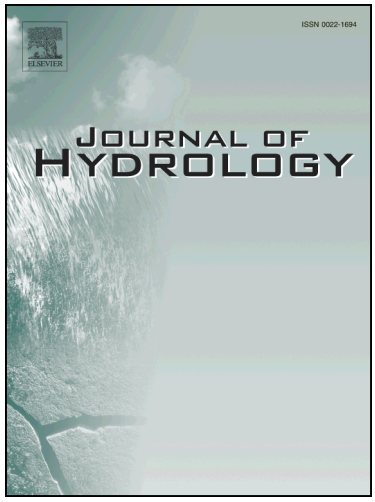


Research papers

Quantifying rainfall-derived inflow and infiltration in sanitary sewer systems based on conductivity monitoring

Mingkai Zhang, Yanchen Liu, Xun Cheng, David Z. Zhu, Hanchang Shi, Zhiguo Yuan

PII: S0022-1694(18)30002-7
DOI: https://doi.org/10.1016/j.jhydrol.2018.01.002
Reference: HYDROL 22484

To appear in: *Journal of Hydrology*

Received Date: 22 May 2017
Revised Date: 7 November 2017
Accepted Date: 1 January 2018

Please cite this article as: Zhang, M., Liu, Y., Cheng, X., Zhu, D.Z., Shi, H., Yuan, Z., Quantifying rainfall-derived inflow and infiltration in sanitary sewer systems based on conductivity monitoring, *Journal of Hydrology* (2018), doi: https://doi.org/10.1016/j.jhydrol.2018.01.002

Quantifying rainfall-derived inflow and infiltration in sanitary sewer systems based on conductivity monitoring

Mingkai Zhang[a], Yanchen Liu[a]*, Xun Cheng[a], David Z. Zhu[b], Hanchang Shi[a], Zhiguo Yuan[a,c]

a. State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China, 100084

b. Department of Civil and Environmental Engineering, University of Alberta, T6G2W2, Edmonton, Alberta, Canada

c. Advanced Wastewater Management Centre (AWMC), The University of Queensland, St. Lucia, Queensland 4072, Australia

*Email: **liuyc@mail.tsinghua.edu.cn**. Tel: +86-10-62796953; Fax: +86-10-62771472

Abstract: Quantifying rainfall-derived inflow and infiltration (RDII) in a sanitary sewer is difficult when RDII and overflow occur simultaneously. This study proposes a novel conductivity-based method for estimating RDII. The method separately decomposes rainfall-derived inflow (RDI) and rainfall-induced infiltration (RII) on the basis of conductivity data. Fast Fourier transform was adopted to analyze variations in the flow and water quality during dry weather. Nonlinear curve fitting based on the least squares algorithm was used to optimize parameters in the proposed RDII model. The method was successfully applied to real-life case studies, in which inflow and infiltration were successfully estimated for three typical rainfall events with total rainfall volumes of 6.25 mm (light), 28.15 mm (medium), and 178 mm (heavy). Uncertainties of model parameters were estimated using the generalized likelihood uncertainty estimation

(GLUE) method and were found to be acceptable. Compared with traditional flow-based methods, the proposed approach exhibits distinct advantages in estimating RDII and overflow, particularly when the two processes happen simultaneously.

Keywords: Conductivity; Inflow; Infiltration; Sanitary sewer overflow; Sewer system

1 Introduction

Sewer systems transport municipal wastewater from cities to treatment facilities. Rainfall-derived inflow and infiltration (RDII) can cause serious operational problems in sanitary sewer systems and downstream wastewater treatment plants. These processes amplify pumping costs, increase the wastewater load, and dilute the wastewater. Additionally, RDII increases the probability of sanitary sewer overflows (SSOs) and risk of drainage pipe collapse (Vallabhaneni et al. 2008). The inflow and infiltration volumes can reach more than 100% of the ordinary wastewater quantity, and this characteristic directly decreases the pollutant removal efficiency (Ellis and Bertrand-Krajewski 2010). These phenomena seriously threaten infrastructure security and wastewater treatment plant (WWTP) operations. Thus, monitoring and assessing the inflow and infiltration processes are crucial to solve the problems associated with the operation and management of sewer systems.

