not analyse this observation in detail. But here it is clear that this observation corresponds well with what we already observed in variograms in Fig. 3.7. These variograms show higher nugget-to-sill ratio at lower intensity due to high TB error consequently causing higher uncertainty in the prediction of AARI. At intensity class 0-5 mm/hr the nugget-to-sill ratio was almost one (nugget variogram) and as a result the derived CV values are significantly higher than other two intensity classes. It is interesting to note that, in the range of 1- 10 mm/hr, there are few points that are separated from the larger cluster with almost zero CV. It shows a consistent rainfall measurement over the area at these time instants, which results in a very small CV in the predicted AARI.

The above discussion is based on results from 5 min temporal averaging interval. The following section discusses the effect of temporal averaging interval on prediction

56

error. Further, although CV in Fig. 3.9 gets as high as 80%, the corresponding AARI is less than 1 mm/hr, thus the prediction error has a very less significance in urban hydrology. Hence we also analysed the prediction error associated with rainfall events' peaks in the last section.

3.4.2.2 Prediction error vs temporal averaging interval

Having analysed the behaviour of the prediction error CV against predicted AARI, this section presents the effect of temporal averaging interval on the prediction error of AARI. Figure 3.10 shows the kriging predictions with 95 % prediction intervals derived from the prediction standard deviation for temporal averaging intervals of 2 min, 5 min, 15 min and 30 min for event 11. Event 11 has average conditions in terms of event duration and peak intensity. Prediction errors of other events against the temporal averaging interval follow the same pattern of behaviour.
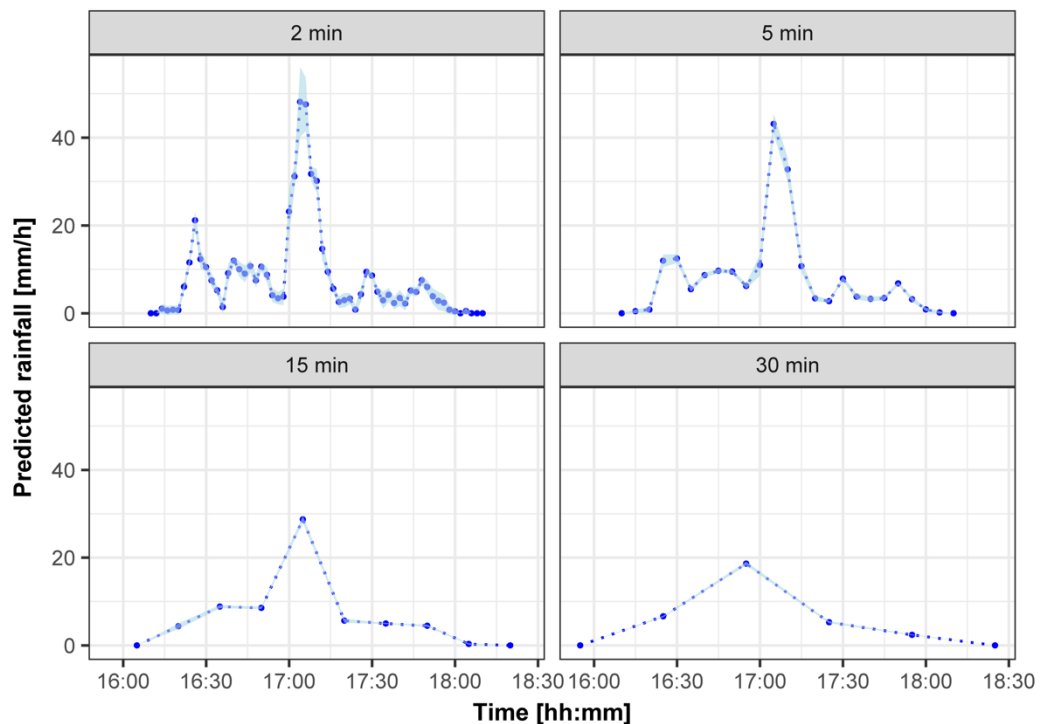
Figure 3.10:   Predictions of AARI (indicated by points) together with 95 % prediction intervals (indicated by grey ribbon) for rainfall event 11 for different averaging intervals.

While short time intervals are of greater interest in urban hydrology, they also lead to large uncertainties. Figure 3.10 shows the smaller the temporal averaging interval, the larger the prediction interval and the larger the level of uncertainty. This is due to the

combined effect of higher spatial variability and larger TB error at lower temporal averaging interval as seen from Fig. 3.7. When the averaging interval is larger than 15 min the prediction interval width becomes negligible. But temporal scales of interest in urban hydrology of similar sized catchment can be as low as 2 min where there is still considerable uncertainty. The 95 % prediction interval shows around ± 13 % of error in rainfall intensity corresponding to a prediction of peak rainfall of 47 mm/h at 2 min averaging interval. While temporal aggregation decreases uncertainty, it obviously leads to a significant reduction of the predicted peaks of AARI. For example, the peak of event 11 gets reduced to around 20 mm/h from around 50 mm/h when averaging interval increases from 2 min to 30 min. Hence a careful trade-off between temporal resolution and accuracy in rainfall prediction is needed to decide the most appropriate time step for averaging point rainfall data for urban hydrologic applications.

The decreasing trend of uncertainty in the prediction of AARI with increasing temporal averaging interval agrees with a previous study by Villarini et al. (2008). Although the spatial extent of their study is much larger (360 km^2), their results also show that the spatial sampling uncertainties tend to decrease with increasing temporal averaging interval due to improvement in measurement accuracy and improved spatial correlation.

3.4.2.3 Prediction error Vs peak rainfall intensity

In addition to rainfall event durations, rainfall event peaks are also of significant interest in urban hydrology as most of the hydraulic structures in urban drainage systems are designed based on peak discharge which is often derived from peak rainfall. Hence it is important to consider the uncertainty in prediction of peaks of AARI. Figure 3.11 presents predicted peaks of AARI for all 13 events presented in Table 3.1, together with labels indicating corresponding CV (%) values. The peak intensities range from 6 mm/h to 92 mm/h at 2 min averaging interval and this range narrows down to 3 mm/h – 21 mm/h at averaging interval of 30 min as a result of temporal aggregation. As expected, temporal aggregation from 2 min to 30 min also results in the reduction of CV. The highest CV at 2 min averaging intervals is 13% for event 4 and reduces to 1.7% at 30 min averaging interval. But it can also be noted that events 5, 6, 8 and 11 show their highest CV at 5 min averaging interval and not at 2 min averaging interval. Tracking back these events, they indeed show more spatial variation over 5 min period compared to 2 min period around the peak.
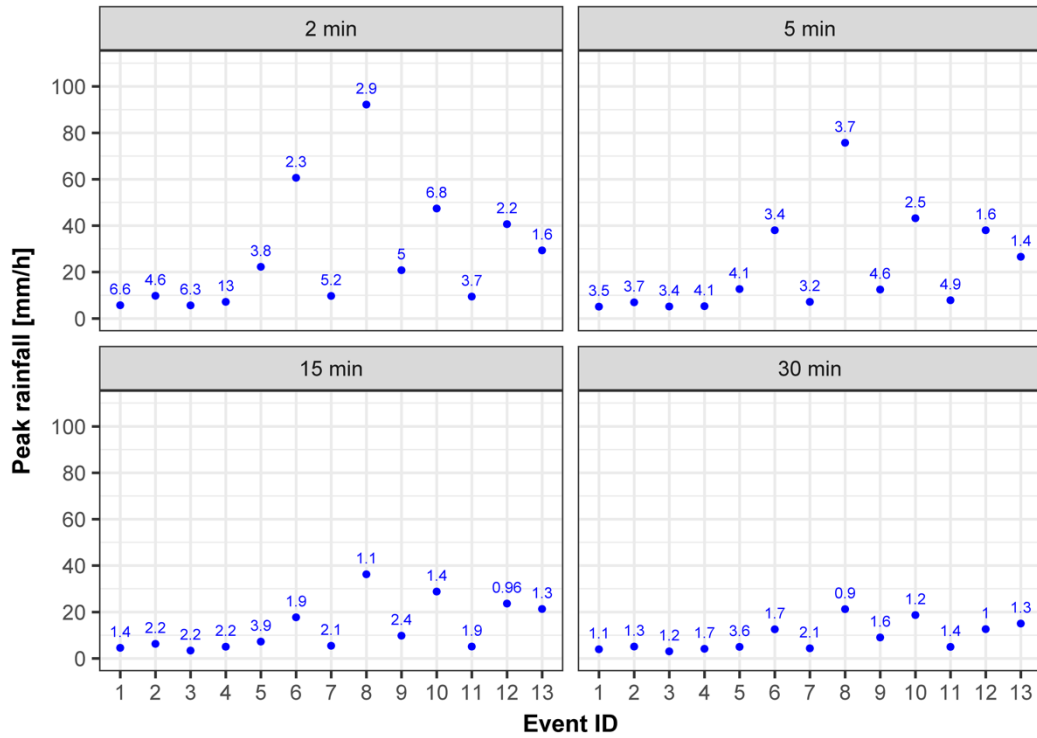
Figure 3.11: Predictions of event peaks of AARI (indicated by points) together with labels indicating corresponding CV (%) values.

As discussed in section 4.2.1, CV decreases with increasing predicted rainfall peaks and this effect is dominant when the averaging interval is at the lowest, i.e. 2 min. This is when the TB error is at its highest. When the temporal averaging interval is 30 min where the TB error is at its lowest, the difference between CV for lower (< 10 mm/h) and higher (> 10 mm/h) intensity becomes smaller. At 30 min averaging interval the mean CV below and above 10 mm/h are 1.7 % and 1.2 % respectively, but they increase to 6.6 % and 3.5 % at 2 min averaging interval. The maximum CV at 2 min averaging interval are 13 % and 6.8 % for lower (< 10 mm/h) and higher (> 10 mm/h) rainfall intensity respectively. Even though these values are significantly less than what we observed from Fig. 3.9 when the rainfall intensity is less than 1 mm/hr, they are still high considering the required accuracy defined in standard guidelines of urban hydrological modelling practice. For example, the current urban drainage verification guideline (WaPUG, 2012) in the UK defines a maximum allowable deviation of 25 % to -15 % in peak runoff demanding more accurate prediction of rainfall which is the main driver of the runoff process in urban areas. A 13% uncertainty in rainfall will result in a similar level of uncertainty in runoff prediction for a completely impervious

surface according to the well-established rational formula (Viessman and Lewis, 1995) which is still widely used for estimating design discharge in small urban catchments.

3.5 Conclusions

Geostatistical methods have been used to analyse the spatial correlation structure of rainfall at various spatial scales, but its application to estimate the level of uncertainty in rainfall upscaling has not been fully explored mainly due to its inherent complexity and demanding data requirements. In this study we presented a method to overcome these challenges and predict AARI together with associated uncertainty using geostatistical upscaling. We used a spatial stochastic simulation approach to address the combination of change of support (from point to catchment) and non-normality of rainfall observations for prediction of AARI and the associated uncertainty. We addressed the issue of scarcity in measurement points by using repetitive rainfall measurements (pooling) to increase the number of spatial samples used for variogram estimation. The methods were illustrated with rainfall data collected from a cluster of eight paired rain gauges in a 400×200 m^2 urban catchment in Bradford, UK. The spatial lag ranges from 21 m to 399 m. As far we are aware these are the smallest lag ranges in which spatial variability in rainfall is examined in an urban area using point rainfall measurements. We defined intensity classes and derived different geostatistical models (variograms) for each intensity class separately. We also used different temporal averaging intervals, ranging from 2 to 30 min, which are of interest in urban hydrology. To the best of our knowledge this is the first such attempt to assign geostatistical models for a combination of intensity class and temporal averaging interval. Finally, we quantified the level of uncertainty in the prediction of AARI for these different combinations of temporal averaging intervals and rainfall intensity ranges.

A summary of the significant findings are listed below:

- Several studies (e.g. Berne et al., 2004; Gebremichael and Krajewski, 2004; Krajewski et al., 2003) used a single geostatistical model in the form of variogram/correlogram for the entire range of rainfall intensity. The current study shows that for small time and space scales the use of a single geostatistical model based on a single variogram is not appropriate and a distinction between

rainfall intensity classes and length of temporal averaging intervals should be made.

- The level of uncertainty in the prediction of AARI using point measurement data essentially comes from two sources; spatial variability of the rainfall and measurement error. The significance and characteristics of the measurement error observed here mainly corresponds to sampling related error of tipping bucket type rain gauges (TB error) and may vary for other types of rain gauges.

- TB error decreases with increasing rainfall intensity. As a result of that, the prediction error decreases with increasing AARI. At 5 min averaging interval the CV values are as high as 80% when the AARI is smaller than 1 mm/hr and they get reduced to less than 10% when AARI is larger than 10 mm/hr

- At smaller temporal averaging intervals, the effect of both spatial variability and TB error is high, resulting in higher uncertainty levels in the prediction of AARI. With increasing temporal averaging interval the uncertainty becomes smaller as the spatial correlation increases and the TB error reduces. At 2 min temporal averaging interval the average CV in the prediction of peak AARI is 6.6 % and the maximum CV is 13 % and they are reduced to 1.5 % and 3.6 % respectively at 30 min averaging interval.

- TB error at averaging intervals of less than 5 min, especially at low intensity rainfall measurements, is as significant as spatial variability. Hence proper attention to TB error should be given in any application of these measurements, especially in urban hydrology where averaging intervals are often as small as 2 min.

Although the spatial stochastic simulation method used in this study needs more computational power (a summary on computation power is presented in Appendix 3B) than block kriging, it is a robust approach and allows data transformation during spatial interpolation and aggregation. Such data transformation is important because rainfall data are not normally distributed for small temporal averaging intervals. The pooling procedure used in this study helps provide a solution to meet the data requirements for geostatistical methods as it extends the available information for variogram estimation. Commenting on the minimum number of measurement points needed to employ this method is difficult, because like any other geostatistical interpolation method, the efficiency of this method also heavily depends on reliable estimation of the

geostatistical model (variogram). Hence, it basically comes down to the question of whether or not a given measurement network can produce a meaningful variogram. As mentioned, Webster and Oliver (2007) advised that around 100 measurement points are needed to adequately estimate a geostatistical model. But there is no single universal rule to define the minimum number of bins and the number of samples for each bin to produce a reliable variogram. Further, since pooling sample variograms of repeated measurements would produce a multiplication of spatial lags, the size of the available data set would also play a role in deciding the minimum number of measurement points.

An urban catchment of this size needs rainfall data at a temporal and spatial resolution which is higher than the resolution of most commonly available radar data (1000 m, 5 min). In addition the level of uncertainty in radar measurements would be much higher than that of point measurements, especially at a small averaging interval (< 5 min, Seo and Krajewski, 2010; Villarini et al., 2008), which are often of interest in urban hydrology. Hence, experimental rain gauge data similar to the ones used in this study are crucial for similar studies focused on small urban catchments.

Results from this study can be used for uncertainty analyses of hydrologic and hydrodynamic modelling of similar sized urban catchments in similar climates as it provides information on uncertainty associated with rainfall estimation which is arguably the most important input in these models. This information will help to differentiate input uncertainty from total uncertainty thereby helping to understand other sources of uncertainty due to model parameter and model structure. This estimate of the relative importance of uncertainty sources can help to avoid false calibration and force fitting of model parameters (Vrugt et al., 2008). This study can also help to judge optimal temporal averaging interval for rainfall estimation of hydrologic and hydrodynamic modelling especially for small urban catchments.

Data availability

"The rainfall intensity data used in this study are freely available at https://doi.org/10.5281/zenodo.291372"

Acknowledgements

Appendix
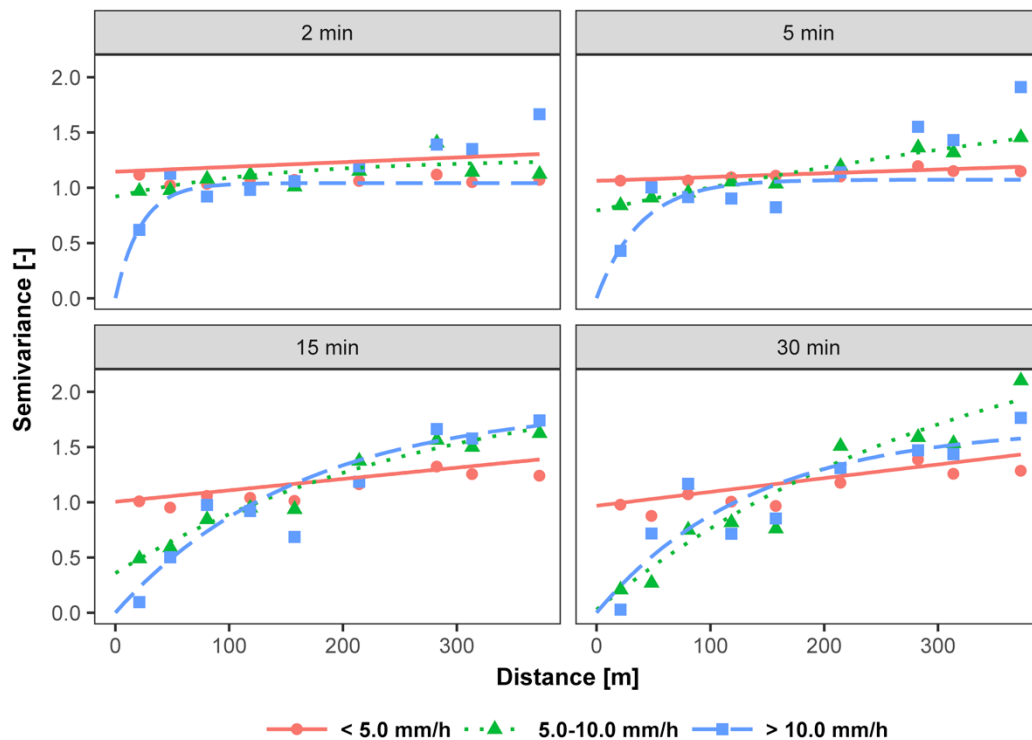
3A: Calculated variograms for each intensity class within each temporal averaging interval using randomly selected 80% of the data

3B: Summary of computational power required for spatial stochastic simulation.

Computer used	Area (m²)	Grid size (m²)	number of simulations	Computational time per time instant (s)
Core i5, 1.7	200 × 400*	25 × 25*	500*	10*
GHz, 4	200 × 400	25 × 25	1000	20
processers , 8	200 × 400	10 × 10	500	60
GB RAM				

*This configuration is used in this study.

4. Improving understanding of the underlying physical process of sediment wash-off from urban road surfaces

Manoranjan Muthusamy[1], Simon Tait[1], Alma Schellart[1], Md Nazmul Azim Beg[2], Rita F. Carvalho[2], João L.M.P de Lima[2]

[1] Department of Civil and Structural Engineering, University of Sheffield, Sheffield, S1 3JD, UK
[2] Marine and Environmental Sciences Centre, Department of Civil Engineering, University of Coimbra, Coimbra, 3030–788, Portugal

Correspondence to: Manoranjan Muthusamy (m.muthusamy@sheffield.ac.uk)

Abstract

Among the urban aquatic pollutants, the most common is sediment which also acts as a transport medium for many contaminants. Hence there is an increasing interest in being able to better predict the sediment washoff from urban surfaces. The exponential wash-off model is the most widely used method to predict the sediment wash-off. Although a number of studies proposed various modifications to the original exponential wash-off equation, these studies mostly looked into one parameter in isolation. This parameter is often the rainfall intensity thereby ignoring the interactions of other parameters corresponding to catchment surface and sediment characteristics. Hence in this study we aim (a) to investigate the effect of rainfall intensity, surface slope and initial load on wash-off load in an integrated and systematic way and (b) to subsequently improve the exponential wash-off equation focusing on the effect of the aforementioned parameters. A series of laboratory experiments were carried out in a full scale setup, comprising of a rainfall simulator, a 1 m^2 bituminous road surface, and a continuous wash-off measuring system. Five rainfall intensities ranging from 33 - 155 mm/hr, four slopes ranging from 2 - 16 % and three initial loads ranging from 50 - 200 g/m^2 were selected based on values obtained from literature. Fine sediment with a size range of 300 – 600 µm was used for all of the tests. Each test was carried out for one hour with at least 9 wash-off samples per test collected. Mass balance checks were carried out for all the tests as a quality control measure to make sure that there is no significant loss of sand during the tests. Results show that the washed off sediment load at any given time is proportional to initial load for a given combination of rainfall intensity and surface slope. This indicates the importance of dedicated modelling of build-up so as to subsequently predict wash-off load. Further, results also show that the wash-off fraction increases with both rainfall intensity and surface slope. The negative-

inverse-exponential (NIE) trend due to the effect of a first flush is more pronounced at combinations of catchment slopes steeper than 8% and rainfall intensities higher than 75 mm/hr. It was also observed that the maximum fraction that is washed off from the surface increases with both rainfall intensity and the surface slope. This observation leads to the second part of the study where the existing wash-off model is modified by introducing a capacity factor which defines this maximum fraction. This capacity factor is derived as a function of wash-off coefficient making use of the correlation between the maximum fraction and the wash-off rate. Values of the modified wash-off coefficient are presented for all combinations of rainfall intensities and surface slopes, which can be transferred to other urban catchments with similar conditions.

4.1 Introduction

Pollutant wash-off is the process by which non-point source pollutants including sediment, nutrients, bacteria, oil, metals and chemicals are removed from urban surfaces by the action of rainfall and runoff. Among the transported pollutants, the most common is sediment which plays a major role in water quality issues of inland water bodies in urban areas (Guy, 1970; Collins and Ridgeway, 1980; Chiew and Vaze, 2004). Sediment also contributes to urban floods by filling up drainage systems and reducing the hydraulic capacity of these systems that are designed to rapidly carry water away from roads and properties (Ivan, 2001). Hence, accurate modelling of sediment wash-off is important for water-quality-based decision-making. But modelling sediment wash-off is not a straightforward exercise as it often involves empirically calibrated equations containing parameters with a highly variable nature such as rainfall, catchment surface and particle characteristics. There are two main processes involved in the transport of sediment from an impervious surface: Build-up and Wash-off. Build-up is a process in which sediment accumulates during dry weather. Wash-off is the process where accumulated sediment deposition is removed from impervious surfaces by rainfall and runoff and then incorporated in the run-off flow over the surface (Francey *et al.*, 2011). Modelling of pollutant wash-off ranges from simple EMC (Event Mean Concentration) (Charbeneau and Barrett, 1998; Kayhanian *et al.*, 2007) to more sophisticated BUWO (Build-up Wash-off models).

One of the earliest studies on sediment wash-off was carried out by Sartor and Boyd (1972). They derived separate build-up and wash-off functions based on an

experimental study of runoff pollution in eight US cities. The original exponential wash-off equation proposed by Sartor and Boyd (1972) is given below

$$w_t = w_0\left(1 - e^{-kit}\right) \qquad (4.1)$$

where w_t is transported sediment load after time t, w_0 is initial load of the sediment on the surface; i is rainfall intensity; and k is the wash-off coefficient. This equation is widely used in several models with or without modifications. These modifications are mainly focused on k. It has been shown that k needs to be calibrated for each catchment as it depends on many parameters corresponding to surface characteristics (Sonnen, 1980; Nakamura, 1984), rainfall and runoff characteristics (Ammon, 1979; Sonnen, 1980; Nakamura, 1984) and particle size (Ammon, 1979; Sonnen, 1980). Apart from refinement in the estimation of k, some studies also suggest other forms of modifications. For instance, a power term to i was suggested to be able to predict the increase in concentration that corresponds to an increase in rainfall rate during an event (Huber and Dickinson, 1992). Another major modification suggested by Egodawatta et al. (2007) is the inclusion of a multiplicative capacity factor on the right side of the Eq.(4.1) varies with rainfall intensity for a better modelling of sediment removal. However, most of the above-mentioned refinements are very site specific and not easily transposed or generalised. Also most of these studies paid attention to one single parameter in isolation, thereby ignoring the effect and interactions of other parameters. For instance, although the introduction of a capacity factor by Egodawatta et al. (2007), is shown to be a meaningful modification, has only been investigated against rainfall intensity. An integrated approach which is lacking in these studies is necessary to investigate the combined effect of dominant parameters associated with rainfall characteristics, surface characteristics and sediment characteristics. Another interesting observation is the lack of attention given to the surface slope in the above studies. Two processes that drive sediment mobilisation are impact energy from rainfall drops (Coleman, 1993) and shear stress from overland flow (Akan, 1987; Deletic *et al.*, 1997) both of which are sensitive to surface slope especially the latter. With the exception of Nakamura, (1984) none of the above studies paid attention to the effect of slope. Nakamura (1984) results show that k increases with surface slope, but this study was based only on two randomly selected slopes and was not extensive enough to be used in subsequent studies or in practical applications.

In addition to the calibration of parameter k, another important input to the exponential washoff equation is the initial load w_0. Sartor and Boyd (1972) provided an exponential equation to calculate the build-up load, which is essentially the initial sediment load in the wash-off prediction. They modelled sediment build-up against antecedent dry days. Although this approach of modelling build-up mainly using antecedent dry days has been used in some models (Bertrand-Krajewski *et al.*, 1993), it has also been criticised, especially in recent studies (Charbeneau and Barrett, 1998; He *et al.*, 2010; Shaw *et al.*, 2010). Among these studies, Shaw et al. (2010) provided an overview of a number of studies which indicated that the mass of washed-off particulate matter during a storm event is relatively insensitive to the time between storm events. This was confirmed by He et al. (2010) who studied the quality of storm-water runoff from a semi-arid, urban residential catchment in Calgary, Alberta. They could not find any relationship between the event mean values of total suspended solids and the antecedent dry weather period. Despite these criticisms, the effect of build-up on wash-off has not been explored in depth in any of the above studies. Hence the question of whether there is a need to model build-up remains unanswered.

Considering the above gaps and room for improvements in sediment wash-off modelling we designed and carried out a series of laboratory experiments to:

- Study the effect of three dominant parameters corresponding to rainfall, surface and sediment characteristics in an integrated and systematic way. These parameters are, rainfall intensity (i), surface slope and initial load (w_0) respectively. , and
- Improve Eq. (4.1) using the experimental results focusing on the effect of the above three parameters.

4.2 Methodology

4.2.1 Experimental set up

Experiments were conducted in a full scale laboratory setup, described in Fig. 4.1, comprising of a rainfall simulator (used in, for example, Carvalho et al. 2014; de Lima et al. 2013; Isidoro and Lima 2013; Montenegro et al. 2013), a 1 m^2 bituminous road surface and a continuous wash-off measuring system. Steady artificially simulated rainfall was employed in order to eliminate the dependency on naturally occurring

rainfall. This approach provides better control over influential variables such as rainfall intensity and duration. Consequently, the use of simulated rainfall enables the generation of a large volume of data in a relatively short period of time (Herngren *et al.*, 2005).

.
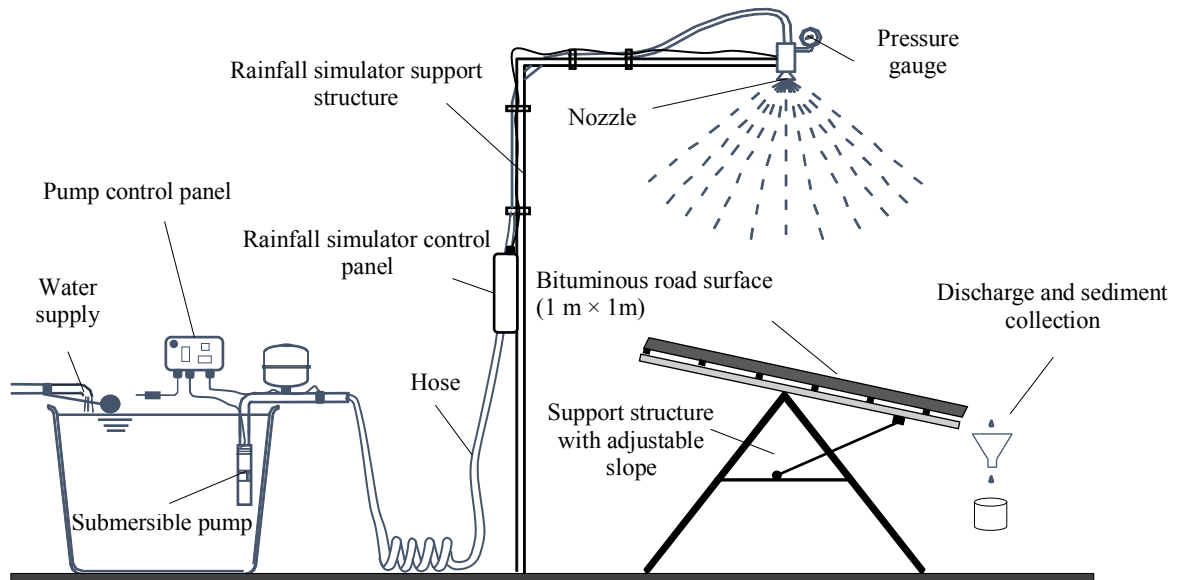
Figure 4.1: Sketch of the experimental setup
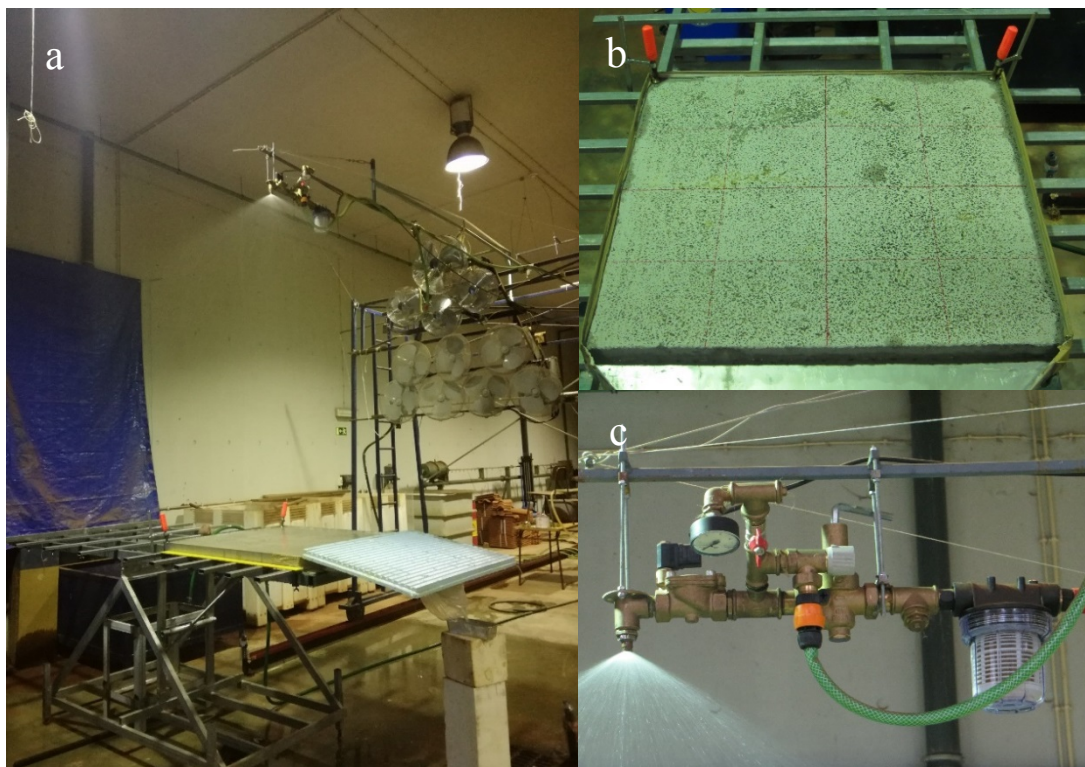
Figure 4.2: Photographs of (a) Experimental set up during data collection (b) Bituminous road surface with grids and (c) Nozzle with pressure gauge during the experiment

A typical urban road surface of 1 m^2 was prepared for the experiments by using bituminous asphalt concrete (Fig. 4.2b). The surface was tested for texture and impermeability before the experiments. Surface texture was measured using sand patch tests (Highway Department UK, 1989) on 16 equally divided grids. The mean texture depth index is 0.4 mm with a standard deviation of 0.03 mm. This surface texture is an average representation of wide ranges of impervious urban surfaces where the mean texture depth index varies from ~0 (tiled pavements) to ~1.0 mm (road surfaces). Mass balance of surface runoff was carried out to check the impermeability and the results show that the surface is completely impermeable. This surface was fixed on a metal support structure with adjustable slope as shown in Fig. 4.1.

The rainfall simulator (Fig. 4.1) has a pressurised hydraulic system comprised of: (i) a steady downward oriented full-cone nozzle (1/4-HH-14W FullJet from Spraying Systems Co., USA), with 3.58 mm orifice diameter, positioned 2.2 m above the geometric centre of the surface; (ii) a hydraulic system attached just in front of the nozzle to eliminate pressure fluctuations (more details in Isidoro & Lima, (2013); and (iii) a submerged pump (76.2 mm SQ from Grundfos Holding A/S, Denmark), installed in a constant head reservoir supplied with tap water. This system allows a steady operating pressure at the nozzle to produce rainfall with consistent intensity, with a spray angle of 120° (wide angle). The pressure at the nozzle is adjusted to change the rainfall intensity. D_{10} and D_{90} of the sand used in the experiment are 300 μm and 600 μm respectively. It is a washed, dried and accurately graded sand, free from organics, clay, silt or metallic inclusions and has a sub-angular to semi-rounded shape. This sediment size is selected to provide a well characterised sediment as this reduces the potential for size sorting and so allows us to link the wash-off behaviour to a particular sediment size. Further, the D_{50} is around 450 μm which is similar to the mean D_{50} of road sediments reported in e.g. Butler and Clark (1995) and Bertrand-Krajewski et al. (1993).

The effect of three parameters: rainfall intensity, surface slope and initial sediment load on sediment wash-off were tested. Five intensities ranging from 33-155 mm/hr, four slopes ranging from 2-16 % and three initial loads ranging from 50 - 200 g/m^2 were selected. These upper limits cover the extreme values derived from literature. For example, the highest ever recorded one hour (note that all simulations were carried out

for one hour, Table 4.1) rainfall intensity in UK is 92 mm/hr (MetOffice UK, 2016). Further the UK Department of Transport suggests a maximum gradient of 10% for most types of the road other than in exceptional circumstances (Manual for Streets, 2009). Finally, the average of `ultimate` sediment loads found in 8 selected urban sites located in Lambeth, UK is 172 g/m^2 (Butler and Clark, 1995). The lower limits were selected using trial simulations to be able to produce a measurable amount of wash-off. Sampling times are adjusted based on the corresponding intensities and at least nine samples were collected for each simulation, see Table 4.1. Note that for the 2% slope the wash-off load was found to be less than 2% of initial load even for the highest intensity of 155 mm/hr; hence only simulations with an initial load of 200 g/m^2 were carried out for this slope. All wash-off samples were collected using numbered foil containers and then these foil containers were dried using standard laboratory moisture extraction ovens until they are completely dry. All dried samples were then weighed using a high precision (accuracy of 0.1 g) laboratory measuring scale.

4.2.2 Quality control

The bituminous road surface was sub divided into 16 equal grid squares (Fig. 4.2b) to aid distribution of the sediment uniformly over the surface. Initially trial tests were repeated with the same conditions (rainfall intensity, surface slope and initial load) to confirm that the experimental setup gave consistent results. Comparing results from these repeated tests showed that the difference was within ±2%. At the end of both the trial and the actual tests the remaining sand from the surface was collected by washing off the surface to carry out a mass balance check. In all cases the mass loss was found to be less than 2% of the original sediment load ensuring that there is no significant loss of sand during the tests.

Table 4.1: **Summary of experimental conditions and sampling frequency**

Slope (%)	Initial load (g/m^2)	Sampling times (min)
2%	200	5, 10, 17, 25, 31, 38, 45, 52, 60
4%	50,100,200	[for intensities 33 mm/hr and 47 mm/hr]
8%	50,100,200	2, 5, 8, 13, 19, 25, 31, 38, 45, 52, 60
16%	50,100,200	[for intensities 75 mm/hr, 110 mm/hr and 155 mm/hr]

4.3 Results and Discussion

4.3.1 Experimental results

To compare the results from different initial loads on a common scale, we used a normalised measure, the wash-off fraction (F_w) which is a ratio between transported sediment load after time t (w_t) and initial load of the sediment (w_0) (Eq. 4.2). Figure 4.3 shows the wash-off fraction plotted against the duration for all of the tests summarised in Table 4.1.

$$F_w = \frac{w_t}{w_0} \qquad (4.2)$$

The most interesting observation is the effect of initial load on F_w. Initial load does not affect F_w until the slope gets steeper (8% and 16%). Even in the case of 8% slope, initial load has an effect only when the rainfall intensity is higher than 110 mm/hr. In these cases, there is an increasing pattern of values of F_w with increasing initial load. These combinations of high rainfall intensity and steep slope where the initial load has an impact on F_w are very rare in reality (MetOffice UK 2017; Manual for Streets 2009). It implies that the effect of initial load on F_w is negligible for most general combinations of rainfall intensity and surface slope. This essentially means the actual mass of sediment washed off at any given time (w_t) is proportional to initial load for a given rainfall intensity and surface slope. Hence the prediction of build-up is perhaps the most preferred way to subsequently predict wash-off compared to the methods presented in recent studies (e.g. Shaw et al. 2010). But on the other hand, as Shaw et al. (2010) correctly pointed out, it is a challenging task to model the build-up process due to unpredicted occurrences of activities like construction work or the input of vegetative debris from wind storms. Despite these challenges the strong correlation observed between build-up load and wash-off load indicates the importance of modelling the build-up process. This observation does not necessarily invalidate the criticisms on the build-up model of Sartor and Boyd (1972) by Charbeneau and Barrett (1998), Shaw et al. (2010) and He et al. (2010) as their criticism is mainly on the use of antecedent dry days as the main parameter controlling the build-up process. Rather this finding calls for more attention to be paid on modelling of build-up process taking more parameters (Wijesiri et al. 2015; Morgan et al. 2017) into consideration in addition to antecedent dry days.

Looking at the effect of intensity and slope, for a given intensity, F_w increases with increasing slope regardless of initial load. Similarly, for a given slope, F_w increases with increasing intensity regardless of the initial load. At 2% slope the wash-off load is negligible for all the rainfall intensities with a maximum F_w of 0.018 at the highest rainfall intensity of 155 mm/hr. The highest F_w after one hour is ~0.9 for the extreme case where intensity, slope and initial load are 155 mm/hr, 16% and 200 g/m^2 respectively.

Another important observation from Fig. 4.3, especially at steeper slopes (8% and 16%), is that only a certain fraction of the available sediment is mobilised during a simulated rain event before the curve becomes almost flat and this maximum fraction increases with rainfall intensity and surface slope. This behaviour suggests a rainfall event for a given surface slope has the capacity to mobilise only a fraction of sediment from the road surface and once it reaches that capacity, as observed during the experiments, wash-off becomes almost zero even though a significant fraction of the original sediment is still available on the surface. Although at milder slopes (2% and 4 %) the wash-off fraction has not reached its maximum value within the duration of the test, it would have reached this value if the tests were long enough. This trend was also observed in a similar study by Egodawatta et al. (2007) in which they analysed this maximum fraction against rainfall intensity. Hence there are two parameters which characterise these curves; wash-off rate and maximum fraction both of which increase with increasing slope and increasing intensity. The negative inverse exponential pattern (NIE) of these curves can mainly be attributed to the first flush effect. The concept of first flush is that the initial period of storm flow carries most of the pollutant including sediments from the urban surface (Helsel *et al.*, 1979; Sansalone and Buchberger, 1997; Bertrand-Krajewski *et al.*, 1998). The most common value of k - 4.6 in^{-1} (0.18 mm^{-1}) in Eq. (4.1) is basically derived from the concept that the first one-half inch of total run-off in one hour would wash-off 0.9 fraction of the initial load (Huber and Dickinson, 1992).

But such generalisation is not valid for all the conditions as can be seen in Fig. 4.3 where the strength of first flush changes with rainfall intensity and surface slope. We believe, in addition to rainfall intensity and surface slope, surface texture and sediment size also plays a major role in deciding the strength of the first flush. Fig. 4.3 shows the effect of first flush is negligible during smaller intensities and milder slopes. This can

mainly be attributed to the surface texture depth and/or sediment size used in the experiments. If it is a smoother surface typically associated with roofs, the effect of the first flush will possibly be magnified (Farreny *et al.*, 2011). Also with a smaller sediment size one would expect a more pronounced first flush and a higher wash-off fraction. For instance Egodawatta & Goonetilleke, (2008) in their study found that the most of the sediment that is washed off initially is the finer sand (<200 μm). But it can be noted that the range of sediment sizes used in Egodawatta and Goonetilleke (2008) is much wider (0-1000 μm) compared to the sediment size used in this study which also explains the higher wash-off fractions they observed in a similar experimental set up. Although we would expect an increase in wash-off fraction with smaller sediment size there is also a possibility for smaller sediment to get trapped in the pores of the surface due to the surface roughness and so significantly reduce their mobility. But on the other hand, if the sediment size is bigger than the surface texture depth, it might be too large for the rainfall impact to mobilise the particles and for the runoff process to transport it especially on mild surface slopes. When the surface is rougher similar to the one used in this study the interaction between sediment size and texture depth becomes complicated and it needs to be explored in depth.
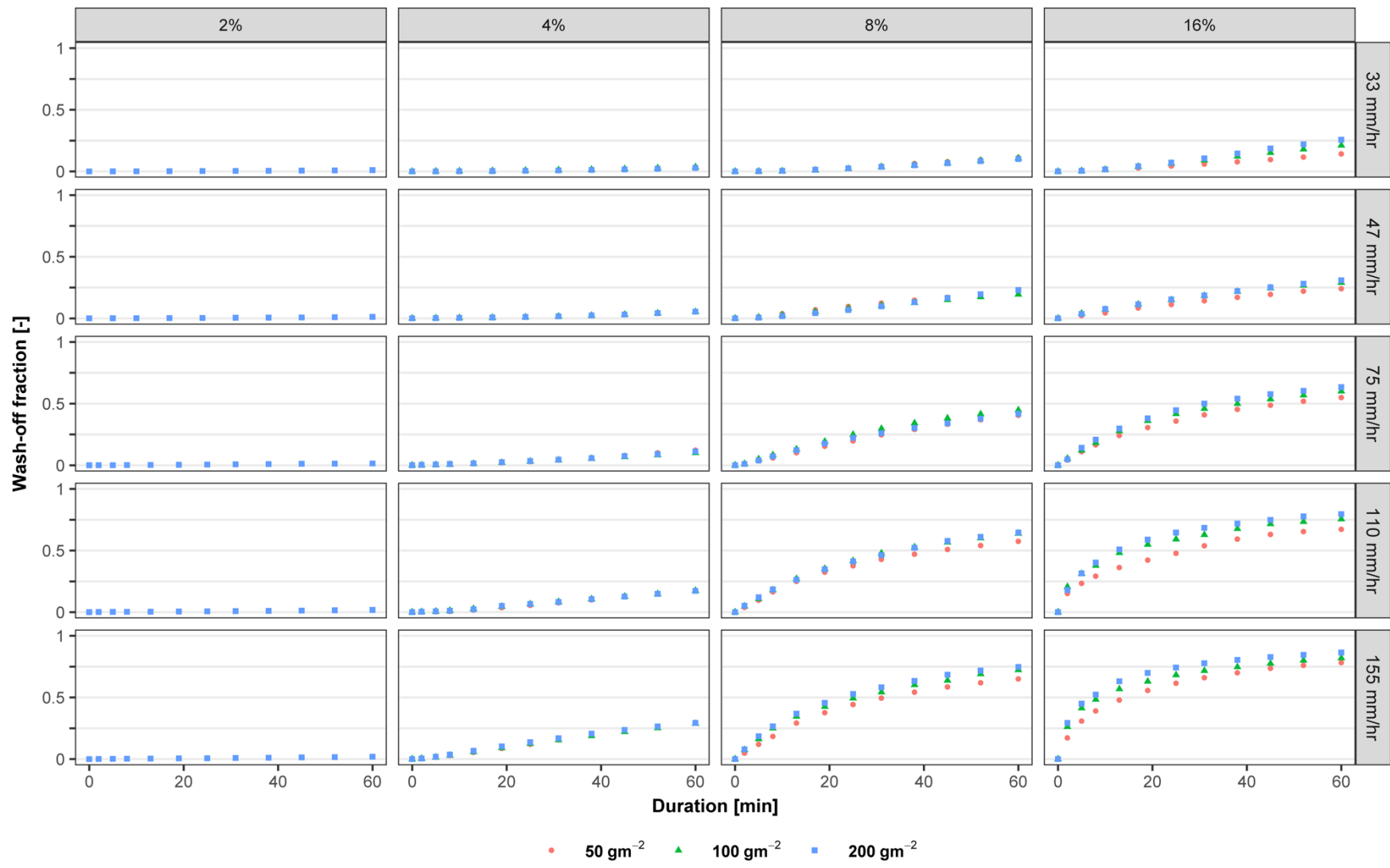
Figure 4.3: Wash-off fraction for all combinations of rainfall intensity, surface slope, and initial load

4.3.2 Model improvement

We attempt to modify Eq.(4.1) based on the experimental data discussed in Section 3.1. From Fig. 4.3 and the corresponding discussion, it is clear that the effect of initial load on the wash-off fraction is negligible for most cases. Hence the effect of initial load has not been considered in this section and a modification in Eq. (4.1) is proposed based only on experimental results from a constant initial load of 200 g/m^2.

As discussed in the previous section, only a certain fraction of the available sediment is mobilised during a simulated rain event before the curve becomes almost flat and this fraction increases with rainfall intensity and surface slope. To replicate this behaviour in the modelling of wash-off, Egodawatta et al. (2007) introduced a new parameter called the capacity factor (C_F), ranging from 0 to 1, into Eq. (4.1) as shown in Eq. (4.3).

$$\frac{w_t}{w_0} = C_F\left(1 - e^{-kit}\right) \qquad (4.3)$$

But due to the limitations of their study, they concluded that C_F primarily varies with rainfall intensity, disregarding the effect of other parameters such as slope. But from Fig. 4.3 it is clear that this fraction of sediment which a rainfall event has the capacity to wash-off also strongly depends on the surface slope in addition to rainfall intensity. This implies C_F needs to be adjusted based on surface slope too. Hence C_F which is the maximum fraction available and k which defines the wash-off rate both need to be calibrated for all combinations of rainfall intensities and surface slopes. From Fig. 4.3 it can also be noted that the higher the maximum fraction, the faster the F_w reaches the maximum fraction meaning these two parameters are dependent. Figure. 4.4 is a simplified version of the experimental results to illustrate this concept where the maximum wash-off fractions are indicated by F_{w1}, F_{w2}, and F_{w3} and the time taken to reach these fractions are indicated by t_1, t_2 and t_3 respectively. This figure shows that $F_{w1} < F_{w2} < F_{w3}$ and consequently $t_1 > t_2 > t_3$. Applying this concept in to Eq. (4.3) suggests that C_F and k are dependent. Therefore, it was decided to make C_F a function of k as shown in Eq.(4.4) instead of introducing a new C_F altogether as in Egodawatta et al. (2007). This gives more physical meaning to this empirical equation and also avoids the compensation of two independent parameters in order to over fit the experimental results. Such compensation between two independent parameters could lead to identifiability problems (Sorooshian and Gupta, 1983).

$$\frac{w_t}{w_0} = f(k)\left(1 - e^{-kit}\right) \qquad (4.4)$$
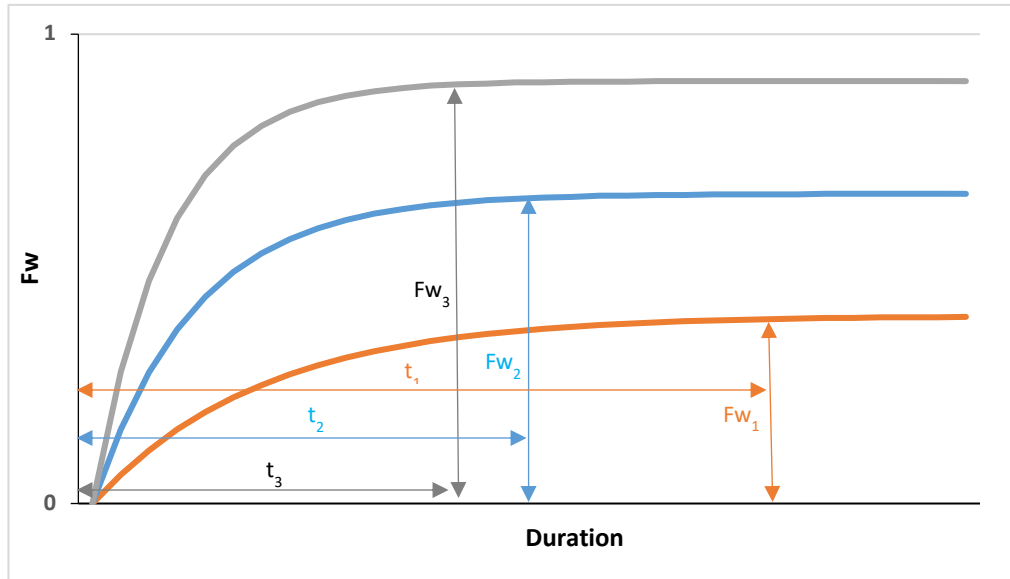
Figure 4.4: Variation of maximum wash-off fraction and corresponding duration

Having introduced a new C_F in the form of $f(k)$, the next step is to estimate this $f(k)$ and subsequently estimate the k values for each combination of slope and intensity. The following steps explains the procedure to estimate $f(k)$ and k values.

1. The first step is to find $f(k)$ which best fits the experimental results. To keep the new equation as simple as possible, $f(k)$ is assumed as a factor of k which leads to the following equation:

$$\frac{w_t}{w_0} = ck'\left(1 - e^{-k'it}\right) \qquad (4.5)$$

Where c is a constant with a unit of mm as unit of k' is mm^{-1}. Note that k is changed to k' since the new values for k' will be different from conventional k values.

2. The next step is to estimate the value of c (constant) and k' (varies with slope and intensity) which gives the smallest residual sum-of-squares between the fitted models and experimental results. Hence for a given value of c, the residual sum-of-squares are calculated for 20 fitted curves derived from 20 k' values each corresponding to a combination of a slope and an intensity. The objective function is to minimise the sum of all residual sum-of-squares derived from these 20 curves for different c values. There are two constraints. The first constraint is that both c and k' cannot have negative values and the second constraint is that the product of c and k' cannot exceed the maximum possible fraction which is 1.

Figure 4.5 shows the sum of residual sum-of-squares plotted against the range of c. It can be seen that the sum of residual sum-of-squares is at its minimum when c is 20. The corresponding fitted curves with different k' are shown in Fig. 4.6 for all the combinations of intensity and slope where the initial load is 200 g/m^2. The sum of the residual sum-of-squares for all these fitted curves is only 0.13 which shows the model fits well with the experimental results.

The k' values derived from the fitted models corresponding to a c value of 20 are plotted against intensity for each slope in Fig. 4.7 (a). Fig. 4.7 (b) shows the surface plot that is obtained by linearly interpolating k' values over the domain. From both plots, it can be noted that the rate of change in k' values against slope increases with increasing rainfall intensities. At 2% slope the change of k' against rainfall intensity is negligible due to the negligible difference in the wash-off fraction against rainfall intensity at this slope. At 8% and 16% slopes the rate of change in k' values after 110 mm/hr shows a drop. This is a reflection of the similar drop in the increase in the wash-off fraction as can be seen in Fig. 4.6. The k' values range from 2.6×10^{-3} to 4.2×10^{-2} which gives a range of 0.05 to 0.84 for C_F ($= 20 k'$). The highest C_F of 0.84 corresponds to the extreme case where intensity and slopes are 155 mm/hr and 16% respectively.

When transferring these c and k' values to other catchments other parameters has to be taken into account especially the sediment size and surface texture. Both the capacity factor ($c\times k'$) and wash-off rate (represented by k') would most likely increase with decreasing sediment size and/or decreasing surface texture depth. Nevertheless, the improved model as shown in Eq. (4.5) is expected to perform well for any sediment