Flow information can be readily used to assess inflow and infiltration into a sewerage system. These processes are commonly analyzed based on the principle of balancing annual or daily flow rates. For example, the “triangle method” ranks all flow values (daily mean values) in ascending order to estimate inflow and infiltration (Weiss et al. 2002). Minimum nightly flow was also used to determine infiltration (Franz 2006). Moreover, several reformative methods have used filters to extract time series of total flow and determine the

quantity of inflow and infiltration (Vaes et al. 2005). These flow-based methods do not accurately quantify when sewer overflow occurs as a result of RDII.

Additionally, water quality monitoring has been adopted to assess inflow and infiltration coupled with flow data. Traditional pollutant index such as chemical oxygen demand (COD) (Kracht and Gujer 2005, Bareš et al. 2009), total nitrogen (TN) (Shelton et al. 2011) and total phosphorus (TP) (Mattsson et al. 2016) were used to assess inflow and infiltration. TN and TP were found to be more reliable water quality parameters. Several stable intrinsic tracers have also been applied to assess RDII as alternative to normal water quality indicators. Naturally occurring stable isotopes of water ($^{18}O/^{16}O$) (Kracht et al. 2007, De Bénédittis et al. 2005, Houhou et al. 2010) was successfully used to determine the source and the amount of rainwater or groundwater entering an drainage system. These methods can provide accurate information regarding inflow and infiltration, but they require intensive measurements. Online collection of such water quality data is difficult, especially in large sewer systems.

Complex models based on hydrological and hydraulic mechanisms have also been applied to simulate inflow and infiltration. Distributed hydrological models are generally used to describe surrounding hydrogeological processes and their interactions with the sewer network. Sewer flow is commonly assumed to be composed of sewage, inflow, and infiltration. Inflow and infiltration can be further separated into three different components as follows: i) groundwater infiltration (GWI); ii) rain-induced infiltration (RII); and iii) direct storm water inflow, which is also called rain-derived inflow (RDI) (Bennett 1999, Staufer et al. 2012). Hydrological models are commonly used to illustrate the transformation of precipitation into runoff at the subcatchment scale. Belhadj et al. (1995) proposed a simple multiparameter conceptual model to simulate RII in a sewer system using hourly measurements of rainfall and flow rates. Gustafsson et al. (1999) developed

the MouseNAM model, in which the fast runoff component from impervious surfaces and the slow runoff component caused by infiltration into the sewer system from the surrounding soil were both considered. The Run-off Routing (RORB) model is a similar model based on several tanks in series that can be used to distinguish various flow components (Laurenson et al. 1990). Karpf and Krebs (2011) presented a method for quantifying inflow and infiltration based on a multimodel approach. In the Stormwater Management Model (SWMM) and the Sanitary Sewer Overflow Analysis and Planning toolbox, RDII can be presented using the so-called RTK method based on a synthetic unit hydrograph. This method assumes that RDII occurs in a sewer in response to a specific precipitation volume over a given period based on the sewershed characteristics (Lai 2008). Boukhemacha et al. (2015) studied urban groundwater flow using a mathematical model with geospatial analysis. Mao et al. (2016) proposed an approximate point source method for measuring soil infiltration and verified the method using three soil types in laboratory experiments. However, similar to flow-based methods, these hydrologic/hydraulic methods are not suitable for flow situations where overflow and/or backflow occur.

Furthermore, several statistical and probabilistic methods have been introduced to study the relationship between RDII and rainfall volume. Zhang (2005) proposed an autoregressive regression framework, in which flow and rainfall data were considered to realize reliable estimation of RDII. Pate and Rahman (2010) investigated the applicability of Monte Carlo simulation with the RORB model. Mikalson (2011) developed two conceptualizations to estimate RDII using a derived probability distribution theory. However, these complex statistical and probabilistic models require large computational runtimes and are difficult to implement in real-time simulations based on online monitoring.