size and surface texture as the underlying physical processes will be the same as those on which the equation was developed.
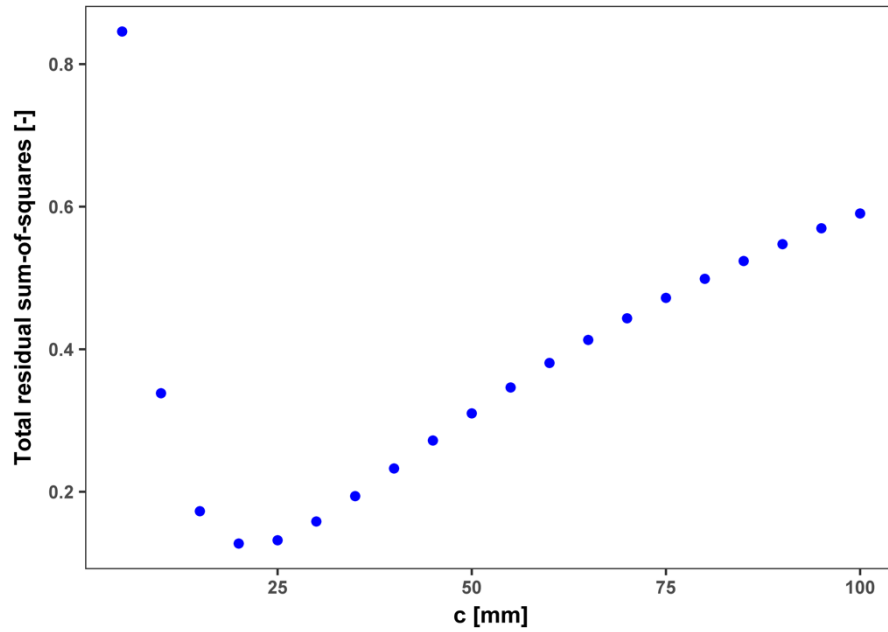
Figure 4.5: Total sum of residual-sum-of-squares plotted against *c* values ranging from 0 to 100, the dashed line shows the *c* value at which the total residual sum-of-squares is minimum
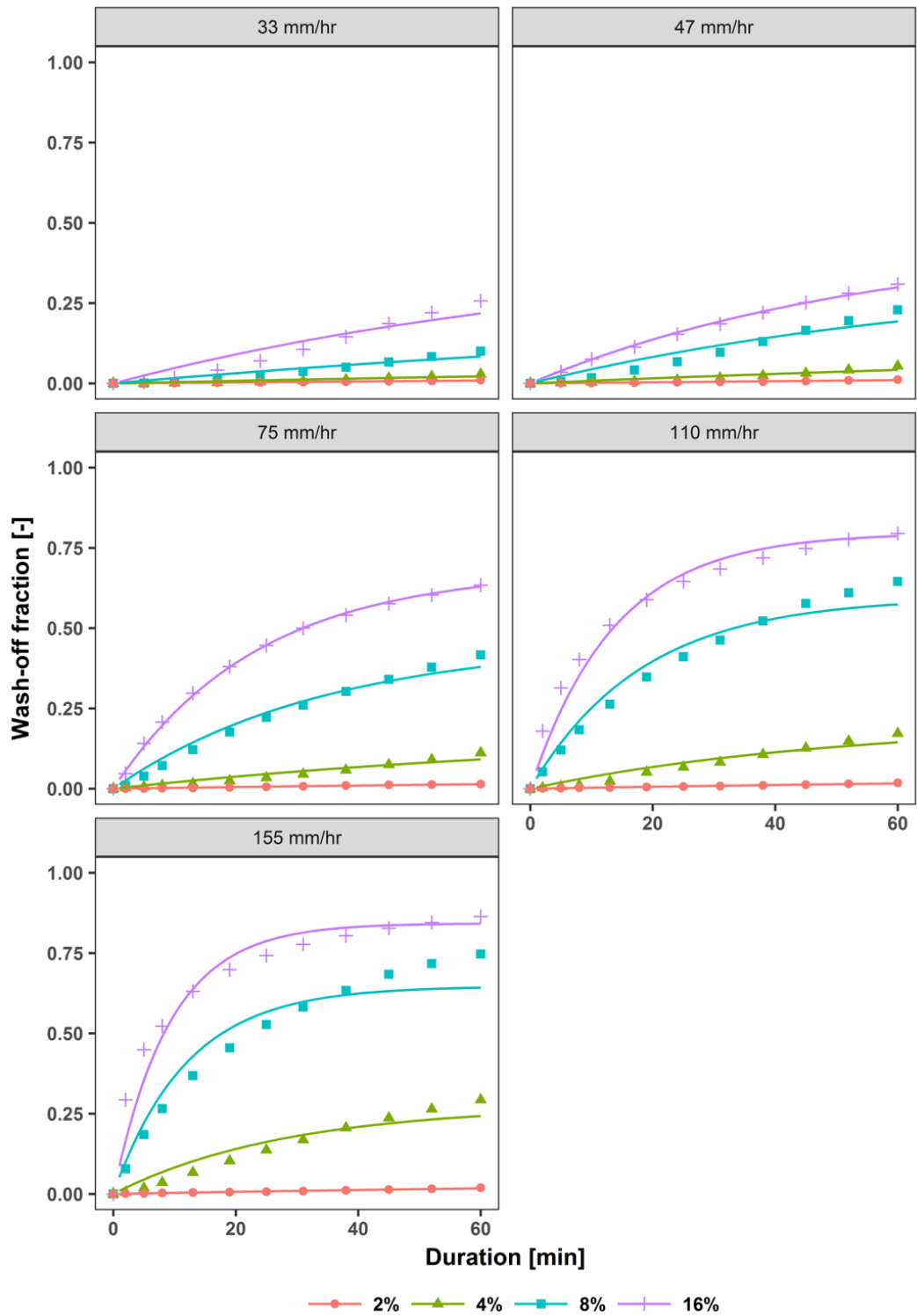
Figure 4.6: Measured wash-off fraction (points) and corresponding fitted curves (lines) derived from Eq.(4.5) (for *c* = 20 and *k'* values as shown in Fig 4.7.) for all combinations of rainfall intensity and surface slopes where initial load is 200 g/m^2
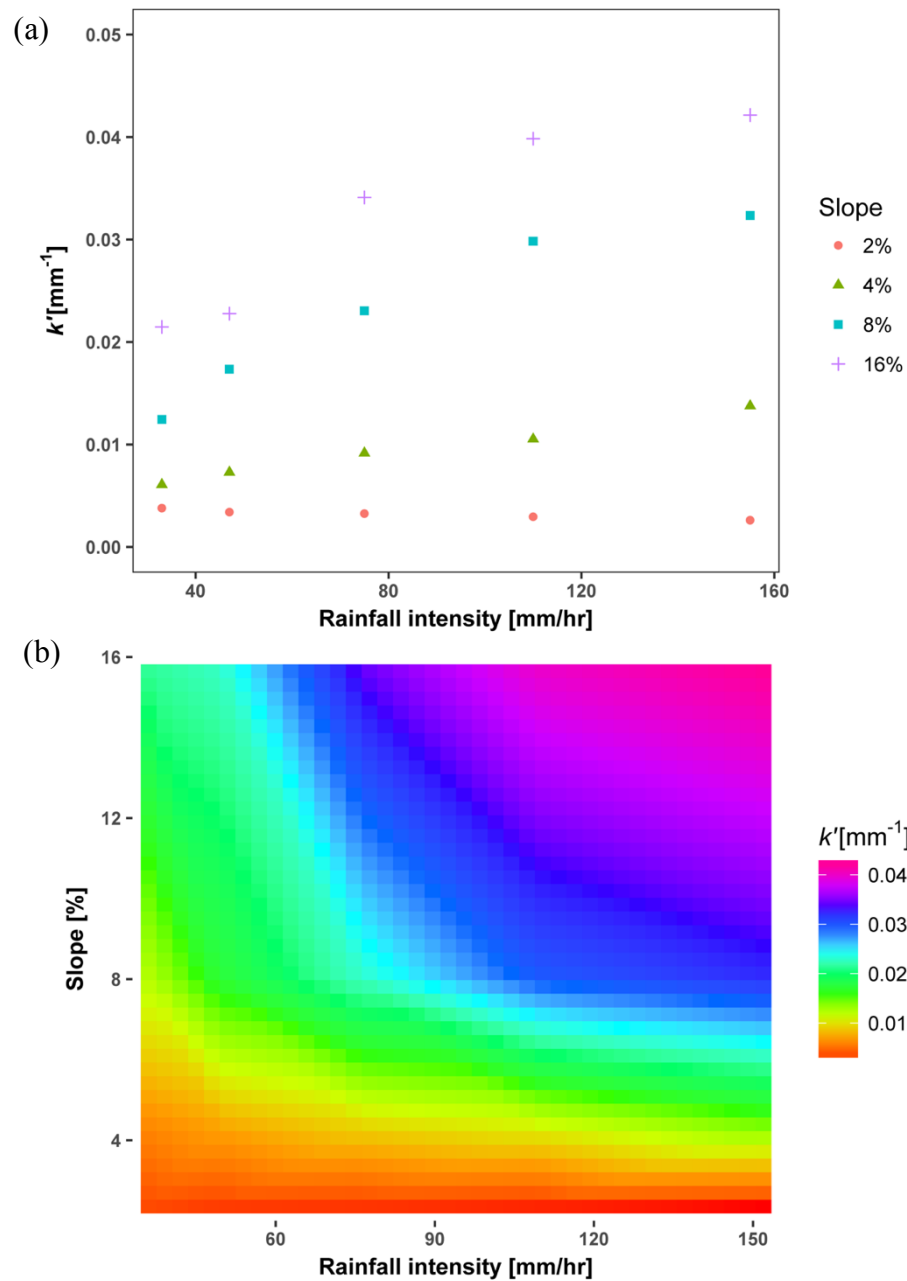
Figure 4.7: (a) Derived *k'* values for all the combinations of rainfall intensity and surface slope and (b) raster image of interpolated k' values over the domain

4.4 Conclusions

In this study, we investigated the effect of rainfall intensity, surface slope and initial load on sediment wash-off using an artificial rainfall generator and a typical urban road surface of 1 m². There has not been a previous experimental study which explored the effect of all the above three dominant parameters on wash-off in an integrated and systematic way.

The experimental results show that:

- The effect of initial load on wash-off fraction at any given time is negligible for most general combinations of rainfall intensity and surface slope. This essentially means that the washed off load at any given time is proportional to initial load for a given combination of a rainfall intensity and a surface slope. Hence, a dedicated modelling approach to predict build-up to help subsequently predict wash-off, despite the challenges mentioned in Shaw et al. (2010) should not be overlooked.

- The negative-inverse-exponential (NIE) trend due to the effect of first flush is clearly observed at combinations of catchment slopes steeper than 8% and rainfall intensities higher than 75 mm/hr. For combinations of milder slope and lower rainfall intensity the effect of first flush becomes negligible.

- A rainfall event has the capacity to mobilise only a fraction of sediment from the road surface and once it reaches that capacity, as observed during the experiments, wash-off becomes almost zero even though a significant fraction of sediment is still available on the surface. The maximum fraction that can be washed off from the surface increases with both rainfall intensity and the surface slope.

This final observation above led us to the second part of the study where the existing wash-off model is modified by introducing a capacity factor which defines the maximum fraction. This capacity factor is derived as a function of wash-off coefficient making use of the correlation between maximum fraction and the wash-off rate. This new improved equation is expected to perform better compared to the original equation as it models the underlying physical process better. Values for the wash-off coefficient are derived for combinations of rainfall intensity and slope which can be transferred to other urban catchments with similar conditions. In the future, in addition to the initial load, rainfall intensity and surface slope, it would be interesting to examine the effect of surface texture and sediment size on the wash-off process. This way a complete matrix of values for capacity factor and wash-off coefficient can be derived which can be transferred to any urban catchments.

4.5 Acknowledgements

5. Accounting for uncertainty propagation in enhanced sediment wash-off modelling within a Bayesian framework

Manoranjan Muthusamy[1], Omar Wani[2,3], Alma Schellart[1] and Simon Tait[1]

[1] Department of Civil and Structural Engineering, University of Sheffield, Sheffield, UK

[2] Institute of Environmental Engineering, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

[3] Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Dübendorf, Switzerland

Correspondence to: Manoranjan Muthusamy (m.muthusamy@sheffield.ac.uk)

Abstract

Exponential wash-off models are the most widely used method to predict sediment wash-off from urban surfaces. In spite of many studies, there is still a lack of knowledge on the effect of external drivers such as rainfall intensity and surface slope on the wash-off prediction and this consequently leads to the lack of knowledge on the effect of uncertainty in external drivers on the wash-off predictions. In this study, a more physically realistic "structure" is added to the original exponential wash-off model (OEM) by replacing the invariant parameters with functions of rainfall intensity and catchment surface slope, so that the model can better represent catchment and rainfall conditions without the need of lookup table and interpolation/extrapolation. In the proposed new exponential model (NEM), two such functions are introduced. One function describes the maximum fraction of the initial load that can be washed off by a rainfall event for a given slope and the other function describes the wash-off rate during a rainfall event for a given slope. The parameters of these functions are estimated using data collected from a series of laboratory experiments carried out using an artificial rainfall generator, a 1 m^2 bituminous road surface and a continuous wash-off measuring system. These experimental data contain high temporal resolution measurements of wash-off fractions for combinations of five rainfall intensities ranging from 33-155 mm/hr and three catchment slopes ranging from 2-8 %. Bayesian inference which allows the incorporation of prior knowledge is implemented to estimate parameter values. Explicitly accounting for model bias and measurement errors, a likelihood function representative of the wash-off process is formulated, and the uncertainty in the prediction of the NEM is quantified. Finally, the propagation of rainfall uncertainty through NEM is investigated in detail using sub-kilometre rainfall data with uncertainty due to spatial variability and measurement error. The results of this study show: 1) even

when the OEM is calibrated for every experimental condition, the NEM's performance, with a parameter values defined by functions, is comparable to the OEM. 2) Verification indices for estimates of uncertainty associated with the NEM suggest that the error model is able to capture the uncertainty well. 3) The level of uncertainty in the prediction of wash-off load due to rainfall uncertainty can be smaller, similar or higher to the level of uncertainty in rainfall depending on the intensity range and the "first-flush" effect.

5.1 Introduction

Urban surface sediment's ability to act as a transport medium to many contaminants makes it one of the major source of pollutants in an urban environment (Guy, 1970; Collins and Ridgeway, 1980; Mitchell *et al.*, 2001; Lawler *et al.*, 2006). Hence there is an increasing interest in being able to better predict the sediment washoff from urban surfaces. But, modelling sediment wash-off is not a straightforward exercise as it requires the understanding of complex interactions between external drivers with a highly variable nature such as rainfall, catchment surfaces and particle characteristics (Sartor and Boyd, 1972; Deletic *et al.*, 1997; Egodawatta and Goonetilleke, 2008). Currently, the most widely used wash-off models are originally developed using laboratory experiments and consequently include empirical parameters without clear physical interpretations. The exponential wash-off equation (Eq.(5.1)) proposed by Sartor and Boyd (1972) is one such model whose performance is highly dependent on the accurate estimation of parameter k:

$$w_t = w_0\left(1 - e^{-kit}\right) \qquad (5.1)$$

Where w_t is the total transported sediment load up to time t; w_0 is initial load of sediment on the catchment surface; i is rainfall intensity; and k is an empirical wash-off coefficient. Equation (5.1) has widely been used in several software packages (e.g. SWMM) with or without modifications (e.g. Zug et al. 1999; Huber and Dickinson 1992). Recently, Egodawatta et al. (2007) suggested a multiplication of a capacity factor for improved modelling of sediment removal which gives a more physical interpretation to the empirically calibrated original model shown in Eq.(5.1). According to Eq.(5.1), if the rainfall continues for long enough regardless of the rainfall intensity, it can wash off all the sediment available at the beginning of the event. In

other words, the maximum wash-off fraction ($^{w_t}/_{w_0}$) is always one. But Egodawatta et al. (2007) showed that a storm event has the capacity to wash-off only a fraction of sediments available and once this maximum fraction is reached the wash-off becomes almost zero, even though a significant fraction of sediment is still available on the surface. They suggested the introduction of an additional term referred to as the 'capacity factor (C_F) to replicate this finding in the model equation. With the inclusion of C_F Eq. (5.1) becomes

$$\frac{w_t}{w_0} = C_F(1 - e^{-kR_t}) \qquad (5.2)$$

Although the above modification was shown to be a meaningful refinement, C_F was investigated against rainfall intensity in isolation in Egodawatta et al. (2007). Muthusamy et al. (2018) further showed that C_F also varies with catchment surface slope in addition to rainfall intensity. In spite of the modifications suggested by various studies including Egodawatta et al. (2007) and Muthusamy et al. (2018), the calibration parameters k and the newly introduced C_F still need to be calibrated for the conditions of each catchment. In general, this is achieved by using a combination of look up tables/charts and interpolation/extrapolation of existing data.

Furthermore, none of the abovementioned studies include any information on the uncertainty in the estimation of the calibration parameters and their dependency structure which needs to be accounted in the prediction of wash-off using these parameters. Although adequate treatment of propagation of uncertainties in model prediction is a currently heavily researched area in hydrology, there are only a few studies on uncertainty related to wash-off modelling. In this regard, Dotto et al. (2012) compared a number of uncertainty techniques applied in urban water storm water quality modelling and found that a Bayesian approach, although computationally demanding, to be one of the preferable uncertainty assessment technique. A Bayesian approach helps to identity different sources of uncertainty such as parameter uncertainty, model bias and measurement noise and consequently helps to separately analyse them, though this requires knowledge about the error process (Dotto *et al.*, 2012). This estimate of the relative importance of uncertainty sources can help to avoid false calibration and forced fitting of model parameters (Vrugt *et al.*, 2008). Furthermore, it also helps to better understand the propagation of error in external drivers through the model. Among these external drivers, rainfall data is an essential

input in the prediction of sediment wash-off. But due to highly variable nature of rainfall over a wide range of scales, it is not always possible to measure rainfall at an appropriate temporal and spatial resolution which is required by specific hydrological modelling application including sediment wash-off modelling. This is common case in both rural and urban catchments. But the effect of inadequate representation of rainfall variability is magnified in urban areas which are characterised by smaller catchment size with a higher proportion of impervious area resulting in a smaller catchment reaction time and higher surface run-off (Al and Elson, 2005; Segond *et al.*, 2007; Gires *et al.*, 2012; Ochoa-Rodriguez *et al.*, 2015; Muthusamy *et al.*, 2015). For example, Ochoa-Rodriguez et al. (2015) showed that the error in peak discharge predictions due to inadequate representation of rainfall variability is up to 250% for drainage areas in the order of 1 ha and up to 50% for drainage areas of ~800 ha. Furthermore, Ochoa-Rodriguez et al. (2015) and Muthusamy et al. (2017) also showed that measurement and sampling error associated with rainfall is as significant as spatiotemporal variability of rainfall at temporal averaging interval of less than 5 min which is typically the timescale of interest in sediment wash-off modelling. Since the wash-off process is also mainly driven by rainfall and runoff (Sartor and Boyd, 1972; Deletic *et al.*, 1997; Egodawatta *et al.*, 2007), any uncertainty in rainfall would have a direct effect on sediment wash-off predictions as well. This demands more attention to be paid to the investigation of rainfall uncertainty propagation in wash-off prediction which is also an area that has not been explored in depth. This can be mainly attributed to the lack of knowledge of the variability of calibration parameters against rainfall intensity or volume.

Considering the above research gaps, in this study, first, the exponential wash-off model presented in Eq.(5.2) was further developed to add a more physically realistic structure by replacing the calibration parameters with functions of external drivers associated with catchment surface and rainfall characteristics. This will not only avoid the need of such empirical look up table/charts and interpolation/extrapolation of data, but it will also introduce some transparency in the parameter estimation which is otherwise a "black box" approach. Second, the uncertainty due to model bias, parameters and measurement noise was separately estimated as all of them need to be accounted in the prediction of wash-off. Further, replacing the invariant calibration parameters with functions of external drivers (i.e. rainfall intensity and surface slope) makes it easier to investigate the propagation of errors in the external drivers (e.g.

rainfall intensity) as these external drivers will now be explicitly defined in the new equation. Finally, by taking advantage of this feature, the propagation of rainfall uncertainty through the new improved model developed in this study was also investigated.

5.2 Material and Methods

5.2.1 Wash-off Data

The data used in this study was collected from a series of laboratory experiments carried out using an artificial rainfall generator, a 1 m^2 bituminous road surface and a continuous wash-off measuring system. Data were collected from experimental conditions with different combinations of rainfall intensity, catchment surface slope and initial sediment load. Five intensities ranging from 33-155 mm/hr, four slopes ranging from 2-16 % and three initial loads ranging from 50 - 200 g/m^2 were tested. More details on the experimental set up and data collection can be found in Muthusamy et al. (2018). As reported in Muthusamy et al. (2018) the effect of initial load on wash-off process was found to be negligible. Hence in this study, experimental results from a constant initial load of 200 g/m^2 as presented in Fig. 5.1 are used. This figure shows the variation of cumulative wash-off fraction ($F_w = {}^{w_t}/_{w_0}$) against rainfall intensity and surface slope.

Note that the 16% slope was eliminated from the data, given that such slopes on road surfaces are extreme scenario and exist only in rare locations. For example, the Department of Transport in UK suggests a maximum gradient of 10% for roads other than in exceptional circumstances (Manual for Streets, 2009). Since one of the aims of the study is to develop a single model with a fixed set of parameters, inclusion of results from such an extreme scenario in the calibration may affect the performance of the model for more general cases.
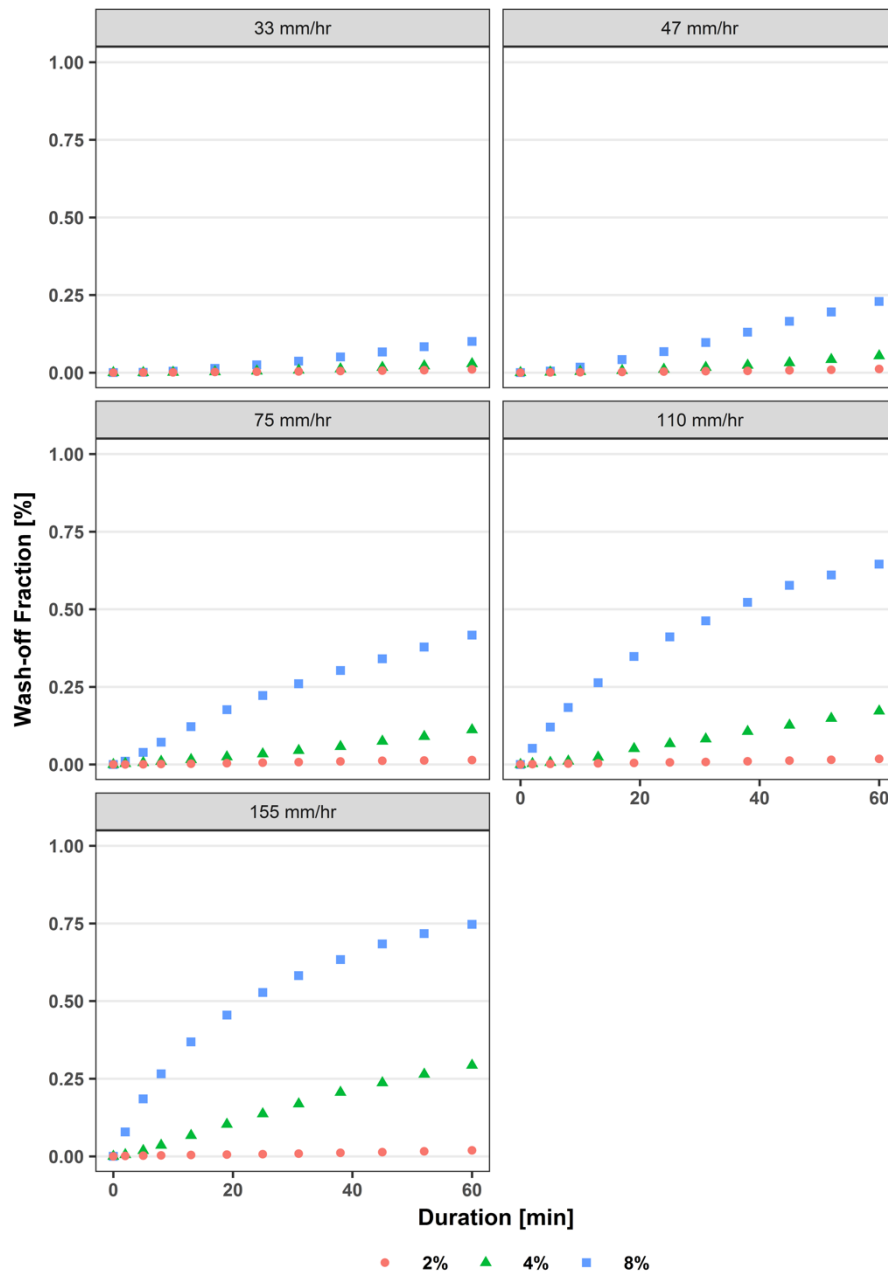
Figure 5.1: Selected results from Muthusamy et al. (2018): Variation of wash-off fraction for different combinations of rainfall intensity and surface slope

5.2.2 The modified wash-off model structure and its rationale

The main objective is to replace the calibration parameters in Eq. (5.2) with functions of surface slope and rainfall intensity, consequently adding a more physically realistic structure to the model. This should make the model robust to new combinations of rainfall intensity and surface slope. To do so the properties of the model that are sensitive to such parameters need to be identified and understood. From Eq. (5.2) there are two parameters which define the characteristics of a wash-off curve. The first parameter, C_F, defines the highest wash-off fraction for a given combination of rainfall

intensity and a slope. The second, k, defines how fast the wash-off curve reaches the maximum fraction for a given surface slope and rainfall intensity, and hence reflects the erosion rate from the catchment surface. Hence, C_F and k are proposed to be replaced with functions of surface slope and rainfall intensity, as shown in Eq. (5.3) and Eq. (5.4).

$$C_F = c_1 i_m{}^{c_2} s^{c_3} \qquad (5.3)$$

$$k = c_4 s \qquad (5.4)$$

Where $c_1, .., c_4$ are constants[2], i_m is the representative rainfall intensity of a rainfall event, s is the catchment surface slope. The following criteria were considered when defining Eq. (5.3) and Eq. (5.4), while also trying to keep these functions as simple as possible to reduce the number of constants:

- C_F – as explained before C_F is a capacity factor which defines the maximum fraction from the initially available sediment that can ever be washed off from a rainfall event for a given slope. Hence, C_F ranges from 0 to 1 and increases with both surface slope and (representative) rainfall intensity of the event. When either of the representative rainfall intensity or slope is zero C_F is zero.

- $k - k$ defines the wash-off rate and it also increases with rainfall intensity and surface slope. But it should be noted that R_t in the exponential term is cumulative rainfall depth at time t, i.e. $i_t t$ which is already a function of average rainfall intensity over time t, i_t. Hence k was taken as a (linear) function of slope only. The complete exponential term reads as $c_4 s i_t t$ which is function of both rainfall intensity and surface slope.

Hereafter this new exponential model will be referred as NEM and the original exponential model as shown in Eq. (5.1) will be referred as OEM.

5.2.3 Estimation of model parameters and associated uncertainty

Bayesian inference was used to estimate the parameter probability distribution, which allows prior knowledge on the parameters to be incorporated in the estimation and also formally quantifies uncertainty in the estimation. In addition, it also helps to capture

---

[2] Although $c_1, .., c_4$ are constant, in Bayesian inference they are referred to as model parameters to aid the readers follow the procedure easily.

the dependence structure between parameters. Bayesian inference requires the definition of the likelihood function and the prior distribution of the parameters.

5.2.3.1 The likelihood function

In addition to finding the best estimate of the parameters, we are also interested in the uncertainty associated with the parameter estimation and consequently the uncertainty in the prediction of wash-off fraction. One way of doing this is to include the error terms which represent the dominant sources of uncertainty explicitly in the likelihood function. We used an error model which accounts for errors due to model structural deficit (model bias, B_M) and measurement noise (E). B_M is modelled as an autoregressive stationary random process and E modelled as an independent identically distributed (IID) process. Hence, an observed output, Y can be formulated as

$$Y(x, \theta, \psi) = y(x, \theta) + B_M(x, \psi) + E(\psi) \qquad (5.5)$$

Where x is the external drivers, θ is deterministic model parameters, ψ error model parameters and y (x, θ) is deterministic model output. In this case, Y is observed wash-off fractions (F_w) and y is the deterministic model output predicted from NEM (f_w). x represents rainfall intensity and surface slope. θ is $sd.B$, $sd.E$ and l are error model parameters (ψ) in which $sd.B$ and l are standard deviation and correlation length respectively which characterise the autoregressive stationary random process and, $sd.E$ is the standard deviation of the measurement noise. Given the error description of Eq. (5.5), we define B_M as a multivariate Gaussian distribution with covariance matrix $\sum(x, \psi)$ and $E(\psi)$ as independent, identical normal noise. Therefore, the analytic formulation of the likelihood function with n number of observation can be formulated as

$$P(Y|x, \theta, \psi) = \frac{(2\pi)^{-\frac{n}{2}}}{\sqrt{\det(\sum(\psi, x))}} \exp\left(-\frac{1}{2}[Y - y(\theta, x)]^T \sum(\psi, x)^{-1}[Y\right.$$
$$\left. - y(\theta, x)]\right) \qquad (5.6)$$

An autoregressive error model represents model structural deficit better than IID as it accounts for the "memory" in the error time series (Del Giudice *et al.*, 2013). This autoregressive bias error model was originally suggested in other generic statistical applications (Craig *et al.*, 2001; Kennedy and O'Hagan, 2001; Higdon *et al.*, 2004;

Bayarri *et al.*, 2007) and later adapted for environmental engineering applications (Reichert and Schuwirth, 2012).

5.2.3.2 Prior distribution of parameters and constraints

Since the introduced parameters $c_1, .., c_4$ are all new, there is no previous estimation of the exact parameters, but a range for each parameters can be derived using our knowledge of the wash-off process, observational data, and the prior belief about values of C_F and k.

Values of c_4 were derived from previous estimations of k as c_4 equals to k/s. The list of k values derived from previous studies is given in Table. 5.1. From the table the range of $0 - 10$ were selected for k. In the absence of any information on slope in most of these studies same range for c_4 was used considering a minimum slope of 1%. Hence a uniform prior with the range 0-1000 was used as a prior distribution for c_4. A uniform prior distribution of model parameters is recommended when there is not enough evidence available to choose a different type of distribution (Freni and Mannina, 2010; Dotto *et al.*, 2012)

Table 5.1: k **values from literature**

Reference	Land use/catchment type	Value k (mm^{-1})
Alley (1981)	Urban catchment	0.036-0.43
Nakamura (1984)	Various	0.05-10
Huber and Dickinson (1992)	General	0.04-0.4
Millar (1999)	Residential	0.21
Egodawatta et al. (2007)	Concrete and asphalt roads	$5.6 \times 10^{-4} - 8.0 \times 10^{-4}$

As discussed previously the range of C_F is 0-1. This leads to the constraint $0 \leq c_1 i_m{}^{c_2} s^{c_3} \leq 1$. However, implication of this constraint in the definition of prior probability is not straightforward, hence this constraint was used in the estimation of likelihood probability.

It is challenging to define prior distributions for the error model parameters ($sd.B, sd.E$ and l) especially in the case of wash-off modelling as examples from such applications in literature are currently lacking. Out of the three parameters, some information on the measurement noise represented by $sd.E$ can be obtained by frequentist tests, i.e. repeating the experiments sufficiently large number of times. But

it is not always possible given the limitation in allocated resources and time. In the absence of much information on any of the error parameters, a uniform prior with the range from 0 to 1(= maximum wash-off fraction) was used for both $sd.B, sd.E$ and a uniform prior with the range of $0 - 200$ min was used for correlation length. This range is selected as error correlation is expected to be insignificant beyond such time length.

5.2.3.3 Bayesian inference

Once the prior distributions (the probability of deterministic and error model parameter, θ and ψ without considering the observed output, Y), $P(\theta, \psi),$ and the likelihood function (the probability of seeing the observed output, Y, as generated by a model with deterministic and error model parameter, θ and ψ), $P(Y|x, \theta, \psi)$, are defined, the posterior distribution of the deterministic and error model parameters (the conditional probability of θ and ψ once the observed output, Y has been taken into account) can be formulated as,

$$P(\theta, \psi|Y, x) = \frac{P(\theta, \psi)P(Y|x, \theta, \psi)}{\int P(Y|\theta, \psi)P(x, \theta, \psi)d\theta d\psi} \qquad (5.7)$$

Since the direct analytical calculation of $P(\theta, \psi|Y, x)$ is generally not possible, numerical techniques such as Markov Chain Monte Carlo (MCMC) simulations have to be applied. MCMC techniques generate a random walk through the parameter space which will converge to the posterior distribution. In this study we used robust adaptive Metropolis MCMC sampler presented in Vihola (2012) which is implemented in a R package, *adaptMCMC* (Scheidegger, 2017)

5.2.4 Propagation of rainfall uncertainty

Since rainfall is the main external driver of the wash-off process, rainfall uncertainty propagation through the NEM is investigated in this section. The rainfall data used in this study is obtained from Muthusamy et al. (2017). Muthusamy et al. (2017) presented areal rainfall intensities and associated uncertainty for 13 rainfall events estimated using geo-statistical upscaling of high resolution point measurements collected from 8 stations within an area of 200 × 400 m in Bradford, UK. The main sources of uncertainty in rainfall estimation was shown to be due to sub-kilometre spatial variability and measurement error. More detail on the estimation of areal rainfall intensity together with associated uncertainty can be found in Muthusamy et al. (2017). Further, they provided each event data for different temporal averaging interval ranging

from 2 min to 30 min. This information is used here to investigate the effect of temporal aggregation of rainfall, in addition to spatial aggregation, on the prediction of wash-off fraction. Table 5.2 provides summary of the rainfall events together with the coefficient of variation (CV) associated with the prediction of peak rainfall intensity of each events. All 13 events are presented in Appendix 5A.

Table 5.2: Summary of rainfall events and associated uncertainty presented in Manoranjan Muthusamy et al. (2017)

Event ID.	Network average duration (h)	Network average intensity (mm/h)	Network average rainfall (mm)	Predicted Event peak (mm/hr) / Uncertainty in event peak prediction (CV %)	
				2 min averaging interval	30 min averaging interval
1	6.33	2.20	13.9	5.7/6.6	3.9/1.1
2	6.42	2.55	16.3	9.9/4.6	5.1/1.3
3	8.92	1.79	16.0	5.6/6.3	3.0/1.2
4	6.83	1.99	13.6	7.7/13	4.1/1.7
5	11.4	2.39	27.3	22/3.8	4.8/3.6
6	4.42	5.31	23.4	61/2.3	13/1.7
7	3.25	3.23	10.5	9.8/5.2	4.3/2.1
8	1.50	7.84	11.8	92/2.9	21/0.9
9	3.08	3.35	10.3	21/5.0	9.1/1.6
10	2.00	7.96	15.9	48/6.8	19/1.2
11	7.92	2.14	17.0	9.5/3.7	4.9/1.4
12	1.75	6.51	11.4	41/2.2	1.2/1.0
13	8.17	4.34	35.5	29/1.6	15/1.3

5.3 Results and discussion

5.3.1 Model performance and associated uncertainty

Figure 5.2 shows the model output with the optimal values for $c_1, .., c_4$ (Table. 5.3) with maximum posterior probability density, i.e. the most probable values given the prior and observed data. This model performance is compared against the performance of the OEM described in Eq. (5.1). Experimental data with 2% and 8% slopes are used for calibration and the data from the 4% slope used for verification. The k value of the OEM is calibrated for each and every combination of surface slope and rainfall intensity during the calibration stage and during the validation stage k values are derived by using interpolation. It can be seen from Fig. 5.2 that with calibration data,

NEM with fixed values of parameters $c_1, .., c_4$ performs as well as the OEM which was calibrated for each and every combination of surface slope and rainfall intensity separately. From Table 5.4, it can be seen that the difference in sum of root mean square error (RMSE$_{OEM}$ - RMSE$_{NEM}$) from the ten calibrated set of data is -0.07 (Wash-off fraction). However, the robustness of NEM over OEM can be seen during the verification stage where the NEM performs better than the OEM in several cases. The difference in sum of root mean square error (RMSE$_{OEM}$ - RMSE$_{NEM}$) from 5 sets of data during verification stage is 0.09 (Wash-off fraction). The drawback with OEM is that for a set of new catchment conditions where OEM has not been calibrated before k value needs to be calculated using interpolation/extrapolation. This might lead to the underperformance of OEM during validation stage as shown in the Fig. 5.2. Considering the overall performance, the NEM with only 4 parameters ($c_1, .., c_4$) performs better than OEM with 15 parameters ($k_1, ..., k_{15}$). Hence, the NEM does not only avoid the need of interpolation to predict the calibration parameter values, it also performs as well as the calibrated OEM.

Table 5.3: Optimal values of constants of Eq. (5.3) and Eq. (5.4)[3]

$c_1 [(mmhr^{-1})^{-0.672}]$	$c_2 [-]$	$c_3 [-]$	$c_4 [mm^{-1}]$
3.99	0.672	1.99	0.208

Table 5.4: Performance of OEM and NEM

Model	Sum of root mean square error (RMSE)	
	Calibration	Verification
OEM	0.11	0.20
NEM	0.18	0.11

---

[3] Note that the unit of c_1 depends on the optimal value of c_2
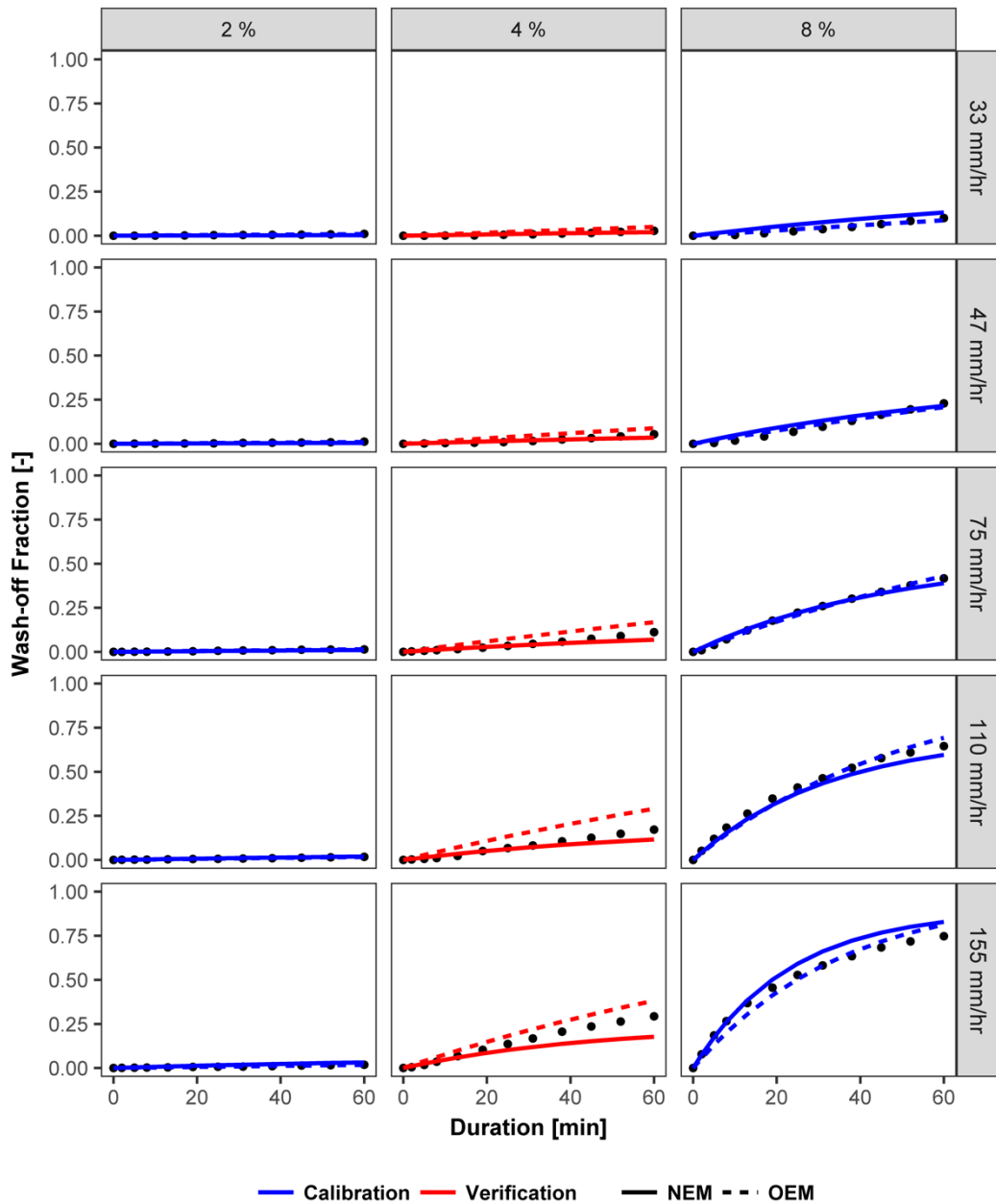
Figure 5.2: Comparison of the model performance

5.3.1.1 Parameter distribution and correlation

This section discusses the posterior distribution of parameters and their multivariate behaviour. Figure 5.3 shows posterior distributions and bivariate matrix of the deterministic and error model parameters. The most likely value of $sd.B$ and $sd.E$ are 0.02 (2%) and 0.002 (0.2%) respectively, showing that most of the uncertainty in the wash-off estimation can be explained by the model bias and that uncertainty due to measurement noise is negligible. Although these are approximate representations of the actual system and corresponding uncertainty, we believe that the experiments were conducted with as high a quality as possible. This is one of the reason why a road

surface as small as 1 sq.m was selected as it gives a better control over the experimental set-up. For example the smaller surface area keeps the spatial variability of the rainfall to the minimum. Furthermore, it also keeps the sediment loss during the experiment to insignificant. The maximum sediment loss observed during an experiment was less than 2% which is an indication of the good quality control.

Looking at the bivariate plots, there is a strong positive correlation between parameters c_1 and c_3 which indicates that these two parameters compensate each other in order to maximise the posterior probability. This can also be seen between parameters c_2 and c_3, but to a lesser extent. Similarly, the strong positive correlation between $sd.B$ and l means that these parameters compensate each other in order to fit the autoregressive error model B_M. Bayesian inference helps resolve such identifiability issues by allowing for informative priors. Therefore, for real cases, where we have reasons to believe that one of the two parameters should be more constrained, the other parameter value will automatically come out to be constrained after joint inference.
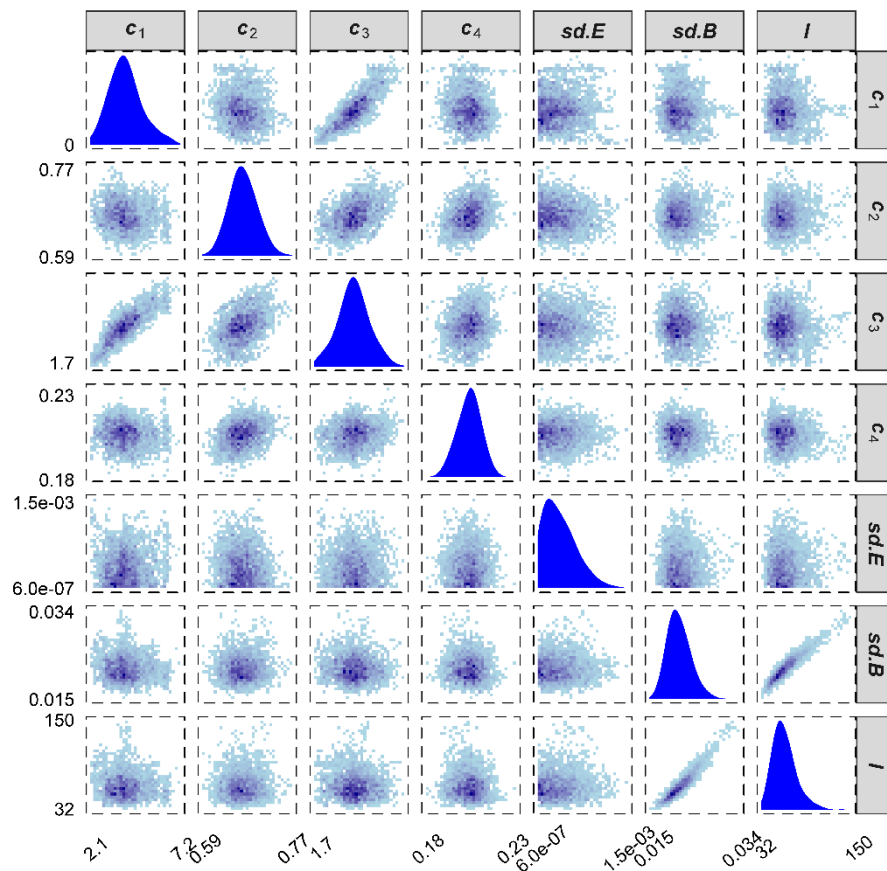

Figure 5.3: Parameter distribution and bivariate correlation

5.3.1.2 Estimation of parameter and predictive uncertainty

Figure 5.4 shows the uncertainty associated with the estimation of wash-off fraction. These uncertainty bands are estimated by drawing samples from the posterior multivariate distribution. Parameter uncertainty was estimated by using deterministic model $(y(x, \theta))$ runs and predictive uncertainty was estimated by using the deterministic model together with error model components $(Y(x, \theta, \psi))$. Since the latter includes the uncertainty due to model bias and measurement noise these bands are wider than the parameter uncertainty. The total predictive uncertainty which accounts for both model bias and measurement noise accounts for ~ 0.1 (10%) uncertainty in the wash-off fraction. This constant trend of predictive uncertainty is a reflection of the fact that the error model used here is not explicitly input-dependent bias model, but rather it is a constant bias (variance) model. On the other hand, parameter uncertainty increases with increasing wash-off fraction as the variance of parameter uncertainty proportionally increases with mean prediction. The parameter uncertainty accounts for a maximum of 0.06 (6%) wash-off fraction when 95% predictive interval is considered.
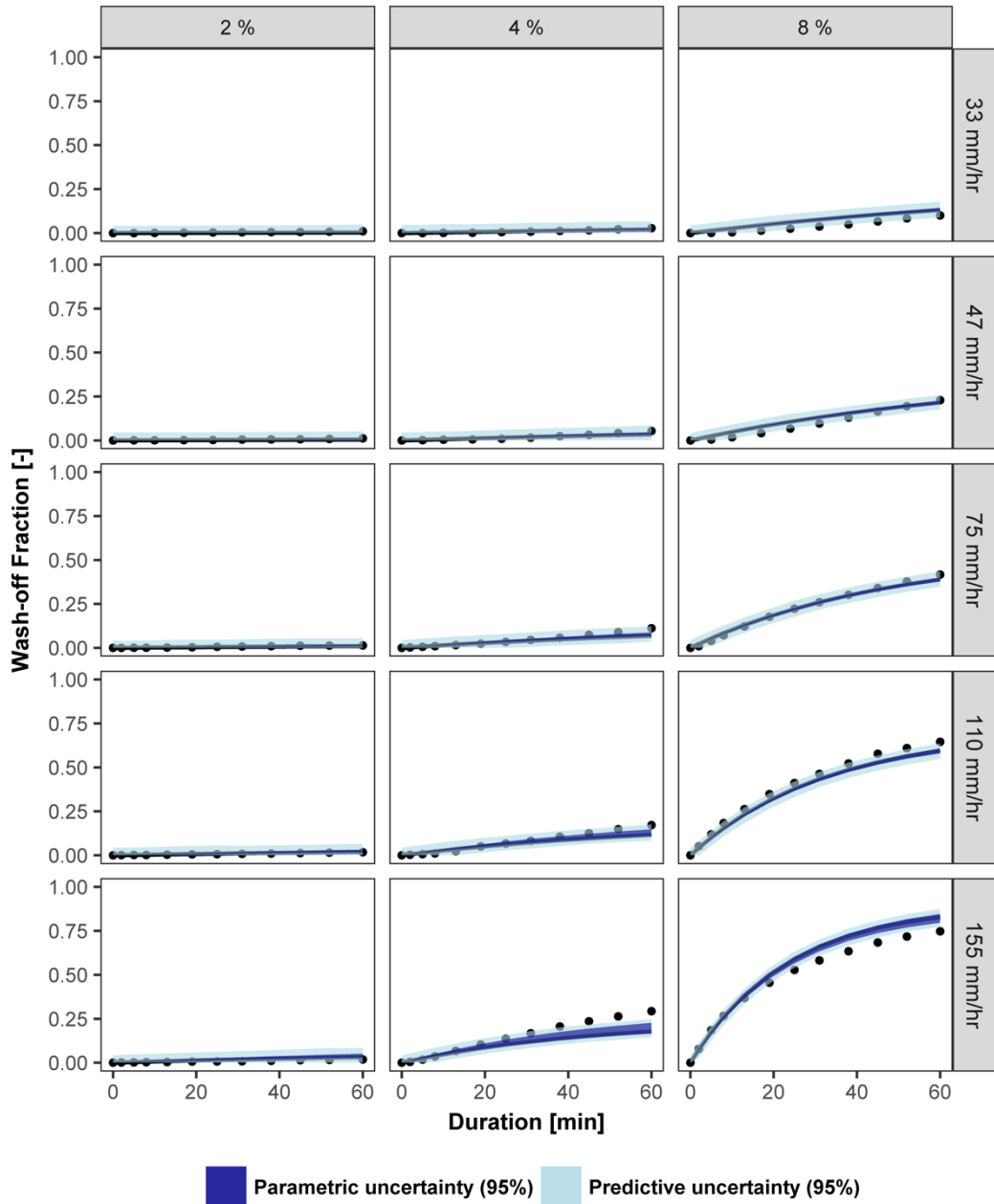
Figure 5.4: Uncertainty associated with the estimation of wash-off fraction

To check the reliability of the uncertainty estimation, prediction interval coverage probability (PICP, Ref Eq.(5.8)) which measures the probability that the observed values lie within the estimated prediction intervals (Shrestha and Solomatine, 2006) was used.

$$PICP = \frac{1}{n}\sum_{i=1}^{n} R * 100\% \qquad where\ R \begin{cases} 1, & PL_t^u \leq O_t \leq PL_t^l \\ 0, & otherwise \end{cases} \qquad (5.8)$$

Where, PL_t^u, PL_t^l are upper and lower boundary of the considered prediction interval at time t for a given slope and rainfall intensity, O_t is corresponding measured wash-off fraction at time t. For a better performance PICP should be close to the considered prediction interval, which is 95% in this case. The calculated PICP during validation stage is 82%, so corresponding accuracy of the uncertainty estimation is around $\sim 85\%$ which essentially means that the error model is able to predict the uncertainty reasonably well.

5.3.2 Effect of rainfall uncertainty in wash-off prediction

Figure. 5.5 presents propagated uncertainty through the NEM for event 11 with a surface slope of 4%. The effect of changes in the slope is discussed later in this section. Event 11 has average conditions in terms of event duration and peak intensity. Prediction errors of other rainfall events follow the same pattern of behaviour and they are presented in Appendix 5A. Figure 5.5 shows predicted values and associated uncertainty in rainfall intensity (first row of the plot), instantaneous wash-off fraction (second row of the plot), cumulative rainfall (third row of the plot) and cumulative wash-off fraction (fourth row of the plot). Further, each column shows the variation of predicted values and associated uncertainty for temporal averaging intervals of 2 min, 5 min, 15 min and 30 min.

It can be seen from Fig.5.5 that the instantaneous wash-off fraction varies almost proportionally against rainfall intensity. The variation in cumulative wash-off fraction plotted against cumulative rainfall also validate this observation. This observed proportional change of instantaneous wash-off fraction against the rainfall intensity throughout the event is due to fact that the 'first-flush' effect is negligible for this event. The concept of "first flush" is that the initial period of storm flow carries most of the pollutant including sediments from the urban surface (Helsel *et al.*, 1979; Sansalone and Buchberger, 1997; Bertrand-Krajewski *et al.*, 1998; Muthusamy *et al.*, 2018). This observation is in agreement with the observed wash-off fraction (Fig. 5.1 and Fig. 5.2) where the wash-off pattern shows a proportional change against smaller rainfall intensities. As seen from Fig.5.1 and as stated in Muthusamy et al (2018) the negative-inverse-exponential trend in total cumulative wash-off fraction due to the effect of first flush is clearly observed at combinations of catchment slopes steeper than 8% and rainfall intensities higher than 75 mm/hr lasting longer than 30 min. But in reality such higher rainfall intensity events are rare and therefore the wash-off is expected to behave

99

as presented in Fig. 5.5 for a similar sediment size and surface roughness condition. In fact none of the events from Muthusamy et al. (2017) show the effect of 'first-flush'.

To demonstrate the effect of "first flush" and to check if the NEM is able to capture the effect of 'first-flush', a synthetic rainfall event is created by using rainfall event 11 and a rainfall multiplier of 6 was used to provide sufficiently large intensities. This synthetic rainfall event is created based on extreme scenario recorded in UK. The highest ever recorded 5 min rainfall intensity in UK is 384 mm/hr (MetOffice UK, 2016) hence the peak of the synthetic rainfall event is kept within 300 mm/hr. Further, a surface slope of 8% is used. This rainfall intensity and surface slope conditions are able to produce a well pronounced "first-flush" effect as shown in Fig. 5.6. This figure shows the peak wash-off fraction corresponds to the first peak of the rainfall event becomes more dominant compare to the next peak due to the more pronounced "first-flush" effect.

Since instantaneous wash-off fraction varies almost proportionally against rainfall intensity, uncertainty in instantaneous wash-off prediction also varies proportionally to uncertainty in rainfall intensity. At 2 min temporal averaging interval where the uncertainty in rainfall prediction is the highest, ~7% CV of peak intensity resulted in a~8% CV on the prediction in instantaneous wash-off fraction peak. The uncertainty associated with the cumulative rainfall at the event at 2 min temporal averaging interval is 1.7% CV and it resulted in a 2.7% CV in the prediction of cumulative wash-off fraction at the end of the event. At 30 min interval the since the uncertainty in the rainfall intensity is negligible, the corresponding uncertainty in the prediction in the wash-off fraction becomes negligible. But note that as mentioned before, this uncertainty is only due to spatial variability and measurement error of rainfall which are negligible for temporal averaging interval of 30 min. But on the other hand, due to temporal aggregation there is a significant reduction in peak wash-off load due to the corresponding reduction in the rainfall intensity peaks. Temporal aggregation of rainfall intensity from 2 min to 30 min reduced the peak rainfall intensity by ~65% and this resulted in ~ 70% reduction in peak of instantaneous wash-off fraction. Although there is not much reduction in cumulative rainfall due to temporal aggregation, there is still a ~20% reduction in corresponding cumulative wash-off fraction. This shows the sensitivity of cumulative wash-off fraction to rainfall peaks and why aggregation

measures such as total rainfall and total runoff often result in under- prediction of wash-off loads.
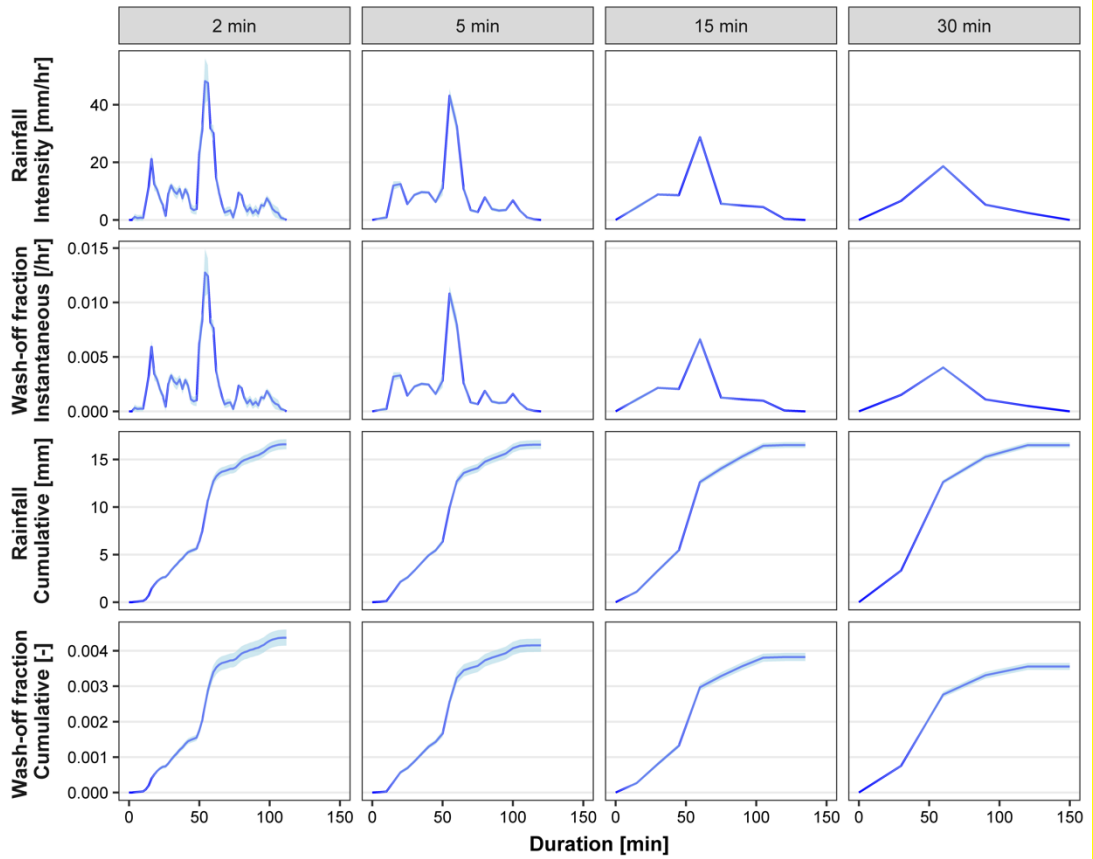

Figure 5.5: Variability in wash-off fraction (instantaneous and total) corresponds to uncertainty in rainfall (intensity and total) for different temporal averaging intervals
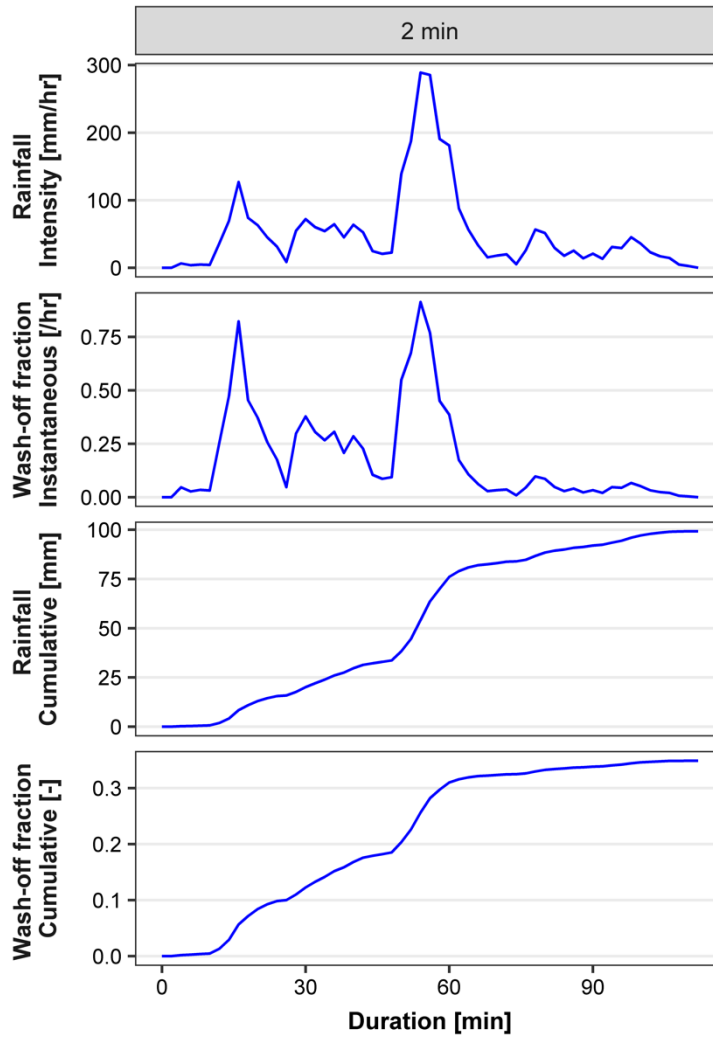
Figure 5.6: Demonstration of "first flush" effect on wash-off using a synthetic rainfall event