This study is performed to develop a simple and cost-effective method to estimate inflow, infiltration, and

sewer overflow in a sanitary sewer system based on conductivity data, which are simple and easy to monitor. The method facilitates the estimation of RDII as RII and RDI separately not only in normal flow conditions, but also when RDII occurs simultaneously with overflow or backflow. The proposed approach enables the quantification of overflow by regenerating flow data based on the estimated RDII.

2 Methods and Materials

2.1 Field trial and data preparation

Online wastewater flow and quality data were collected in a sewer system in Wuxi City, China (Figure S1). The catchment area of the monitoring site was approximately 0.8 km^2. This network was mainly a separate sanitary sewer system with hidden cross-connections. The pipe diameter at the monitoring site was 600 mm, and the dry weather flow varied between 40 m^3/h and 120 m^3/h. The flow meter used in the system was a HACH FL900AV Flow Meter. The depth was measured using a submerged pressure transducer in the flowmeter. Rainfall event was recorded by a Sigma Rain Gauge Tipping Bucket with a Hach FL900 Series Flow Logger. Conductivity was measured using a water quality instrument (Product Model: PLOC100, independent development) that monitors the conductivity, temperature, oxidation-reduction potential (ORP), and pH. Data collection lasted for approximately one year. All data (flow, water level, conductivity, and rainfall) were collected at 10-minute intervals. The quality of the monitoring data was assessed using Local Regression Smoothing (Hastie and Loader 1993) to remove disturbances and outliers.

When RDII exceeded the conveyance capacity of the system, the water level in the system increased and caused backwater or even overflow. Rainfall events could be grouped into three categories, namely, light, medium, or heavy. Light rainfall events were events that raised water level in the sewer system without

backwater formation. Medium rainfall events were events that raised the water level and caused backwater to form (i.e., the flow decreased but was still positive). These phenomena indicated that the inflow and infiltration exceeded the capacity of the sewer, and SSOs occurred. Heavy rainfall events were events that raised the water level and caused backflow (i.e., the flow at the monitoring site become negative). Three representative rainfall events are shown in Table 1.

2.2 Variation in flow and conductivity in dry weather

Dry weather flow and its conductivity can be characterized using spectral analysis and fast Fourier transform (FFT) into components of sine functions as follows:

$$y = \sum_{i=1}^{n} A_i \sin\left(2\pi f_i t + \phi_i\right) + C, \qquad (1)$$

where y is the flow rate or conductivity at the monitoring point; and A_i, f_i, and ϕ_i represent the amplitude, frequency, and phase of the i^{th} sine function, respectively. C is the average value of the diurnal variation in the flow rate or conductivity. The major frequencies (f_i) were determined based on the high amplitudes (A_i) generated by the FFT.

2.3 Model of inflow and infiltration with instantaneous unit hydrograph in wet weather flow

Inflow and infiltration can be simulated as the final outflow of a series of cascaded linear reservoirs. As derived in the Supplementary Information [Equations (S1) and (S2)], the outflow from the cascaded linear reservoirs is as follows (Nash 1957):

$$\mathrm{u}(0,t) = \frac{1}{K \cdot \Gamma(N)} \cdot \left(\frac{t}{K}\right)^{\mathrm{N}-1} \cdot e^{-\frac{t}{K}} \qquad (2)$$

where $u(0,t)$ is the instantaneous unit hydrograph, which represents the RDII flow as the result of unit, instantaneous rainfall. *N* is the number of linear reservoirs, $\Gamma(N)$ is the gamma function of *N*, and *K* is the storage coefficient of the linear reservoir (Supplementary Information).

Rainfall is always continuous rather than instantaneous. Thus, the instantaneous unit hydrograph should be converted to a temporal unit hydrograph, which is used to describe the RDII process produced by per unit rainfall in the record time span (Δt, which is 10 min in our study) as follows:

$$u(\Delta t,t)=\frac{1}{\Delta t}\left[\int_0^t u(0,\tau)\,d\tau-\int_{\Delta t}^t u(0,\tau\text{-}\Delta t)\,d\tau\right] \quad (3)$$

where Δt is the record time interval of rainfall, $u(\Delta t,t)$ represents the outflow as the result of unit rainfall during the period of Δt.