Figure 5.7 shows predicted values and associated uncertainty (CV %, shown in labels) in rainfall event peaks and corresponding wash-off fraction peaks for temporal averaging intervals 2 min and 30 min of all13 actual rainfall events. As already seen from Fig.5.5, temporal aggregation of rainfall intensity from 2 min to 30 min reduces prediction of peaks of wash-off fraction significantly. The highest peak which corresponds to event 8 reduced from 0.024 h^{-1} to 0.0048 h^{-1} showing ~80% of reduction in wash-off fraction due to similar level of reduction in rainfall peaks. On the other hand since the temporal aggregation reduces the uncertainty of rainfall peaks, it consequently reduces the uncertainty in the prediction of wash-off peak as well. The highest uncertainty in the wash-off fraction peak is from event 4 in which the uncertainty in wash-off fraction reduced from 15% to ~2% corresponds to similar amount of reduction of uncertainty in rainfall peak.
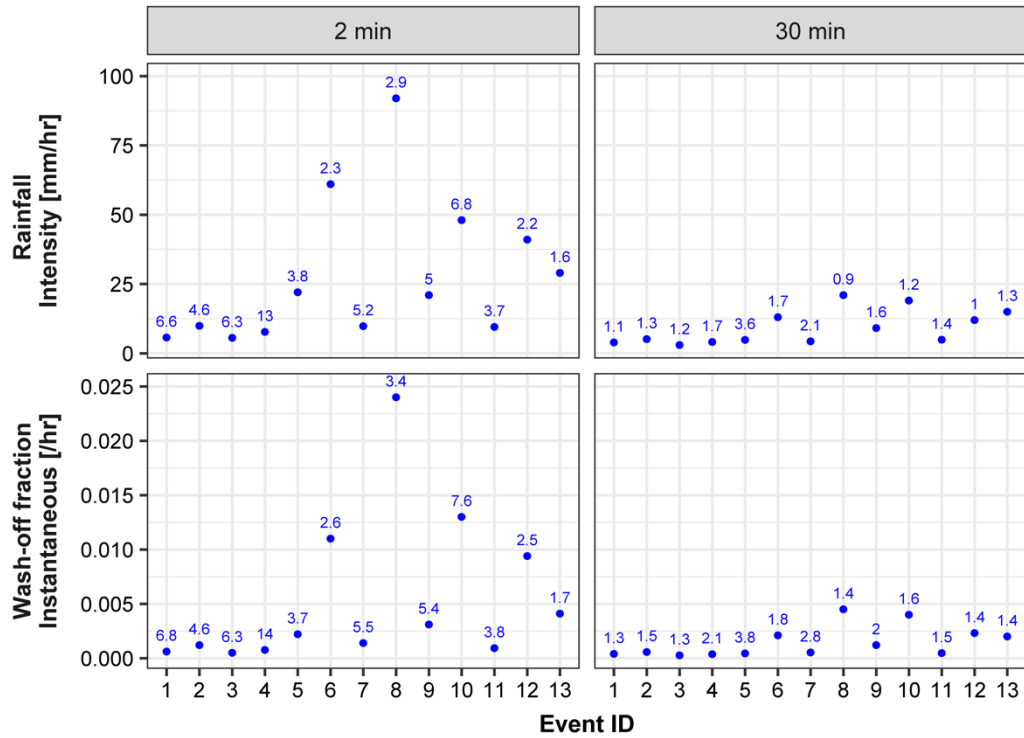
102

Figure 5.7: Uncertainty in peak wash-off fraction corresponding to the uncertainty in peak rainfall intensity of all 13 events presented in Table 5.2

Figure. 5.8 where CV of predicted instantaneous wash-off (CV_w, applicable to both load and fraction) plotted against CV of rainfall intensity (CV_R) for all the data from 13 events at 2 min temporal averaging interval at surface slope of 2%, 4% and 8%. Further, to check if there is an effect of "first-flush" we divided the data into two portions; uncertainty corresponds to rainfall intensities at the start of the event (initial period) and uncertainty corresponds to rainfall intensities from the following period (following period). The division is based on duration of each events. The initial period is the first 10% of the total event duration and the following period is remaining 90% of the event. For example, if the total duration of the rainfall event is 100 min then initial period is first 10 min and following period is remaining 90 min.

First of all, looking at the trend of CV_w against the CV_R, it clearly shows that there is a linear trend (i.e. similar level of uncertainty in instantaneous wash-off due to a certain level of uncertainty in rainfall intensity) up to around 40%. After that CV_w shows a clear decreasing trend against CV_R and around 100% of CV_R, CV_w becomes almost constant against CV_R. The linear trend is expected due the proportional change in instantons wash-off against rainfall intensity as there is no "first-flush" effect as already

discussed. The decreasing trend of CV_w after 40% means that the change in the instantaneous wash-off against rainfall intensity is less than proportional. The only reason could be that the rainfall intensities with $CV_R > 40\%$ are not as high as rainfall intensities with $CV_R < 40\%$ to produce a proportional change in the instantaneous wash-off. From Fig. 5.9 it is indeed clear that $CV_R > 40\%$ belongs to rainfall intensities that are smaller than 5 mm/hr which will produce a very little wash-off, hence the corresponding variability due to the uncertainty in rainfall intensities that are less than 5 mm/hr will be much lesser than higher rainfall intensities where CV_R is > 40%. The reason of high CV_R for rainfall intensities less than 5 mm/hr was shown due to the higher measurement error in Muthusamy et al. (2017).

Looking at the effect of different duration of the rainfall event (initial period and following period), there is no visible difference in the trend indicating the absence of the "first-flush" even at 8% slope. We tried with different initial periods (First 5 %, 20% and 50 % duration of the total duration of the events), but the behaviour was same and there was no visible difference in the trend for any cases.
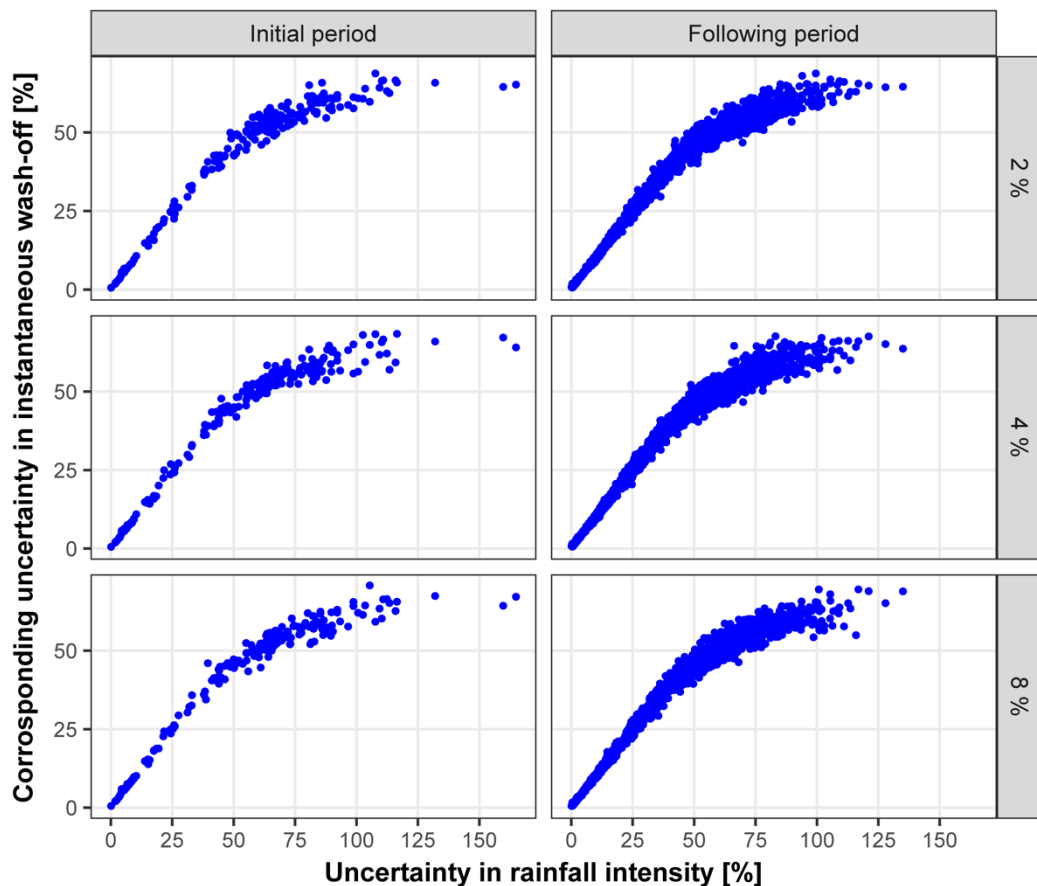

Figure 5.8: CV of predicted wash-off peaks plotted against CV of rainfall peaks
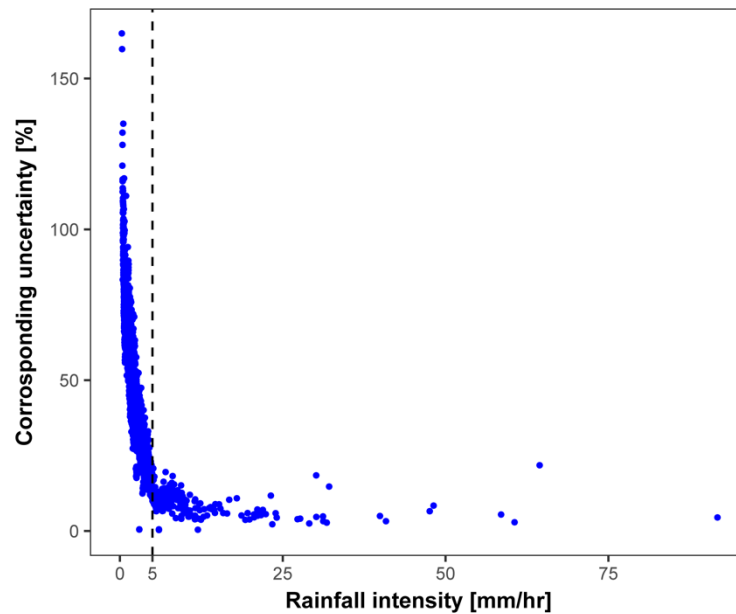
Figure 5.9: Rainfall intensity and corresponding uncertainty at 2 min temporal averaging interval for all the events presented in Table

5.3.3 General discussion

Note that in addition to rainfall intensity and surface slope, other parameters such as sediment size and surface texture will also affect the sediment wash-off, but due to the limitations in the data used in this study, the NEM does not include the effect of these parameters. With smaller sediment sizes and smoother surfaces the wash-off is expected to be higher and this will also effect the occurrence of "first-flush". For example, Egodawatta et al. (2007) in a similar experimental study used a smaller sized sediment resulting in a relatively higher wash-off fraction. Hence the application of the NEM, like the OEM needs to be checked against different sediment sizes and surface textures. Inclusion of the effect of these parameters explicitly might introduce more complexity in the equation, but nevertheless such an equation can be applicable globally regardless of individual catchment conditions. This is one of the research areas in sediment wash-off modelling that requires to be investigated in detail.

IID is the most commonly used form of error model in urban hydrology (Freni *et al.*, 2009; Dotto *et al.*, 2011; Breinholt *et al.*, 2012; Bonhomme and Petrucci, 2017) mainly because of its simplicity. But it requires absence of a serial correlation in the error distribution, which can lead to underestimation of uncertainty and biased parameter

estimates (Del Giudice *et al.*, 2013). This makes IID less robust for different urban hydrological applications. The autoregressive bias error model used in this study, although more complicated, does not have these requirements and is more robust. We assumed a constant bias to keep the autoregressive error model simple, but it is also possible to describe it as an input – dependent bias (Del Giudice *et al.*, 2013) where bias can be a function of both slope and intensity. The advantage of such bias description still needs to be investigated in the uncertainty analysis of wash-off modelling in the future.

5.4 Conclusions

In this study, first, we proposed an improved exponential wash-off model by replacing the calibration parameters of the original exponential model with functions of rainfall intensity and surface slope, making the model more robust to a new set of catchment conditions. This will not only avoid the need of look up tables or charts and interpolation or extrapolation, but it will also introduce some transparency in the parameter estimation which is otherwise a black box approach. This new exponential model (NEM) was calibrated and verified using the experimental data collected for different combinations of surface slopes and rainfall intensities. Bayesian inference, which allows the incorporation of prior knowledge, is implemented to estimate the distribution of the parameters of the newly introduced functions. Second, by statistically describing model bias and measurement noise, predictive uncertainty in the prediction of NEM was estimated. Finally, the propagation of rainfall uncertainty due to sub-kilometre spatial variability and measurement error of rainfall through the NEM was investigated in detail.

Although during the calibration stage OEM performs better than NEM, it has to be taken into account that OEM had to be calibrated for each and every experimental condition separately. Further, at the validation stage, NEM performance improved over OEM, reflecting the ability of the new exponential model to perform better under a range of new catchment conditions. Verification measures show the uncertainty estimates associated with the NEM predictions are plausible, indicating that the use of two error terms, autoregressive error and independently identically distributed error, to represent model bias and measurement noise respectively was a reasonable representation of the error process associated with sediment wash-off modelling. The

total predictive uncertainty which accounts for both model bias and measurement noise accounts for ~ 0.1 (10%) uncertainty in wash-off fraction when 95% predictive interval is considered out of which a maximum of 0.06 (6%) comes from the parameter uncertainty.

The effect of uncertainty in rainfall intensity in sediment wash-off can be concluded as below

1. $CV_R > CV_W$

   This is where the rainfall intensities are too small to produce a proportional change in the instantaneous wash-off against the corresponding change in the rainfall intensity, hence the uncertainty in rainfall intensities (CV_R) will cause a lesser level of uncertainty in wash-off fraction (CV_W)

2. $CV_R = CV_W$

   This is the most commonly observed case where the change in the instantons wash-off is proportional to change in the rainfall intensity, hence any uncertainty in rainfall estimation would produce a similar level of uncertainty in the prediction of instantaneous wash-off. This is when there is no effect of "first-flush"

3. $CV_R < CV_W$

   When there is an effect of "first-flush", then most of the sediment will be washed off at the beginning of the event and corresponding instantaneous wash-off will be larger than previous two cases, therefore a certain level of uncertainty in a rainfall intensity would produce a larger level of uncertainty in corresponding predicted instantaneous wash-off at the beginning of an event.

None of the measured rainfall events from Muthusamy et al. (2017) produced "first flush" effect, hence for most common rainfall and surface slope conditions it was found that any uncertainty in rainfall estimation would produce a similar level of uncertainty in the prediction of instantaneous wash-off ($CV_R = CV_W$). Consequently, the maximum uncertainty in the peak instantaneous wash-off fraction due to measurement and sampling error and spatial variability of rainfall within a spatial extend of 8 ha found to be 13% when a temporal averaging interval of 2 min is considered.

It should be noted that the occurrence of "first-flush" is sensitive to optimal values of $c_1,..,c_4$ in NEM and these values needs to be checked against different sediment sizes and different surface roughness as these are two other major external drivers which would affect the sediment wash-off. Nevertheless, the model structure of NEM would be applicable for any sediment size and surface texture as the underlying physical processes will be the same as those on which the model structure of NEM was developed

Acknowledgement

Appendix

5A: Figures showing propagated uncertainty through NEM for all 13 events mentioned in Table 5.2 for temporal averaging intervals of 2 min and 30 min.
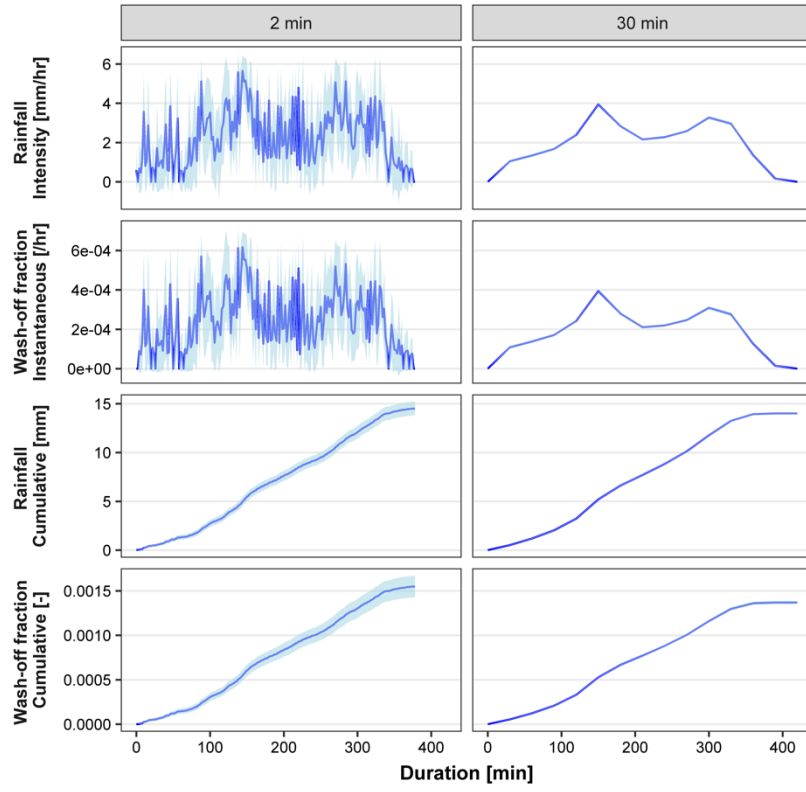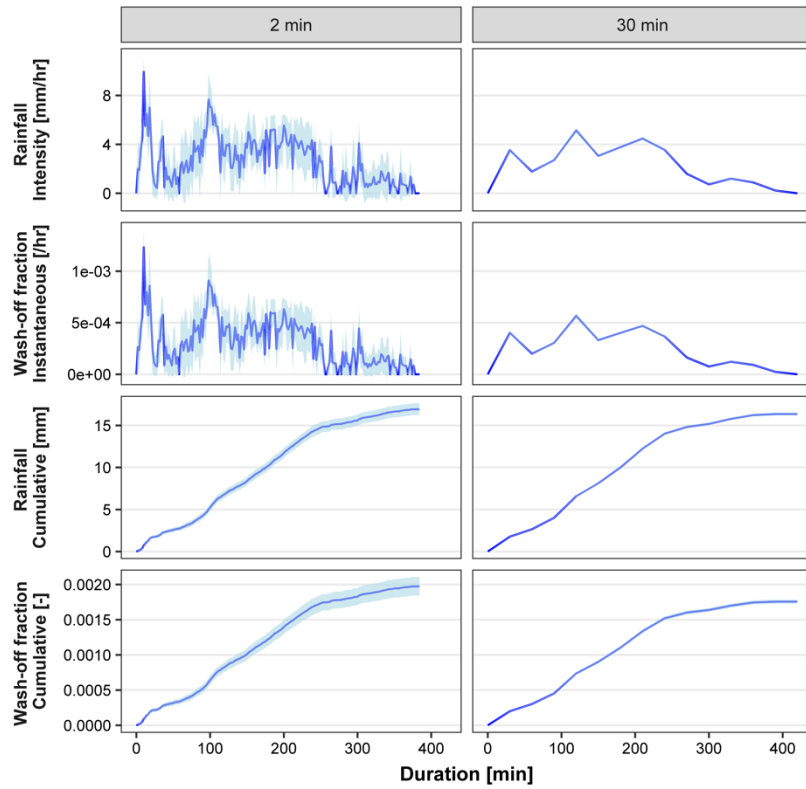
Event ID - 3
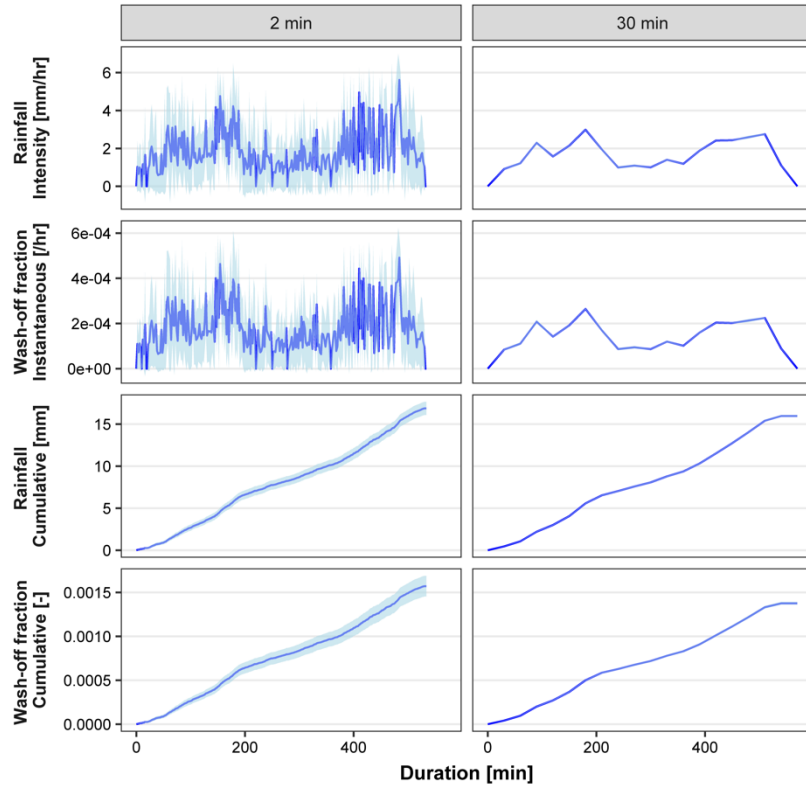
Event ID - 4
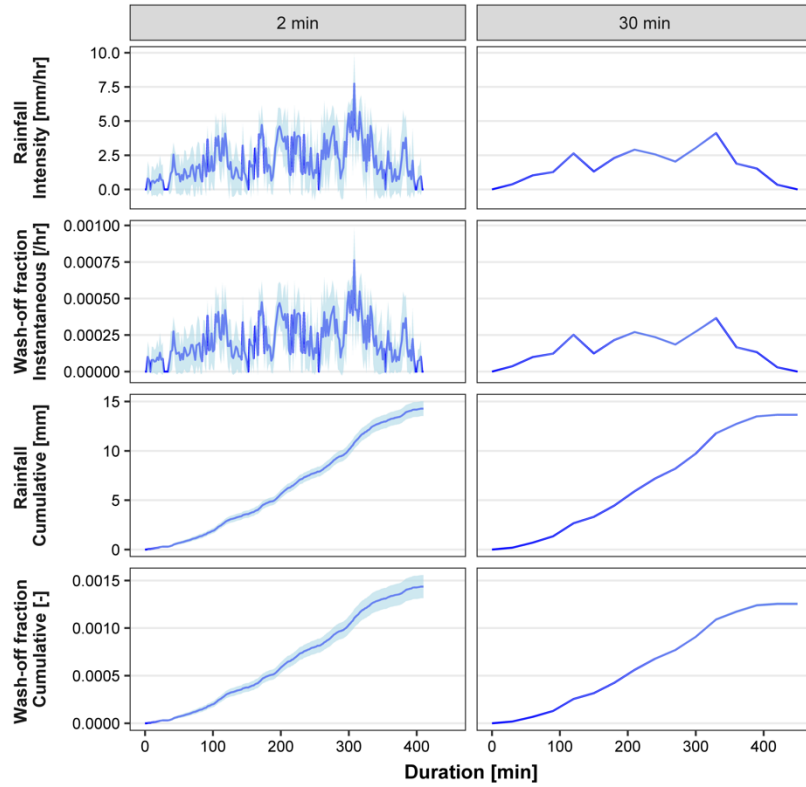
Event ID - 5
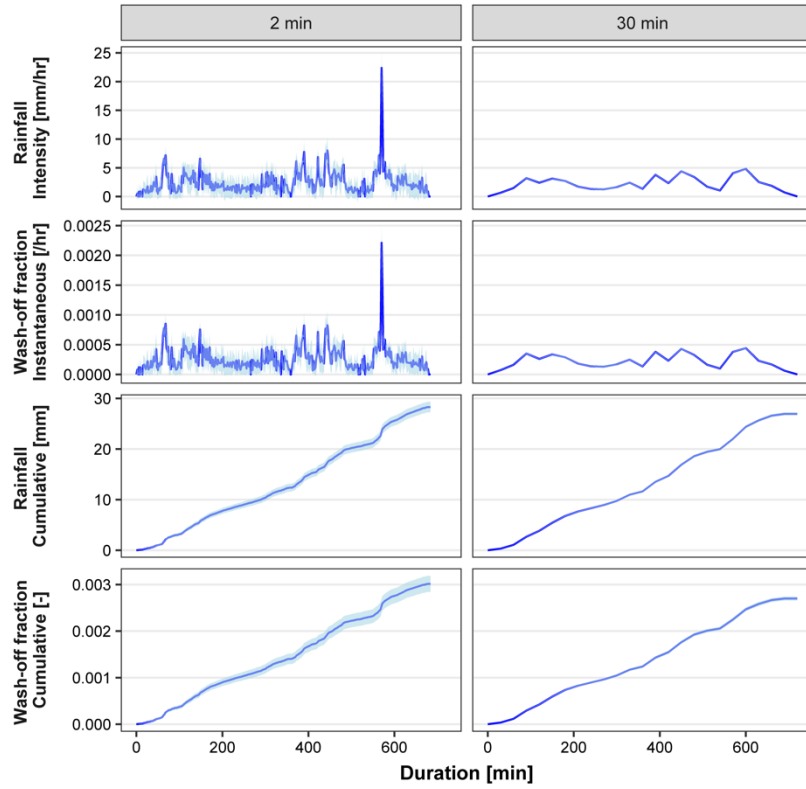
Event ID - 6
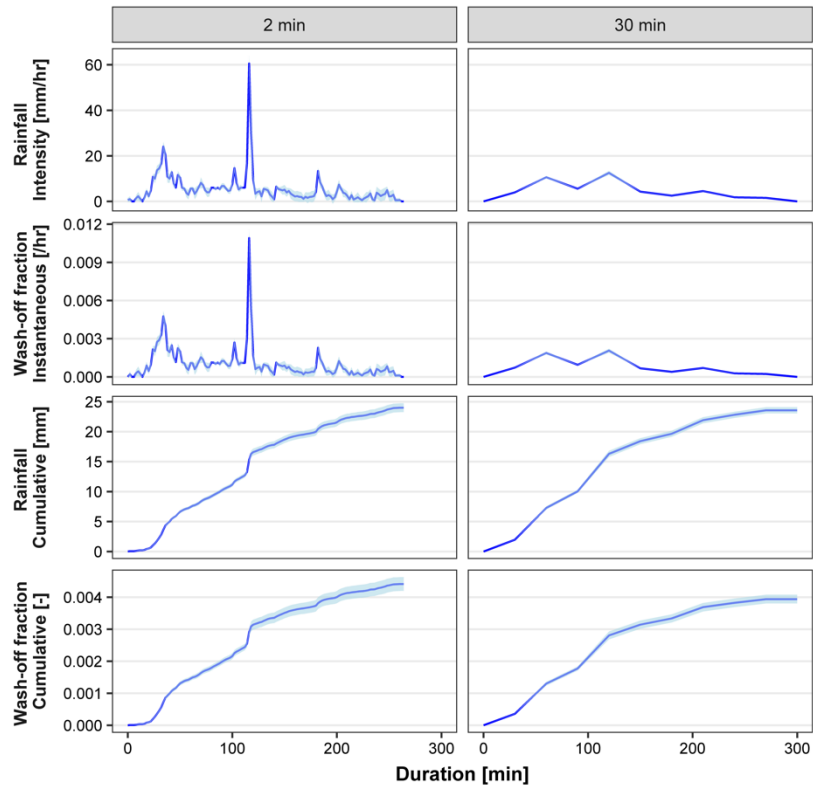
Event ID - 7
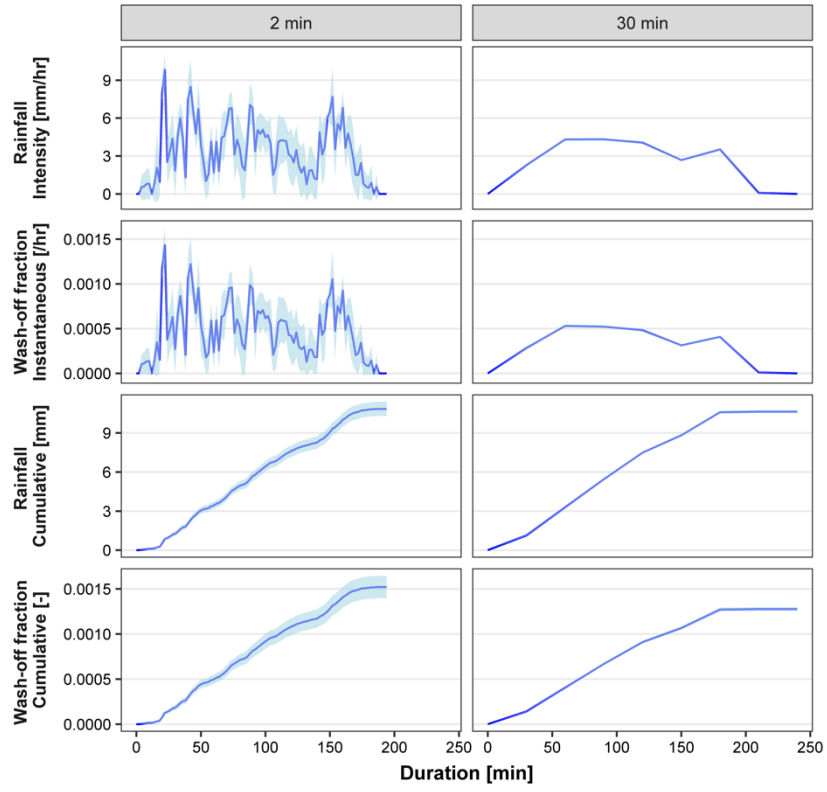
Event ID - 8
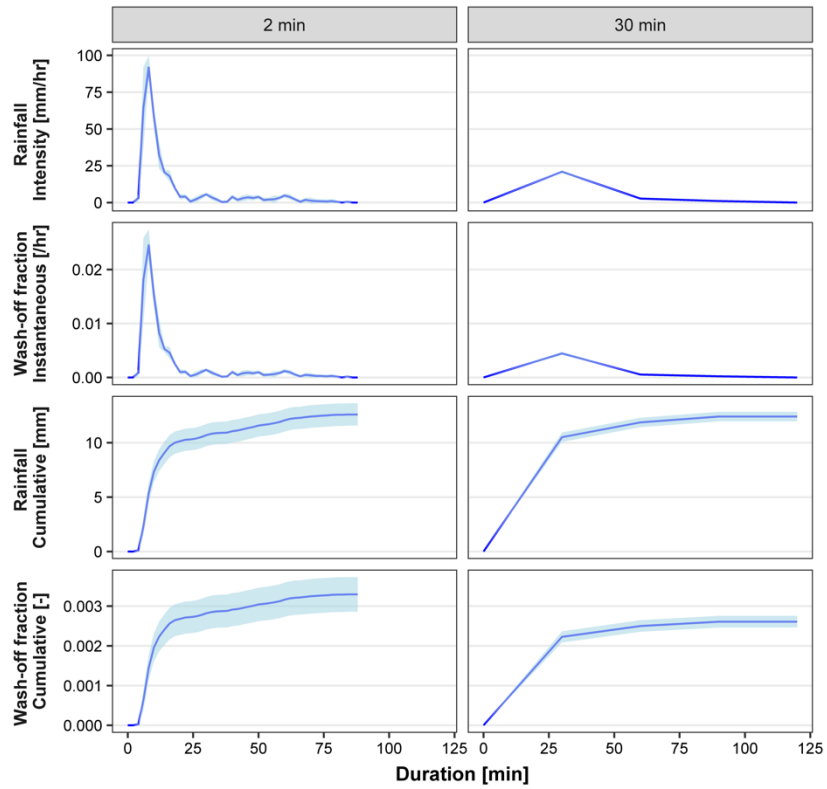
Event ID - 9
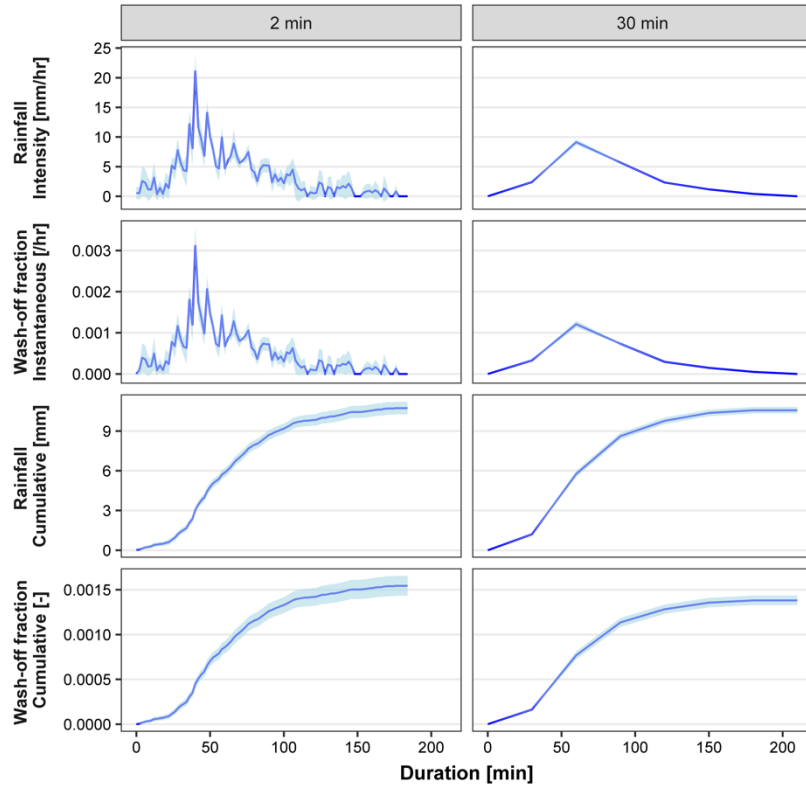
Event ID - 10
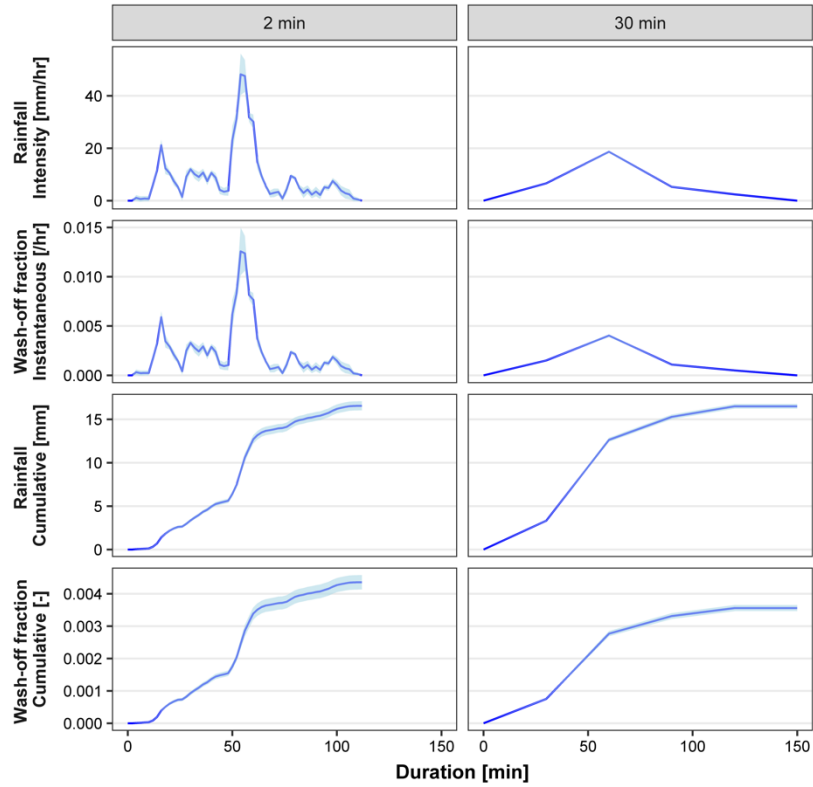
Event ID - 11
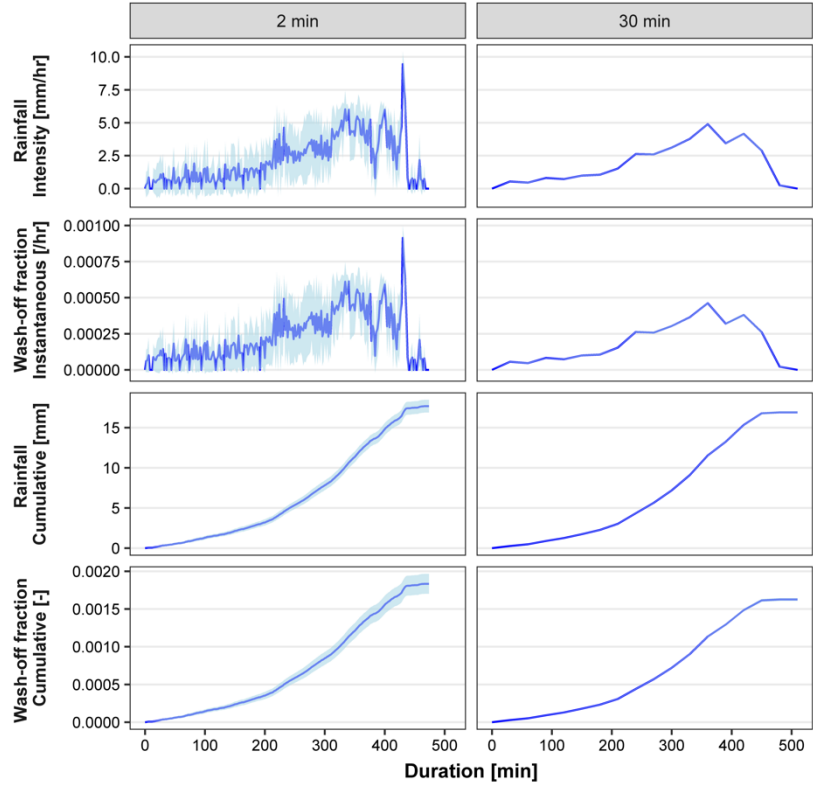
Event ID - 12
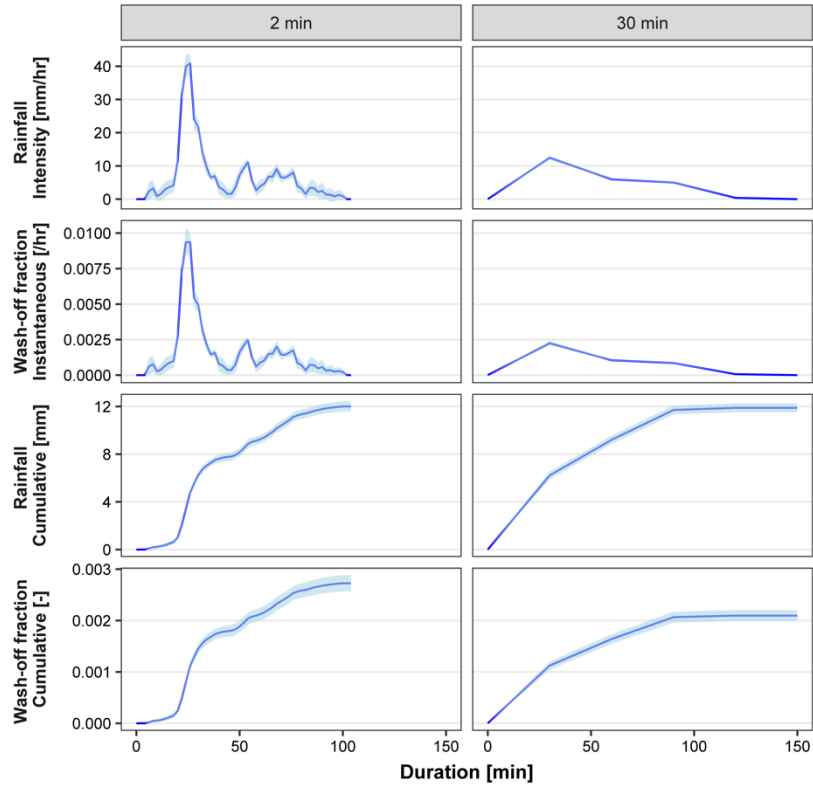
Event ID - 13
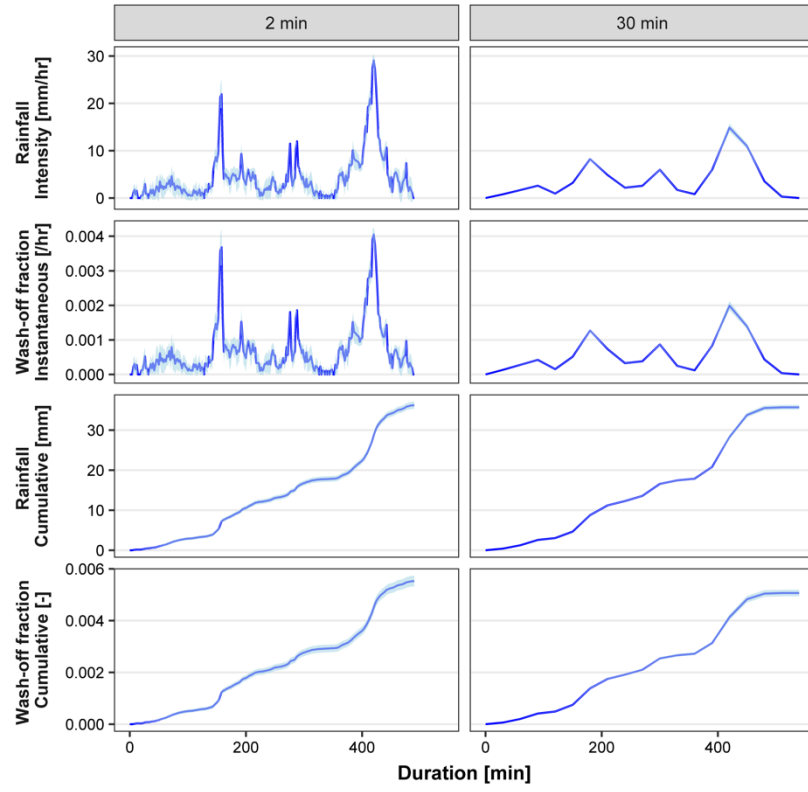
6. Summary, Discussion and future works, conclusions

6.1 Overarching summary

The main aim of the research was to investigate the effect of uncertainty caused by rainfall variability at a selected urban scale in order to improve sediment wash-off predictions from urban catchment surfaces. This research work was carried out in three parts.

The aim of the first part of the study was to obtain a stochastic description of urban-scale spatial variability in rainfall in a way that can be used in lumped sediment wash-off models. Through literature review, it was found that (1) uncertainty due to rainfall variability at a sub-kilometre scale is significant in the modelling of any hydrological process that is driven by rainfall and runoff such as sediment wash-off and (2) geostatistical methods, despite their challenging data requirements, can be modified and developed to study the spatial variability of rainfall. This is due to their capability to take into account the spatial correlation structure of rainfall data and their ability to provide quantification of uncertainty in upscaling. Taking into account these findings, a geostatistical method was developed to estimate the spatially averaged rainfall intensity together with the associated level of uncertainty. High spatial resolution rainfall data collected from a cluster of eight paired rain gauges in a 400m × 200m urban catchment was used to develop this methodology. The spatial lag of the rain gauge network ranges from ~20 to ~400 m. As far as the author is aware, this is the smallest spatial scale in which the variability in rainfall has been examined using high temporal resolution point rainfall measurements in urban hydrology. Unreliable data which were detected by making use of the paired rain gauge set up were omitted prior to geo-statistical analyses. Variogram, which is a widely accepted geo-statistical measure, was used to illustrate the spatial variability of rainfall for different combinations of the temporal averaging interval (2 min, 5 min, 15 min and 30 min) and different range of rainfall intensities (< 5 mm/h, 5-10 mm/hr and > 10 mm/h). This was the first time that geostatistical models such as variograms have been assigned to a combination of rainfall intensity ranges and temporal averaging intervals. These variograms were then used in spatial stochastic simulations to obtain spatially averaged rainfall intensities together with associated uncertainties for the same combinations.

The two main challenges typically associated with rainfall data in an urban catchment addressed in this study were the scarcity of rainfall measurement locations and non-normality of rainfall data, both of which needed to be considered when adopting a geostatistical approach.

The aim of the second part was to improve the understanding of sediment wash-off from urban surfaces and to establish the correlation between calibration parameters and external drivers in the current wash-off model. From the literature, it was understood that the current wash-off models still need to be improved in terms of representation of the interaction between the external drivers associated with rainfall, catchment surface and sediment characteristics. It was also noted that the current sediment wash-off model structure needs to be improved in order to be able to differentiate and quantify different sources of errors and their propagation, a feature that will be required when rainfall error propagation is investigated. Hence, before investigating the propagation of rainfall error quantified in the first part of the thesis, the widely used exponential wash-off model currently in practice was improved. Taking the research gaps identified through literature review into consideration, laboratory experiments were conducted to investigate the effect of three selected external drivers, rainfall intensity, surface slope and initial load on wash-off load, in an integrated and systematic way. The experimental set-up comprised of a rainfall simulator, a 1 m^2 bituminous road surface, and a continuous wash-off measuring system. Five rainfall intensities ranging from 33 to 155 mm/h, four slopes ranging from 2 to 16% and three initial loads ranging from 50 to 200 g/m^2 were selected based on values obtained from the literature. Fine sediment with a size range of 300–600 μm was used for all of the tests. This was the first time where the effect of all the above three dominant parameters on wash-off load is investigated in an integrated and systematic way. Using the experimental results the original exponential equation which is still in practice was improved by establishing the correlation of two calibration parameters, capacity factor and wash-off coefficient, against rainfall intensity and catchment surface slope.

In the final part of the study, the propagation of different sources of uncertainty, including rainfall uncertainty, in improved sediment wash-off modelling was investigated. This task was carried out in three steps. First, the wash-off model derived from the second part was improved further by replacing the calibration parameters with functions of rainfall intensity and surface slope making the model more robust to new

catchment conditions. Further, replacing the invariant calibration parameters with functions of external drivers (i.e. rainfall intensity and surface slope) made it easier to investigate the propagation of errors in the external drivers (e.g. rainfall intensity) as these external drivers are now explicitly defined in the new equation. Bayesian inference, which allows the incorporation of prior knowledge, was implemented with Markov Chain Monte Carlo (MCMC) sampling method to estimate the posterior probability distribution of the parameters of the newly introduced functions. In the second step, different sources of error in the prediction of this newly improved sediment wash-off model were separately quantified. Uncertainty due to model bias and measurement noise was separately quantified by explicitly modelling them as an autoregressive bias term and an independent error term respectively in the likelihood function of the Bayesian framework. In the final step, the propagation of rainfall uncertainty obtained in the first part of the study was propagated through the new improved wash-off model and its impact was investigated. This uncertainty propagation was investigated (1) for different temporal averaging intervals (2) for different surface slope conditions and (3) for different periods of the rainfall events.

6.2 Discussion and future works

Rainfall uncertainty analysed in this study essentially comes from two main sources: Natural spatial variability of rainfall and measurement and sampling error. Quality control measures such as paired gauges were used in this study are able to limit the measurement errors of rainfall measurement. However, it is impossible to completely avoid measurement error (and sampling error in case of tipping bucket type rain gauges). Hence, in this study, the effect of spatial variability is studied together with the inherent measurement and sampling error associated with tipping bucket type rain gauges. Although variograms provide information on the total measurement error and microscale spatial variability (the nugget effect), it does not have provision to separately quantify them. However, in Chapter 3 it was seen that when the rainfall intensities are higher (> 10 mm/hr) nugget effect becomes smaller indicating smaller measurement and sampling errors. This implies that most of the total uncertainty in this rainfall intensity range is due to natural spatial rainfall variability. Hence, the uncertainty in wash-off peaks caused by rainfall intensity peaks which are mostly for rainfall intensities > 10 mm/hr is strongly related to the natural variability of rainfall.

Separating uncertainty caused by measurement and sampling error from uncertainty caused by natural spatial variability is not a straightforward task. Such separation needs quantification of measurement and sampling error associated with tipping bucket type rain gauges and this quantity also depends on other factors such as rainfall intensity range and site conditions. This makes the exact quantification of measurement and sampling error of point rainfall data an area which needs to be investigated on its own merit. Such a detail investigation would help to separately quantity the two major sources of uncertainty in upscaling of rainfall data which are its natural variability and measurement and sampling error. However, it is clear from the nugget values indicated in the field data set that at the higher intensities that the measurement and sampling uncertainty dropped and that the effect of the natural rainfall variability increased.

As seen from Chapter 3, a distinction between intensity classes is important when analysing the spatiotemporal variability of rainfall. In this study, the entire rainfall intensity range was divided into three classes considering the available data. Another possible extension could be to develop different geo-statistical models based on rainfall type in addition to intensity range. Convective, transitional, and frontal are three different rainfall types which can be separated and since the nature of each rainfall type is different, the variability of rainfall is expected to be different even for the same intensity range (Jaffrain and Berne, 2012). By doing so, the uncertainty due to rainfall variability can be derived for different rainfall types which could be useful when uncertainty propagation of separate rainfall events in a hydrological prediction is investigated.

The inclusion of capacity factor in sediment wash-off modelling is based on the finding from previous studies (Egodawatta *et al.*, 2007; Egodawatta and Goonetilleke, 2008) that a rainfall event can remove only a fraction of sediment from the surface and once a rainfall event reached the capacity, there is no more wash-off even when there is a significant amount of sediment is remaining on the surface. The underlying physical interpretation is that a rainfall event can mobilise only the particles that are smaller than a specific size and this size increases with increasing rainfall intensity. In this PhD, in addition to rainfall intensity, the effect of surface slope on capacity factor was investigated. The surface slope was chosen as the underlying physical processes of wash-off - rainfall drop impact and shear stress from runoff - are both functions of the surface slope. Hence, one of the hypotheses of chapter 4 was that the surface slope will

have an effect on capacity factor. Results also showed that the hypothesis was correct and slope indeed has a significant effect on capacity factor and consequently on wash-off load.

Regarding the sediment size used in this study, a D_{50} of ~450 μm was selected based on field observations on urban road sediment size distribution (Bertrand-Krajewski *et al.*, 1993; Butler and Clark, 1995). Further, the range of 300 μm (D_{10}) - 600 μm (D_{90}) was selected so as to provide a well-characterised sediment as this removes the potential for size sorting and so allows us to link the wash-off behaviour of a particular sediment size to the rainfall driver. It also provides the possibility to compare with the effect on the physical wash-off process of another different well-defined sediment size in the future. However, it is acknowledged that the sediment size chosen would have had an effect on the results. It is expected that the transport capacity would be linked with particle size and that the relative size of the particle to catchment surface roughness would affect the amount of available sediment and the threshold of motion.

Most of the sediment wash-off tests were not run long enough to observe a plateau in the cumulative sediment wash-off fraction (Fig 4.6) when the event reaches the maximum capacity. However, most of the tests indicated that the gradient of the cumulative sediment wash-off had decreased very significantly towards the end of the tests and extrapolation of the data indicated that this gradient was trending to a plateau. This extrapolation indicated that some of the tests would take more than 10 hours to reach the maximum available sediment capacity. For example, the test with rainfall intensity of 47 mm/hr and slope of 4% would have taken around 15 hours to reach the maximum capacity according to the Eq.4.5. There were 50 tests altogether and running tests around 15 hours each would have significantly reduced the number of tests and would not have permitted the study of the three external parameters that were found to be dominant. Nevertheless, all the experiments were run for at least an hour longer than most of the experiments reported in previous studies which investigated sediment wash-off using similar experiments (e.g. Egodawatta *et al.*, 2007). Further, when the rainfall intensity is 155 mm/hr and the surface slope 16% it can be seen that the cumulative wash-off fraction almost reaches a plateau (percentage change of ~2% over the last 8 min compare to overall change of ~86% over an hour) even though more than 10% sediment remains in the surface. This shows that even with the most extreme conditions in the experimental series (steepest slope and most intense rainfall) some of

the sediment was not washed off implying the other experimental cases with smaller rainfall intensity and smaller surface slopes would also have reached a smaller maximum capacity if they were run for long enough, as the driving physical processes would be the same. However, it is acknowledged that it would have been better if a few experiments were run long enough to physically observe the plateau. This would have given a better justification of the inclusion of capacity factor. Hence, this needs to be taken into account when deciding the duration of tests in the future wash-off experimental studies.

In Fig. 4.6, in a few cases where rainfall intensity is > 110 mm/hr there is a distinct pattern of underestimation of the model especially after 30-40 min (Eq. 4.5). This underestimation could possibly be due to underestimation of the capacity factor in Eq.4.5. However, this has been resolved when using NEM as presented in Fig 5.2 where there is no distinct behaviour of underestimation or overestimation. Although there is an underestimation in the verification stage when rainfall intensity is > 110 mm/hr, the prediction is still better than OEM where there is a systematic overestimation in all verification cases.

Rainfall intensities used in the experiments ranged from ~30 mm/hr to ~150 mm/hr which is on the higher side of most rainfall intensities observed in the UK. However, the minimum intensity of ~ 30 mm/hr was chosen based on the trial experiments to produce measurable sediment wash-off amounts from the surface. For example, at 2% slope, even the rainfall intensity of 155 mm/hr produced only 6g wash-off total wash-off at the end of 60 min. Although a smaller sediment size (and possibly smoother surface, refer Al Ali *et al.*, 2017) could have produced more sediment wash-off, as mentioned earlier this sediment size was chosen based on previous findings from field case studies. In addition to sediment size and surface roughness, surface size also a deciding factor in the amount of washed off sediment as the larger surface will have a proportionally higher initial sediment load. On the other hand, unlike sediment size and surface roughness, surface size does not affect the underlying physical process and as a result, the wash-off fraction (= washed off load/initial load) will remain same. This provides the flexibility in choosing the surface size for similar wash-off experiments. The small surface size such as the one used in this PhD (1×1 m^2) provides a degree of flexibility to change the experiment conditions (e.g. surface slope, initial load) and makes it possible to run such a large number of experiments. Also, it helps to keep the

rainfall intensity fairly uniform over the surface. Similar sized experimental surfaces have been used in recent studies to take advantage of the above-mentioned points (Egodawatta *et al.*, 2007; Al Ali *et al.*, 2017). However, the trade-off is the physically lesser amount of washed off sediment from the surface and consequently the limitation in testing very mild rainfall conditions in these experiments. Hence, an optimal surface size needs to be chosen in future studies which take into account the flexibilities in the experimental setup and the minimum rainfall intensity that can produce a physically measurable sediment wash-off with limited measurement error.

Rainfall intensities used in these experiments are also comparable to rainfall intensities used in similar previous wash-off studies. For example, Egodawatta *et al.*, (2007) used a rainfall intensity range of 40 mm/hr - 133 mm/hr and 20 mm/hr - 133 mm/hr in their experiments to study the wash-off behaviour. Recently Al Ali *et al.*, (2017) used a constant rainfall intensity of 120 mm/hr in similar experimental settings to study the wash-off behaviour from different surfaces. One of the reasons why such experimental rainfall intensities are widely used is that the pattern of experimental observations indicate that the underlying physical transport process of wash-off are the same within the rainfall intensity range as it includes well-developed transport. Due to the practical difficulty in covering a large range of rainfall intensity in an experimental program, extrapolation of the equation/model outside the experimental conditions is often used. Even the OEM was originally developed for much narrower intensity range of 8 mm/hr – 20 mm/hr (Sartor and Boyd, 1972) and has been used widely for rainfall intensities that are well outside this range. Similarly, in this PhD, the experimental rainfall intensities based on which NEM was developed were much higher than the rainfall intensity collected during a limited period from the urban catchment in Bradford. Hence, NEM was applied to rainfall intensities that were well outside the experimental rainfall intensities. In addition, NEM was also used for a synthetic rainfall event to present the effect of 'first-flush' effect. This event was also well outside the calibrated rainfall intensity range. Although the application of the wash-off models in extrapolated rainfall conditions has been a common practice in the past for the reasons mentioned before, this assumption should be verified in future studies. With such verification, the evidence of the finding in this PhD regards to the rainfall uncertainty propagation in wash-off modelling would have been stronger. However, there was no available data from field studies to verify the performance of NEM outside the experimental intensity range as such verification data will also need information on a

surface slope. In previous literature, surface slope data has neither been collected together with wash-off data nor included in the analysis. Hence, verification of NEM outside the intensity range used in the current experiments was not possible and presently, this is a limitation of this study.

Another challenge in the experimental set up in terms of replicating the real rainfall event is the variability within a rainfall event. NEM was developed based on experimental results which were obtained from rainfall events with constant rainfall intensity throughout the duration of an event. Keeping the rainfall intensity constant makes it easier to understand the physical wash-off process and to consequently modify the wash-off model. In fact, most of the previous studies used a constant intensity rainfall event to understand the wash-off process and consequently apply the results to develop and improve the wash-off equations. These studies include Sartor and Boyd (1972) where the exponential model was originally proposed and Egodawatta et al. (2007) where the capacity factor was first introduced in the exponential wash-off. However, constant intensity rainfall events are never the case in reality. Nevertheless, equation proposed by Sartor and Boyd (1972) and consequent refined version (e.g. Egodawatta et al., 2007) were all shown to be applicable for real case studies too. In this regard, NEM also needs to be checked against temporal data collected in wash-off events resulted from real-time varying rainfall events.

Figure 6.1 shows the change in the maximum fraction against the initial load, rainfall intensity and slope so as to provide a quantification of how each of these variables affects the maximum wash-off fraction after an hour. From the figure, it is clear that the sensitivity of wash-off fraction against a variable is not constant and it depends on the remaining variables. For instance, at 2% slope, the sensitivity of maximum wash-off fraction against rainfall intensity is much smaller than the cases when the slope is > 4%. Similarly, the sensitivity of maximum wash-off fraction to slope change from 8% to 16% is comparably smaller than to other slope changes (i.e. 2% - 4% and 4%-8%). It can also be seen that except in extreme scenarios (when rainfall intensity exceeds more than 110 mm/hr and surface slopes exceeds more than 8%), the maximum change on wash-off fraction affected by the initial load is smaller than 5% most of the time. NEM was developed without the slope 16% to avoid extreme scenarios. From the rest of the results only when the slope is more than 8% and the rainfall intensity is more than 110 mm/hr, the effect of the initial load becomes significant. These are considered

extreme scenarios. The inclusion of initial load in the NEM just to account for these extreme conditions would introduce more parameters which makes the equation more complicated without contributing to an equal improvement in the prediction of wash-off for most practical conditions. Hence, considering the added complexity initial load would introduce, it was not considered in the derivation of NEM. Further, as discussed before, conceptually the fraction of sediment washed off during a rainfall event is dependent on the particle size distribution of the initial load. This particle size distribution does not depend on the amount of sediment Hence, the washed off fraction expected to be the same regardless of the initial load within the common catchment and rainfall conditions as seen from most of the experiments.
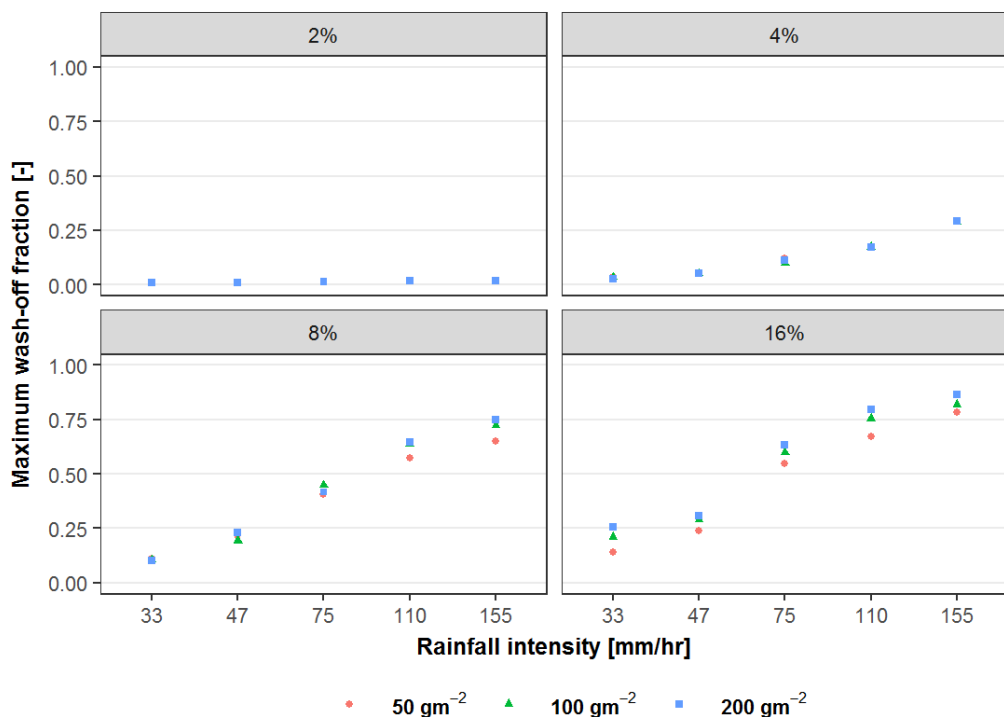
Figure 6.1: Variation in the maximum wash-off fraction against the initial load, rainfall intensity and slope

The above observation also means that the actual mass of sediment washed off at any given time (w_t) is proportional to initial load for a given rainfall intensity and surface slope. Hence the prediction of build-up is perhaps the most preferred way to subsequently predict wash-off compared to the methods presented in recent studies (e.g. Shaw et al. 2010). On the other hand, as Shaw et al. (2010) correctly pointed out, it is a challenging task to model the build-up process due to unpredicted occurrences of activities like construction work or the input of vegetative debris from wind storms. Despite these challenges, the strong correlation observed between build-up load and

wash-off load indicates the importance of modelling the build-up process. This observation does not necessarily invalidate the criticisms on the build-up model of Sartor and Boyd (1972) by Charbeneau and Barrett (1998), Shaw et al. (2010) and He et al. (2010) as their criticism is mainly on the use of antecedent dry days as the main parameter controlling the build-up process. Rather our finding recommends for more attention to be paid on modelling of build-up process accurately, taking more parameters (Wijesiri et al. 2015; Morgan et al. 2017) into consideration in addition to antecedent dry days.

In the improved NEM, $c_{1,....,}$ c_4 are all constants, unlike the OEM where the parameter k has to be calibrated for each case separately. This is the important advantage of NEM over OEM. Hence, although the structure of the NEM looks more complicated than OEM, it avoids the need of any lookup table/plot to estimate the value of the calibration parameter as the external drivers- rainfall intensity and surface slope- are now explicitly included in the NEM as functions. This also introduces some transparency in the parameter estimation which is otherwise a black box approach. Further, replacing the invariant calibration parameters with functions of external drivers (i.e. rainfall intensity and surface slope) makes it easier to investigate the propagation of errors in the external drivers (e.g. rainfall intensity) as these external drivers are now explicitly defined in the new equation. In summary, as NEM tries to incorporate the form of the physical processes, it will always be superior to a simple parsimonious model which in fact merges all physical processes together and so neglects interactions between the external drivers as the catchment conditions change. However, it should be noted that optimal values of $c_1,..,c_4$ needs to be checked against different sediment sizes and different surface roughness as these are two other major external drivers which would affect the sediment wash-off. Any variation in particle size distribution (sorting during an event) will have a direct effect on these parameters. Similarly, mobilisation capacity of a rainfall and runoff will also be expected to change with surface roughness. Hence, these values will be different for a surface with different roughness. Nevertheless, the model structure of NEM would be applicable for any sediment size and surface texture as the underlying physical processes will be the same as those on which the model structure of NEM was developed.

Bayesian inference is more useful when more prior information is available. Unfortunately, the information available on parameter k in literature is not consistent

125

and a wide range of values have been used in the past. This is one of the motivation why it was intended to replace the calibration parameter with external drivers in the NEM. Due to the inconsistency in k values in previous studies, a uniform distribution was chosen for c_4. Further, as stated in chapter 5, since, the other parameters, c_1, c_2 and c_3 are all new and there was no previous literature on these parameters. Hence, uniform distributions were used for other parameters too. This limits the usefulness of Bayesian inference. A more informative prior in the form of normal distribution or truncated normal distribution would have enhanced the usefulness of Bayesian inference and in such case, the uncertainty in the posterior distribution of the parameters is expected to be smaller.   However, the Bayesian inference is capable of using any prior information available including uniform distributions and a number of previous studies have utilised this strategy (e.g. Dotto et al., 2012; Freni and Mannina, 2010). In addition, Bayesian inference also helps to capture the dependence structure which is also one of the other reasons why Bayesian inference was used in this study. Further, it helps to predict different sources of uncertainty as demonstrated in chapter 5 and the uncertainty estimation predicted using Bayesian inference were found to be reasonably accurate (85% accuracy). Optioned values of $c_1,..,c_4$ from this study can be used as prior information in the future and this way a more informative prior can be obtained. This leads to the next discussion point: the inter-correlation between the parameters $c_{1,..,}c_4$. it was observed that some of the parameters are correlated (c_3 against c_1 and c_2). Bayesian inference helps resolve such identifiability issues by allowing for informative priors. Therefore, for real cases, where we have reasons to believe that one of the two parameters should be more constrained, the other parameter value will automatically come out to be constrained after joint inference. As stated already, this is the time first time some of the parameters are introduced and therefore prior information is limited. This is one of the limitations of the model now, however, when more information on these parameters becomes available, the identifiability problem can be solved. Further, a strong correlation between two parameters also reduces the added value of one of those parameters. If the relationship between these parameters is known, one of the parameters can be reduced. Again, due to lack of prior information on any of these parameters, obtaining such relationship in advance was not possible. However, this can be possible in the future when more information becomes available on these parameters.

In this PhD, the uncertainty in the rainfall intensity which is the main hydrological driver of the wash-off process has been investigated. In addition to uncertainty from rainfall variability, other parameters such as sediment size and surface texture are also expected to contribute to the uncertainty in the prediction of wash-off considering their highly variable nature. However, these parameters have not been explicitly included in any wash-off models yet. Hence, prior to investigating the propagation of uncertainty due to these parameters, first and foremost, these parameters need to be included in the wash-off models either explicitly or at least in the form of calibration parameter(s). In addition, the complex interaction between sediment size and surface texture also need to be taken into account when these parameters are investigated. This calls for more integrated experimental studies similar to the one employed in this PhD which is perhaps the only way to investigate these complex interactions.

6.3 Conclusions

6.3.1 On uncertainty in areal rainfall estimation due to sub-kilometre rainfall variability and measurement and sampling error

Spatial variability of rainfall at a sub-kilometre scale was found to be intensity-dependent in the case of the Bradford rainfall data. Hence the assumption of constant spatial variability across intensity ranges, which is a commonly found assumption in previous studies (e.g. Berne et al. 2004; Krajewski et al. 2003), was found to be invalid in the case of the Bradford rainfall data. Hence such assumptions need to be validated in future studies. The uncertainty in the upscaling of rainfall data using point measurements essentially comes from two sources: spatial variability of the rainfall and measurement error. The significance and characteristics of the measurement error observed here mainly corresponds to sampling related error of tipping bucket type rain gauges (TB error) and may vary for other types of rain gauges. TB error at averaging intervals of less than 5min, especially at low-intensity rainfall measurements, is as significant as the spatial variability. Hence, proper attention to TB error should be given in any application of these measurements, especially in urban hydrology, where averaging intervals are often as small as 2min. At smaller temporal averaging intervals, the effect of both spatial variability and TB error is high, resulting in higher uncertainty levels in the areal rainfall estimation, up to 13% at 2 min temporal averaging intervals.

With increasing temporal averaging interval, the uncertainty becomes smaller, i.e. up to 3.6% at 30 min, as the spatial correlation increases and the TB error reduces.

6.3.2 On adapting a geo-statistical method to rainfall upscaling

Although the spatial stochastic simulation method used in this study needs more computational power (a summary on computation power is presented in the Appendix 3B) than block kriging, it is a robust approach and allows data transformation during spatial interpolation and aggregation. Such data transformation is important because rainfall data are not normally distributed especially for small temporal averaging intervals as was shown in this study. Further, the pooling procedure used in this study helps provide a solution to meet the data requirements for geostatistical methods as it extends the available information for variogram estimation.

6.3.3 On improving the understanding of wash-off process

The wash-off load was found to be proportional to initial load irrespective of rainfall intensity and surface slope. Hence, a constant initial load model suggested by Shaw et al. (2010) is not a valid assumption for all catchment and rainfall conditions. Therefore, a dedicated modelling approach to predict build-up to help subsequently predicted wash-off, despite the challenges mentioned in Shaw et al. (2010) should not be overlooked. Furthermore, a rainfall event has the capacity to wash-off only a fraction of initial load, represented by the capacity factor, and once that fraction is washed off from the surface there is no more wash-off even if the rainfall event continues. The maximum fraction that can be washed off from the surface increases with both rainfall intensity and the surface slope.

The effect of the interaction of different variables in the wash-off process was clearly observed. For instance, for a rainfall intensity of 75 mm/hr the wash-off fraction after an hour is only 0.13 when the surface slope is 4%, but this increased 0.42 when the slope was changed to 8%. This clearly indicates the advantages of integrated experimental studies which investigates the effect of multiple variables together over isolated studies where only one variable is investigated (e.g. Egodawatta et al. 2007) in developing better understanding of the wash-off process.

6.3.4 On new exponential wash-off model

The calibration parameters of the exponential model (Eq. 5.2), capacity factor and wash-off coefficient, are both sensitive to the surface slope and rainfall intensity but

are not sensitive to initial load. Hence, in the new exponential model, these calibration parameters have been replaced with functions of rainfall intensity and slope. Although during the calibration stage, original exponential model performs better than the new exponential model, it has to be taken into account that the original exponential model has to be calibrated for each and every experimental condition separately. Further, at the validation stage, the new exponential model performance improved over the original exponential model, reflecting the ability of the new exponential model to perform better under new catchment conditions.

6.3.5 On uncertainty associated with new improved wash-off model

Verification measures show the uncertainty estimates associated with the new improved wash-off model predictions were plausible, indicating that the autoregressive error model accounting for the structural deficit with constant bias accounting for measurement noise was a reasonable representation of the error process associated with sediment wash-off modelling. The 95% predictive uncertainty shows a maximum of 10% variability in the prediction of a total wash-off fraction out of which 6% comes from parameter uncertainty and the remainder came from the structural deficit and measurement noise.

6.3.6 On the effect of rainfall uncertainty in wash-off prediction

The level of uncertainty in the prediction of wash-off load due to rainfall uncertainty can be smaller, similar or higher to the level of uncertainty in rainfall depending on the intensity range and the "first-flush" effect as explained in detail below. When the rainfall intensities are too small to produce a proportional change in the instantaneous wash-off against the corresponding change in the rainfall intensity, a certain level of uncertainty in rainfall intensities would cause a lesser level of uncertainty in wash-off fraction. When rainfall intensities are bigger, but still not big enough to produce a "first-flush" effect then the change in the instantaneous wash-off is proportional to change in the rainfall intensity, hence any uncertainty in rainfall estimation would produce a similar level of uncertainty in the prediction of instantaneous wash-off. When the rainfall intensities are big enough to produce a "first-flush", then most of the sediment will be washed off at the beginning of the event and corresponding instantaneous wash-off will be larger than previous two cases, therefore a certain level of uncertainty in a rainfall intensity would produce a larger level of uncertainty in corresponding predicted instantaneous wash-off at the beginning of an event i.e. during the "first-flush".

None of the measured rainfall events from Chapter 3 produced a "first flush" effect, hence for most common rainfall and surface slope conditions, it was found that any uncertainty in rainfall estimation would produce a similar level of uncertainty in the prediction of instantaneous wash-off. Consequently, the maximum uncertainty in the peak instantaneous wash-off due to measurement and sampling error and spatial variability of rainfall within an 8 ha catchment in Bradford found to be 13% when a temporal averaging interval of 2 min is considered. Although at 30 min temporal averaging interval, maximum uncertainty in the peak instantaneous wash-off fraction was reduced to 4%, the temporal aggregation, on the other hand, reduced the prediction of peak instantaneous wash-off up to 80%.

These figures demand that both spatial and temporal variability of rainfall at sub-kilometre scale together with measurement and sampling error needs to be taken into account in the prediction of wash-off.

Appendix

Written permission from Faculty of Engineering to submit this thesis in an alternative format is attached below

REQUEST TO SUBMIT AN ALTERNATIVE FORMAT THESIS

Students should use this form to request permission to submit a doctoral thesis which contains sections which are in a format suitable for submission for publication in a peer-reviewed journal (or other appropriate outlet for academic research), alongside more traditional thesis chapters. This avoids the need for students to spend time rewriting publications into thesis chapters, as well as enhancing their writing for publication skills.

Requests to submit an alternative format thesis must be supported by the student's supervisory team and also require approval from the relevant Faculty and must therefore be made well in advance of the student submitting their thesis and no later than the end of the student's 3rd year of study.

Students wishing to submit in this format must read the guidance on structure of an alternative format thesis at: http://www.sheffield.ac.uk/ris/pgr/code/altformat

SECTION 1: STUDENT'S DETAILS	
NAME: Manoranjan Muthusamy	REGISTRATION NUMBER: 140260575
DEPARTMENT: Civil and Structural Engineering	FACULTY: Engineering Faculty

PROPOSED THESIS TITLE:
Effect of rainfall variability on urban sediment behaviour

REASON FOR THE REQUEST: Please use this space to provide details of the published papers that will be incorporated into your thesis and how your thesis will be structured (continue on a separate sheet, if required)

This PhD thesis will consist of the following three main chapters in addition to Introduction, Literature review and Conclusion

1. Quantifying small scale spatial variability of rainfall with regards to urban hydrological modelling
2. Development of an improved sediment wash-off model for urban impervious surfaces
3. Estimation of uncertainty and rainfall error propagation in sediment wash-off modelling

I request to use my published/submitted peer viewed journal papers for the above three chapters. The details of these papers are listed below in the same order. Note that I am

the first author of all the papers.

1. Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models
   Journal: Hydrology and Earth System Sciences (HESS)
   Status: Published
   Reference: Muthusamy, M., Schellart, A., Tait, S., and Heuvelink, G. B. M.: Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models, Hydrol. Earth Syst. Sci., 21, 1077-1091, https://doi.org/10.5194/hess-21-1077-2017 , 2017.

2. Improving the understanding of the underlying physical process of sediment wash-off from urban road surfaces
   Journal: Journal of hydrology
   Status: Under review
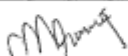
3. Uncertainty analyses and rainfall error propagation through a "global" wash-off model
   Journal (Target): Water Research
   Status: Under preparation. Target submission date – End of September, 2017 (Month 33 from the starting date of the PhD)

The decision to incorporate the above paper into the thesis is taken considering the significant amount of time spent on the preparation of the above papers and to reduce the time spent on rewriting. How the constituent papers form a coherent body of work of the thesis will be explained well in the introduction chapter.

SIGNATURE:	DATE: 27/07/2017

SECTION 2: SUPERVISOR'S COMMENTS

NAME: Alma Schellart	DEPARTMENT: Civil and Structural Engineering

COMMENTS IN SUPPORT OF THE REQUEST:
I support Mano's request as the contents of the 3 proposed papers logically follow through the main body of work for this PhD research, and time otherwise spent in re-writing these papers as chapters can be saved. One paper has already been published, and the second one is under review, I have the confidence that the 3rd paper will also be

132

under review by the time the thesis is submitted. The proposed structure of introduction, literature review, then the three papers, followed by an over-arching conclusions chapter would form a coherent body of work.

SIGNATURE:	DATE: 4/8/2017
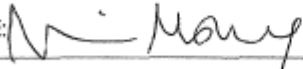
Please return this completed form to one of the following email addresses, depending on your Faculty: Arts & Humanities - pgrarts@sheffield.ac.uk; Engineering - pgreng@sheffield.ac.uk; Medicine, Dentistry & Health - pgrmdh@sheffield.ac.uk; Science - pgrsci@sheffield.ac.uk; Social Sciences - pgrsocsci@sheffield.ac.uk

SECTION 3: FACULTY CONSIDERATION

NAME:	FACULTY:

COMMENTS ON DECISION:

☑	I approve this request for submission of an alternative format thesis	☐	I do not approve this request for submission of an alternative format thesis

SIGNATURE:	DATE: 1 0 AUG 2017

References

Akan, A. O. (1987) 'Pollutant Washoff by Overland Flow', *Journal of Environmental Engineering*. American Society of Civil Engineers, 113(4), pp. 811–823. doi: 10.1061/(ASCE)0733-9372(1987)113:4(811).

Al, S. E. T. and Elson, P. E. A. N. (2005) 'Extraordinary flood response of a small urban watershed', *Bulletin of the American Meteorological Society*, 86(12), pp. 1730–1732. doi: 10.1175/JHM426.1.

Alley, W. M. (1981) 'Estimation of Impervious-Area Washoff Parameters', *Water Resources Research*, 17(4), pp. 1161–1166.

Alley, W. M. and Smith, P. E. (1981) 'Estimation of accumulation parameters for urban runoff quality modeling', *Water Resources Research*, 17(6), pp. 1657–1664. doi: 10.1029/WR017i006p01657.

Ammon, D. C. (1979) 'Urban Stormwater Pollutant Buildup and Washoff Relationships', *Master thesis, Dept. of Environmental Engineering Sciences*. Gainesville, FL.

Ashley, R. M., Wotherspoon, D. J. J., Coghlan, B. P. and McGregor, I. (1992) 'The Erosion and Movement of Sediments and Associated Pollutants in Combined Sewers', *Water Science and Technology*, 25(8), p. 101 LP-114. Available at: http://wst.iwaponline.com/content/25/8/101.abstract.

Baecher, G. B. and Christian, J. T. (2003) *Reliability and Statistics in Geotechnical Engineering*. Wiley. Available at: https://books.google.co.uk/books?id=42r6CKatuwcC.

Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., *et al.* (2007) 'A Framework for Validation of Computer Models', *Technometrics*. Taylor & Francis, 49(2), pp. 138–154. doi: 10.1198/004017007000000092.

Beck, M. B. (1987) 'Water quality modeling: A review of the analysis of uncertainty', *Water Resources Research*, 23(8), pp. 1393–1442. doi: 10.1029/WR023i008p01393.

van de Beek, C. Z., Leijnse, H., Stricker, J. N. M., Uijlenhoet, R., *et al.* (2010) 'Performance of high-resolution X-band radar for rainfall measurement in The Netherlands', *Hydrology and Earth System Sciences*, 14(2), pp. 205–221. doi: 10.5194/hess-14-205-2010.

Berne, A., Delrieu, G., Creutin, J. D. and Obled, C. (2004) 'Temporal and spatial resolution of rainfall measurements required for urban hydrology', *Journal of Hydrology*, 299(3–4), pp. 166–179. doi: 10.1016/j.jhydrol.2004.08.002.

Berne, A. and Krajewski, W. F. (2013) 'Radar for hydrology: Unfulfilled promise or unrecognized potential?', *Advances in Water Resources*. Elsevier Ltd, 51, pp. 357–366. doi: 10.1016/j.advwatres.2012.05.005.

Bertrand-Krajewski, J.-L. (2006) 'Modelling of Sewer Solids Production and Transport. Cours de DEA "Hydrologie Urbaine"', *Transport*, p. 39.

Bertrand-Krajewski, J.-L., Briat, P. and Scrivener, O. (1993) 'Sewer sediment production and transport modelling: A literature review', *Journal of Hydraulic*

Research. Taylor & Francis, 31(4), pp. 435–460. doi: 10.1080/00221689309498869.

Bertrand-Krajewski, J.-L., Chebbo, G. and Saget, A. (1998) 'Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon', *Water Research*, 32(8), pp. 2341–2356. doi: 10.1016/S0043-1354(97)00420-X.

Bogner, K., Pappenberger, F. and Cloke, H. L. (2012) 'Technical Note: The normal quantile transformation and its application in a flood forecasting system', *Hydrology and Earth System Sciences*, 16(4), pp. 1085–1094. doi: 10.5194/hess-16-1085-2012.

Bonhomme, C. and Petrucci, G. (2017) 'Should we trust build-up/wash-off water quality models at the scale of urban catchments?', *Water Research*, 108, pp. 422–431. doi: 10.1016/j.watres.2016.11.027.

Brauer, C. C., Teuling, A. J., Overeem, A., van der Velde, Y., *et al.* (2011) 'Anatomy of extraordinary rainfall and flash flood in a Dutch lowland catchment', *Hydrology and Earth System Sciences*, 15(6), pp. 1991–2005. doi: 10.5194/hess-15-1991-2011.

Breinholt, A., Møller, J. K., Madsen, H. and Mikkelsen, P. S. (2012) 'A formal statistical approach to representing uncertainty in rainfall–runoff modelling with focus on residual analysis and probabilistic output evaluation – Distinguishing simulation and prediction', *Journal of Hydrology*, 472(Supplement C), pp. 36–52. doi: https://doi.org/10.1016/j.jhydrol.2012.09.014.

Brodie, I. M. (2007) 'Prediction of stormwater particle loads from impervious urban surfaces based on a rainfall detachment index', *Water Science and Technology*, 55(4), p. 49 LP-56. Available at: http://wst.iwaponline.com/content/55/4/49.abstract.

Bruni, G., Reinoso, R., Van De Giesen, N. C., Clemens, F. H. L. R., *et al.* (2015) 'On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution', *Hydrology and Earth System Sciences*, 19(2), pp. 691–709. doi: 10.5194/hess-19-691-2015.

Butler, D. and Clark, P. (1995) 'Sediment management in urban drainage catchments', *Report, Construction Industry Research and Information Association (CIRIA)*. London, UK.

Carvalho, S. C. P., de Lima, J. L. M. P. and de Lima, M. I. P. (2014) 'Using meshes to change the characteristics of simulated rainfall produced by spray nozzles', *International Soil and Water Conservation Research*, 2(2), pp. 67–78. doi: 10.1016/S2095-6339(15)30007-1.

Cecinati, F., Wani, O. and Rico-Ramirez, M. A. (2017) 'Comparing Approaches to Deal with Non-Gaussianity of Rainfall Data in Kriging-Based Radar-Gauge Rainfall Merging', *Water Resources Research*. doi: 10.1002/2016WR020330.

Characklis, A. W. G., Bedient, P. B. and Roe, F. (1979) 'A study of stormwater runoff quality', 699(1978), pp. 673–699.

Charbeneau, R. J. and Barrett, M. E. (1998) 'Evaluation of methods for estimating stormwater pollutant loads', *Water Environment Research*, 70(7), pp. 1295–1302. doi: 10.2175/106143098X123679.

Chiew, F. H. S. and Vaze, J. (2004) 'Nutrient Loads Associated with Different

Sediment Sizes in Urban Stormwater and Surface Pollutants', *Journal of Environmental Engineering*. American Society of Civil Engineers, 130(4), pp. 391–396. doi: 10.1061/(ASCE)0733-9372(2004)130:4(391).

Chow, V. Te (1964) *Handbook of Applied Hydrology. A compendium of water-resources technology*. New York: McGraw-Hill.

Christakos, G. (1984) 'On the Problem of Permissible Covariance and Variogram Models', *Water Resources Research*, 20(2), pp. 251–265. doi: 10.1029/WR020i002p00251.

Ciach, G. J. and Krajewski, W. F. (2006) 'Analysis and modeling of spatial correlation structure in small-scale rainfall in Central Oklahoma', *Advances in Water Resources*, 29, pp. 1450–1463. doi: 10.1016/j.advwatres.2005.11.003.

Cole, S. J. and Moore, R. J. (2008) 'Hydrological modelling using raingauge- and radar-based estimators of areal rainfall', *Journal of Hydrology*, 358(3), pp. 159–181. doi: https://doi.org/10.1016/j.jhydrol.2008.05.025.

Coleman, T. J. (1993) 'A comparison of the modelling of suspended solids using SWMM3 quality prediction algorithms with a model based on sediment transport theory.', in *6th Int. Conf. on Urban Storm Drainage*. Reston, VA: ASCE.

Collins, P. G. and Ridgeway, J. W. (1980) 'Urban storm runoff quality in southeast Michigan', *Journal of the Environmental Engineering Division*, 106(EEl), pp. 153–162.

Craig, P. S., Goldstein, M., Rougier, J. C. and Seheult, A. H. (2001) 'Bayesian Forecasting for Complex Systems Using Computer Simulators', *Journal of the American Statistical Association*. [American Statistical Association, Taylor & Francis, Ltd.], 96(454), pp. 717–729. Available at: http://www.jstor.org/stable/2670309.

Cressie, N. (1993) *Statistics for spatial data*. J. Wiley (Wiley series in probability and mathematical statistics: Applied probability and statistics). Available at: https://books.google.co.uk/books?id=4SdRAAAAMAAJ.

Cressie, N. and Hawkins, D. M. (1980) 'Robust estimation of the variogram: I', *Journal of the International Association for Mathematical Geology*, 12(2), pp. 115–125. doi: 10.1007/BF01035243.

Cristiano, E., ten Veldhius, M. and van de Giesen, N. (2016) 'Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas - &amp;ndash; a review', *Hydrology and Earth System Sciences Discussions*, (October), pp. 1–34. doi: 10.5194/hess-2016-538.

Dawson, C. and Gerritsen, M. (2013) *Computational Challenges in the Geosciences*. Springer New York (The IMA Volumes in Mathematics and its Applications). Available at: https://books.google.co.uk/books?id=kk26BAAAQBAJ.

Debo, T. N. and Reese, A. (1995) *Municipal Stormwater Management, Second Edition*. CRC Press. Available at: https://books.google.co.uk/books?id=HuPYQi83j-QC.

Deletic, A., Maksimovic, C. and Ivetic, M. (1997) 'Modelling of storm wash-off of suspended solids from impervious surfaces', *Journal of Hydraulic Research*, 35(1), pp. 99–118. doi: 10.1080/00221689709498646.

Delleur, J. (2001) 'New Results and Research Needs on Sediment Movement in Urban Drainage', *Journal of Water Resources Planning and Management*. American Society of Civil Engineers, 127(3), pp. 186–193. doi: 10.1061/(ASCE)0733-9496(2001)127:3(186).

Dingman, S. L. (2015) *Physical Hydrology: Third Edition*. Waveland Press. Available at: https://books.google.co.uk/books?id=rUUaBgAAQBAJ.

Dirks, K. N., Hay, J. E., Stow, C. D. and Harris, D. (1998) 'High-resolution studies of rainfall on Norfolk Island Part II: Interpolation of rainfall data', *Journal of Hydrology*, 208(3–4), pp. 187–193. doi: 10.1016/S0022-1694(98)00155-3.

Dotto, C. B. S., Kleidorfer, M., Deletic, A., Rauch, W., *et al.* (2011) 'Performance and sensitivity analysis of stormwater models using a Bayesian approach and long-term high resolution data', *Environmental Modelling & Software*, 26(10), pp. 1225–1239. doi: https://doi.org/10.1016/j.envsoft.2011.03.013.

Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., *et al.* (2012) 'Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling', *Water Research*, 46(8), pp. 2545–2558. doi: 10.1016/j.watres.2012.02.009.

Duncan, H. P. (1995) 'A review of urban storm water quality processes', *Report-Cooperative Research Centre for Catchment Hydrology, Melbourne, Australia.*

Egodawatta, P. and Goonetilleke, A. (2008) 'Understanding road surface pollutant wash-off and underlying physical processes using simulated rainfall', *Water Science and Technology*, 57(8), pp. 1241–1246. doi: 10.2166/wst.2008.260.

Egodawatta, P., Haddad, K., Rahman, A. and Goonetilleke, A. (2014) 'A Bayesian regression approach to assess uncertainty in pollutant wash-off modelling', *Science of The Total Environment*, 479–480(Supplement C), pp. 233–240. doi: https://doi.org/10.1016/j.scitotenv.2014.02.012.

Egodawatta, P., Thomas, E. and Goonetilleke, A. (2007) 'Mathematical interpretation of pollutant wash-off from urban road surfaces using simulated rainfall', *Water Research*, 41(13), pp. 3025–3031. doi: 10.1016/j.watres.2007.03.037.

Emmanuel, I., Andrieu, H., Leblois, E. and Flahaut, B. (2012) 'Temporal and spatial variability of rainfall at the urban hydrological scale', *Journal of Hydrology*. Elsevier B.V., 430–431, pp. 162–172. doi: 10.1016/j.jhydrol.2012.02.013.

Fankhauser, R. (1998) 'Influence of systematic errors from tipping bucket rain gauges on recorded rainfall data', *Water Science and Technology*. No longer published by Elsevier, 37(11), pp. 121–129. doi: 10.1016/S0273-1223(98)00324-2.

Farreny, R., Morales-Pinzón, T., Guisasola, A., Tayà, C., *et al.* (2011) 'Roof selection for rainwater harvesting: Quantity and quality assessments in Spain', *Water Research*, 45(10), pp. 3245–3254. doi: 10.1016/j.watres.2011.03.036.

Fiener, P. and Auerswald, K. (2009) 'Spatial variability of rainfall on a sub-kilometre scale', *Earth Surface Processes and Landforms*. John Wiley & Sons, Ltd., 34(6), pp. 848–859. doi: 10.1002/esp.1779.

Francey, M., Duncan, H. P., Deletic, a. and Fletcher, T. D. (2011) 'Testing and Sensitivity of a Simple Method for Predicting Urban Pollutant Loads', *Journal of*

Environmental Engineering, 137(9), pp. 782–789. doi: 10.1061/(ASCE)EE.1943-7870.0000386.

Freni, G. and Mannina, G. (2010) 'Bayesian approach for uncertainty quantification in water quality modelling: The influence of prior distribution', *Journal of Hydrology*, 392(1), pp. 31–39. doi: https://doi.org/10.1016/j.jhydrol.2010.07.043.

Freni, G., Mannina, G. and Viviani, G. (2009) 'Uncertainty assessment of an integrated urban drainage model', *Journal of Hydrology*, 373(3), pp. 392–404. doi: https://doi.org/10.1016/j.jhydrol.2009.04.037.

Gebremichael, M. and Krajewski, W. F. (2004) 'Assessment of the statistical characterization of small-scale rainfall variability from radar: Analysis of TRMM ground validation datasets', *Journal of Applied Meteorology*, 43(8), pp. 1180–1199. doi: 10.1175/1520-0450(2004)043<1180:AOTSCO>2.0.CO;2.

Gebremichael, M. and Krajewski, W. F. (2004) 'Assessment of the Statistical Characterization of Small-Scale Rainfall Variability from Radar: Analysis of TRMM Ground Validation Datasets', *Journal of Applied Meteorology*. American Meteorological Society, 43(8), pp. 1180–1199. doi: 10.1175/1520-0450(2004)043<1180:AOTSCO>2.0.CO;2.

Genton, M. G. (1998) 'Highly Robust Variogram Estimation', *Mathematical Geology*, 30(2), pp. 213–221. doi: 10.1023/A:1021728614555.

Gires, A., Onof, C., Maksimovic, C., Schertzer, D., *et al.* (2012) 'Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study', *Journal of Hydrology*. Elsevier B.V., 442–443, pp. 117–128. doi: 10.1016/j.jhydrol.2012.04.005.

Gires, A., Tchiguirinskaia, I., Schertzer, D., Schellart, A., *et al.* (2014) 'Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data', *Atmospheric Research*. Elsevier B.V., 138, pp. 125–138. doi: 10.1016/j.atmosres.2013.11.008.

Del Giudice, D., Honti, M., Scheidegger, A., Albert, C., *et al.* (2013) 'Improving uncertainty estimation in urban hydrological modeling by statistically describing bias', *Hydrology and Earth System Sciences*, 17(10), pp. 4209–4225. doi: 10.5194/hess-17-4209-2013.

Glasbey, C. A. and Nevison, I. M. (1997) 'Modelling Longitudinal and Spatially Correlated Data', in Gregoire, T. G. et al. (eds). New York, NY: Springer New York, pp. 233–242. doi: 10.1007/978-1-4612-0699-6_20.

Goovaerts, P. (2000) 'Estimation or simulation of soil properties? An optimization problem with conflicting criteria', *Geoderma*, 97(3–4), pp. 165–186. doi: 10.1016/S0016-7061(00)00037-9.

Gringarten, E. and Deutsch, C. V (2001) 'Teacher's Aide Variogram Interpretation and Modeling', *Mathematical Geology*, 33(4), pp. 507–534. doi: 10.1023/a:1011093014141.

Guy, H. P. (1970) 'Sediment Problems in Urban Areas', *Geological survey circular 601-E, U.S. Geological Survey*. Washington.

Habib, E., Krajewski, W. F. and Kruger, A. (2001) 'Sampling Errors of Tipping-Bucket

Rain Gauge Measurements', *Journal of Hydrologic Engineering*, 6(2), pp. 159–166. doi: 10.1061/(ASCE)1084-0699(2001)6:2(159).

Halls-Moore, M. (no date) *Bayesian Statistics: A Beginner's Guide*. Available at: https://www.quantstart.com/articles/Bayesian-Statistics-A-Beginners-Guide (Accessed: 13 November 2017).

He, J. X., Valeo, C., Chu, a and Neumann, N. F. (2010) 'Characterizing Physicochemical Quality of Storm-Water Runoff from an Urban Area in Calgary, Alberta', *Journal of Environmental Engineering-Asce*, 136(November), pp. 1206–1217. doi: 10.1061/(ASCE)EE.1943-7870.0000267.

Heal, K. V., Hepburn, D. A. and Lunn, R. J. (2006) 'Sediment management in sustainable urban drainage system ponds', *Water Science and Technology*, 53(10), pp. 219–227. doi: 10.2166/wst.2006.315.

Helsel, D. R., Kim, J. I., Grizzard, T. J., Randall, C. W., *et al.* (1979) 'Land Use Influences on Metals in Storm Drainage', *Journal (Water Pollution Control Federation)*. Water Environment Federation, 51(4), pp. 709–717. Available at: http://www.jstor.org/stable/25039892.

Herngren, L., Goonetilleke, A., Sukpum, R. and Silva, D. Y. de (2005) 'Rainfall Simulation as a Tool for Urban Water Quality Research', *Environmental Engineering Science*. Mary Ann Liebert, Inc., publishers, 22(3), pp. 378–383. doi: 10.1089/ees.2005.22.378.

Higdon, D., Kennedy, M., Cavendish, J., Cafeo, J., *et al.* (2004) 'Combining Field Data and Computer Simulations for Calibration and Prediction', *SIAM Journal on Scientific Computing*. Society for Industrial and Applied Mathematics, 26(2), pp. 448–466. doi: 10.1137/S1064827503426693.

Highway Department UK (1989) 'Guidance Notes on Road Testing', *Publication No. RD/GN/009*, 1.

Hirsch, R. M. and Slack, J. R. (1984) 'A Nonparametric Trend Test for Seasonal Data With Serial Dependence', *Water Resources Research*, p. 727. doi: 10.1029/WR020i006p00727.

Huber, W. C. and Dickinson, R. E. (1992) *Storm Water Management Model, Version 4 : User's Manual, Rep. No. EPA/600/3-88/001a*. Athens, Ga.

Isaaks, E. H. and Srivastava, R. M. (1989) *An Introduction to Applied Geostatistics*. New York, USA: Oxford University Press.

Isidoro, J. M. G. P. and Lima, J. L. M. P. de (2013) 'Analytical Closed-Form Solution for 1D Linear Kinematic Overland Flow under Moving Rainstorms', *Journal of Hydrologic Engineering*. American Society of Civil Engineers, 18(9), pp. 1148–1156. doi: 10.1061/(ASCE)HE.1943-5584.0000740.

Ivan, A. (2001) *Guidelines on non-structural measures in urban flood management, IHP-V | Technical Documents in Hydrology | No. 50*. Paris. doi: 10.4000/vertigo.11074.

Jaffrain, J. and Berne, A. (2012) 'Quantification of the small-scale spatial structure of the raindrop size distribution from a network of disdrometers', *Journal of Applied Meteorology and Climatology*, 51(5), pp. 941–953. doi: 10.1175/JAMC-D-11-

0136.1.

Jensen, N. E. and Pedersen, L. (2005) 'Spatial variability of rainfall: Variations within a single radar pixel', *Atmospheric research*. NETHERLANDS: Elsevier, 77(1–4), pp. 269–277. doi: 10.1016/j.atmosres.2004.10.029.

Jewski, W. F. K. R. A., Ciach, G. J. and Habib, E. (2016) 'An analysis of small-scale rainfall variability in different climatic regimes different climatic regimes', 6667(May). doi: 10.1623/hysj.48.2.151.44694.

Journel, A. G. and Huijbregts, C. (1978) *Mining geostatistics*. Academic Press. Available at: https://books.google.co.uk/books?id=w5UZAQAAIAAJ.

Kayhanian, M., Suverkropp, C., Ruby, A. and Tsay, K. (2007) 'Characterization and prediction of highway runoff constituent event mean concentration', *Journal of Environmental Management*, 85(2), pp. 279–295. doi: 10.1016/j.jenvman.2006.09.024.

Kennedy, M. C. and O'Hagan, A. (2001) 'Bayesian calibration of computer models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. Blackwell Publishers Ltd., 63(3), pp. 425–464. doi: 10.1111/1467-9868.00294.

Krajewski, W. F., Ciach, G. J. and Habib, E. (2003) 'An analysis of small-scale rainfall variability in different climatic regimes', *Hydrological Sciences Journal*, 48(2), pp. 151–162. doi: 10.1623/hysj.48.2.151.44694.

Krajewski, W. F., Ciach, G. J., McCollum, J. R. and Bacotiu, C. (2000) 'Initial Validation of the Global Precipitation Climatology Project Monthly Rainfall over the United States', *Journal of Applied Meteorology*. American Meteorological Society, 39(7), pp. 1071–1086. doi: 10.1175/1520-0450(2000)039<1071:IVOTGP>2.0.CO;2.

Lanza, L. G. and Vuerich, E. (2009) 'The WMO Field Intercomparison of Rain Intensity Gauges', *Atmospheric Research*, 94(4), pp. 534–543. doi: https://doi.org/10.1016/j.atmosres.2009.06.012.

Lawler, D. M., Petts, G. E., Foster, I. D. L. and Harper, S. (2006) 'Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK', *Science of the Total Environment*, 360(1–3), pp. 109–126. doi: 10.1016/j.scitotenv.2005.08.032.

de Lima, J. L. M. P., Carvalho, S. C. P. and P. de Lima, M. I. (2013) 'Rainfall simulator experiments on the importance o of when rainfall burst occurs during storm events on runoff and soil loss', *Zeitschrift für Geomorphologie, Supplementary Issues*. E. Schweizerbart'sche Verlagsbuchhandlung, 57(1), pp. 91–109. doi: 10.1127/0372-8854/2012/S-00096.

Lloyd, C. D. (2005) 'Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain', *Journal of Hydrology*, 308(1–4), pp. 128–150. doi: 10.1016/j.jhydrol.2004.10.026.

Ly, S., Charles, C. and Degré, A. (2013) 'Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale : a review', *Biotechnology, Agronomy, Society and Environment (BASE)*, 17(2), pp. 392–406.

Mair, A. and Fares, A. (2011) 'Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island', (April). doi: 10.1061/(ASCE)HE.1943-5584.0000330.

Matheron, G. (1963) 'Principles of geostatistics', *Economic Geology*, 58(8), p. 1246 LP-1266. Available at: http://economicgeology.org/content/58/8/1246.abstract.

MetOffice UK (2016) *Online.* Available at: https://www.metoffice.gov.uk/public/weather/climate-extremes/#?tab=climateExtremes (Accessed: 25 November 2015).

Millar, R. G. (1999) 'Analytical determination of pollutant wash-off parameters', *Journal of Environmental Engineering*, Vol. 125,(No. 10 (Technical Note)), pp. 989–992.

Mitchell, G., Lockyer, J. and McDonald, A. . (2001) 'Pollution Hazard from Urban Nonpoint Sources: A GIS-model to Support Strategic Environmental Planning in the UK', *Technical Report, School of Geography, University of Leeds*, 1,2, p. 240pp.

Montanari, A., & Brath, A. (2004) 'A stochastic approach for assessing the uncertainty of rainfall-runoff simulations', *Water Resources Research*, 40(1), pp. 1–11. doi: 10.1029/2003WR002540.

Montenegro, A. A. A., Abrantes, J. R. C. B., de Lima, J. L. M. P., Singh, V. P., *et al.* (2013) 'Impact of mulching on soil and water dynamics under intermittent simulated rainfall', *CATENA*, 109, pp. 139–149. doi: 10.1016/j.catena.2013.03.018.

Moys, G. D., Osborne, M. P. and Payne, J. A. (1988) 'Mosqito1. Modelling of stormwater quality including tanks and overflows. Design specifications', *Tech. Rep. SR 184 Hydraulics Research Limited, Wallingford, UK*, (August).

Muthusamy, M., Patania, A., Schellart, A. and Tait, S. (2015) 'Analysis of sub-kilometre variability of rainfall in the context of urban runoff modelling', in *10th International Urban Drainage Modelling Conference*, pp. 201–205.

Muthusamy, M., Schellart, A., Tait, S. and Heuvelink, G. B. M. (2017) 'Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models', *Hydrology and Earth System Sciences*, 21(2), pp. 1077–1091. doi: 10.5194/hess-21-1077-2017.

Muthusamy, M., Tait, S., Schellart, A., Beg, M. N. A., *et al.* (2018) 'Improving understanding of the underlying physical process of sediment wash-off from urban road surfaces', *Journal of Hydrology*, 557, pp. 426–433. doi: https://doi.org/10.1016/j.jhydrol.2017.11.047.

Nakamura, E. (1984) 'Factors affecting stormwater quality decay coefficient', in P. Balmer, P. Malmquist, A. S. (ed.) *Proceedings of the Third International Conference on Urban Storm Drainage*. Goteborg, Sweden, pp. 979 – 988.

Nalder, I. a. and Wein, R. W. (1998) 'Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest', *Agricultural and Forest Meteorology*, 92, pp. 211–225. doi: 10.1016/S0168-1923(98)00102-6.

Novotny, V. and Olem, H. (1994) *Water Quality: Prevention, Identification, and Management of Diffuse Pollution.* Wiley. Available at: https://books.google.co.uk/books?id=KqwDAAAACAAJ.

Ochoa-Rodriguez, S., Wang, L. P., Gires, A., Pina, R. D., *et al.* (2015) 'Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation', *Journal of Hydrology*. Elsevier B.V., 531, pp. 389–407. doi: 10.1016/j.jhydrol.2015.05.035.

Overeem, A., Buishand, T. A. and Holleman, I. (2009) 'Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar', *Water Resources Research*, 45(10), p. n/a-n/a. doi: 10.1029/2009WR007869.

Overeem, A., Leijnse, H. and Uijlenhoet, R. (2011) 'Measuring urban rainfall using microwave links from commercial cellular communication networks', *Water Resources Research*, 47(12), p. n/a-n/a. doi: 10.1029/2010WR010350.

Pebesma, E. J. (2004) 'Multivariable geostatistics in S: The gstat package', *Computers and Geosciences*, 30(7), pp. 683–691. doi: 10.1016/j.cageo.2004.03.012.

Pedersen, L., Jensen, N. E., Christensen, L. E. and Madsen, H. (2010) 'Quantification of the spatial variability of rainfall based on a dense network of rain gauges', *Atmospheric Research*. Elsevier B.V., 95(4), pp. 441–454. doi: 10.1016/j.atmosres.2009.11.007.

Peleg, N., Ben-Asher, M. and Morin, E. (2013a) 'Radar subpixel-scale rainfall variability and uncertainty: Lessons learned from observations of a dense rain-gauge network', *Hydrology and Earth System Sciences*, 17(6), pp. 2195–2208. doi: 10.5194/hess-17-2195-2013.

Peleg, N., Ben-Asher, M. and Morin, E. (2013b) 'Radar subpixel-scale rainfall variability and uncertainty: Lessons learned from observations of a dense rain-gauge network', *Hydrology and Earth System Sciences*, 17(6), pp. 2195–2208. doi: 10.5194/hess-17-2195-2013.

Reichert, P. and Schuwirth, N. (2012) 'Linking statistical bias description to multiobjective model calibration', *Water Resources Research*, 48(9). doi: 10.1029/2011WR011391.

Reinertsen, T. R. (1981) 'Quality of stormwater runoff from streets. Urban stormwater quality, management, and planning', *Proc., 2nd Int. Conf. on Urban Storm Drainage, Water Resources Publications, Littleton, CO.*

Sage, J., Bonhomme, C., Berthier, E. and Gromaire, M.-C. (2016) 'Assessing the Effect of Uncertainties in Pollutant Wash-Off Dynamics in Stormwater Source-Control Systems Modeling: Consequences of Using an Inappropriate Error Model', *Journal of Environmental Engineering*, 143(August), pp. 1–9. doi: 10.1061/(ASCE)EE.1943-7870.0001163.

Sansalone, J. J. and Buchberger, S. G. (1997) 'Partitioning and First Flush of Metals in Urban Roadway Storm Water', *Journal of Environmental Engineering*. American Society of Civil Engineers, 123(2), pp. 134–143. doi: 10.1061/(ASCE)0733-9372(1997)123:2(134).

Sartor, J. D. and Boyd, B. G. (1972) 'Water pollution aspects of street surface contaminants.' Washington, D.C., p. EPA Rep. 11024 DOC 07-71, (NTIS PB-203289).

Scheidegger, A. (2017) 'adaptMCMC: Implementation of a Generic Adaptive Monte Carlo Markov Chain Sampler'. Available at: http://cran.r-

project.org/package=adaptMCMC.

Schilling, W. (1991) 'Rainfall data for urban hydrology: what do we need?', *Atmospheric Research*, 27(1–3), pp. 5–21. doi: 10.1016/0169-8095(91)90003-F.

Schuurmans, J. M. and Bierkens, M. F. P. (2006) 'Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model', *Hydrology and Earth System Sciences*, 3(4), pp. 2175–2208. doi: 10.5194/hessd-3-2175-2006.

Segond, M. L., Wheater, H. S. and Onof, C. (2007) 'The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK', *Journal of Hydrology*, 347(1–2), pp. 116–131. doi: 10.1016/j.jhydrol.2007.09.040.

Seo, B.-C. and Krajewski, W. F. (2010) 'Scale Dependence of Radar Rainfall Uncertainty: Initial Evaluation of NEXRAD's New Super-Resolution Data for Hydrologic Applications', *Journal of Hydrometeorology*, 11(5), pp. 1191–1198. doi: 10.1175/2010JHM1265.1.

Sevruk, B. and Hamon, W. R. (1984) *International comparison of national precipitation gauges with a reference pit gauge, instruments and observing methods*. Geneva.

Shaw, S. B., Stedinger, J. R. and Walter, M. T. (2010) 'Evaluating Urban Pollutant Buildup/Wash-Off Models Using a Madison, Wisconsin Catchment', *Journal of Environmental Engineering*, 136(February), pp. 194–203. doi: 10.1061/(ASCE)EE.1943-7870.0000142.

Shrestha, D. L. and Solomatine, D. P. (2006) 'Machine learning approaches for estimation of prediction interval for the model output.', *Neural networks : the official journal of the International Neural Network Society*, 19(2), pp. 225–35. doi: 10.1016/j.neunet.2006.01.012.

Smith, J. A., Baeck, M. L., Villarini, G., Welty, C., *et al.* (2012) 'Analyses of a long-term, high-resolution radar rainfall data set for the Baltimore metropolitan region', *Water Resources Research*, 48(4), p. n/a-n/a. doi: 10.1029/2011WR010641.

Sonnen, M. B. (1980) 'Urban Runoff Quality: Information Needs', *Journal of the Technical Councils ASCE*, Vol. 106, pp. 29–40.

Sorooshian, S. and Gupta, V. K. (1983) 'Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness', *Water Resources Research*, 19(1), pp. 260–268. doi: 10.1029/WR019i001p00260.

Tabios III, G. Q. and Salas, J. D. (1985) 'A comparative analysis of techniques for spatial interpolation of precipitation', *Journal of the American Water Resources Association*, 21(3), pp. 365–380.

Todini, E. (2008) 'A model conditional processor to assess predictive uncertainty in flood forecasting', *International Journal of River Basin Management*, 6(2), pp. 123–137. doi: 10.1080/15715124.2008.9635342.

Vaze, J. and Chiew, F. H. S. (2003) 'Comparative evaluation of urban storm water quality models', *Water Resources Research*, 39(10), pp. 1–10. doi: 10.1029/2002WR001788.

Viessman Jr., W. and Lewis, G. L. (1995) 'Introduction to hydrology', in *Introduction to hydrology*. 4th edn. Harper Collins, p. 311.

Vihola, M. (2012) 'Robust adaptive Metropolis algorithm with coerced acceptance rate', *Statistics and Computing*, pp. 997–1008. doi: 10.1007/s11222-011-9269-5.

Villarini, G., Mandapaka, P. V., Krajewski, W. F. and Moore, R. J. (2008) 'Rainfall and sampling uncertainties: A rain gauge perspective', *Journal of Geophysical Research: Atmospheres*, 113(11), pp. 1–12. doi: 10.1029/2007JD009214.

Vrugt, J., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., *et al.* (2008) 'Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation', *Water Resources Research*, 44(12). doi: 10.1029/2007WR006720.

Van der Waerden, B. L. . (1953) 'Order tests for two-sample problem and their power I-III', *Indagationes Mathematicae*, 14,15, p. 453–458 (14), 303–316(15).

WaPUG (2002) 'Code of practice for the hydraulic modelling of sewer systems.', *Wastewater Planning Users Group, Birmingham, England.*, 2002(November).

Webster, R. and Oliver, M. a (2007) *Geostatistics for environmental scientists*. Second edi. West Sussex, England: John Wiley & Sons, Ltd.

Weerts, A. H., Winsemius, H. C. and Verkade, J. S. (2011) 'Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales)', *Hydrology and Earth System Sciences*, 15(1), pp. 255–265. doi: 10.5194/hess-15-255-2011.

Wilde, B. (2010) *MIN E 310 Class Notes*. University of Alberta.

Wischmeier, L. D. M. and W. H. (1969) 'Mathematical Simulation of the Process of Soil Erosion by Water', *Transactions of the ASAE*. St. Joseph, MI: ASAE, 12(6), p. 754. doi: https://doi.org/10.13031/2013.38945.

Wood, S. J., Jones, D. a. and Moore, R. J. (2000) 'Accuracy of rainfall measurement for scales of hydrological interest', *Hydrology and Earth System Sciences*, pp. 531–543. doi: 10.5194/hess-4-531-2000.

Young, A. (2009) 'Manual for Streets', *Department for Transport, and Communities and Local Government.*, 162(3), pp. 129–131. doi: 10.1680/muen.2009.162.3.129.

Zhao, J., Chen, Y., Hu, B. and Yang, W. (2015) 'Mathematical Model for Sediment Wash-Off from Urban Impervious Surfaces', *Journal of Environmental Engineering*, 142(4), p. 4015091. doi: 10.1061/(ASCE)EE.1943-7870.0001058.

Zug, M., Phan, L., Bellefleur, D. and Scrivener, O. (1999) 'Pollution wash-off modelling on impervious surfaces: Calibration, validation, transposition', in *Water Science and Technology*. No longer published by Elsevier, pp. 17–24. doi: 10.1016/S0273-1223(99)00004-9.