Parameter *R* was used to represent the fractions of rainfall volume that enter the sewer system as RDII. The whole monitoring time can be divided into *n* parts by the record time interval, and *i* is the ordinal number from 1 to *n*. P_i is the rainfall depth per unit time (rainfall record time interval), and the net rainfall entering into the sanitary sewer system could be presented as the product of *R* and P_i in each unit time. *F* is the sub-catchment area upstream of the monitoring site. By summing the $u(\Delta t,t)$ produced by each net rainfall, the discharge of RDII can be expressed as follows:

$$Q(t)=F\sum_{i=1}^{n}R\cdot P_i\cdot u_i(\Delta t,t) \quad (4)$$

RDI and RII were described using the same principle of cascading linear reservoirs to reduce the computational complexity. For the inflow process, R_{RDI}, K_{RDI}, and N_{RDI} were used to substitute for *R*, *K*, and *N*

in Equations (2) and (4). Moreover, $u(\Delta t,t)$ in Equation (3), and *Q(t)* in Equation (4) would be substituted by $u_{RDI}(\Delta t,t)$ and $Q_{RDI}(t)$, respectively. Similarly, substituting R_{RII}, K_{RII}, and N_{RII} for *R*, *K*, and *N*, $u(\Delta t,t)$ will be become $u_{RII}(\Delta t,t)$, and *Q(t)* will be $Q_{RII}(t)$. R_{RDI} and R_{RII} represent the ratio of rainfall that enters the sewer system as RDI and RII, respectively, while K_{RDI} and K_{RII} represent the storage coefficients of RDI and RII, respectively, while N_{RDI} and N_{RII} represent the number of linear reservoirs of RDI and RII, respectively. Then, the wet weather flow $Q_{WWF}(t)$ is the sum of the dry weather flow $Q_{DWF}(t)$ [from Equation (1)], RDI flow $Q_{RDI}(t)$, and RII flow $Q_{RII}(t)$, as shown by Equation (5) as follows.

$$Q_{WWF}(t) = Q_{DWF}(t) + F\sum_{i=1}^{n} R_{RDI}\cdot P_i \cdot u_{RDI_i}(\Delta t,t) + F\sum_{i=1}^{n} R_{RII}\cdot P_i \cdot u_{RII_i}(\Delta t,t) \quad (5)$$

2.4 Pollutant balance model in wet weather flow

Conductivity was chosen as the pollutant indicator in the model. RDI and RII normally decrease the conductivity of domestic sewage, as expressed in the material balance equation. The components of wet weather conductivity can be combined as Equation (6) as follows:

$$c_{WWF}(t) = \frac{c_{DWF}(t)Q_{DWF}(t) + c_{RDI}(t)F\sum_{i=1}^{n} R_{RDI}\cdot P_i \cdot u_{RDI_i}(\Delta t,t) + c_{RII}(t)F\sum_{i=1}^{n} R_{RII}\cdot P_i \cdot u_{RII_i}(\Delta t,t)}{Q_{DWF}(t) + F\sum_{i=1}^{n} R_{RDI}\cdot P_i \cdot u_{RDI_i}(\Delta t,t) + F\sum_{i=1}^{n} R_{RII}\cdot P_i \cdot u_{RII_i}(\Delta t,t)} \quad (6)$$

c_{WWF} is the conductivity of the wet weather flow; c_{DWF}, c_{RDI}, and c_{RII} are the conductivities of dry weather flow, RDI, and RII, respectively. The conductivity of RDI and RII were monitored in a catch-basin nearby.

2.5 Data analysis to determine RDII

Q_{DWF} and C_{DWF} can be described by Equation (1). The parameters related to dry weather would be determined

first by the dry weather flow and conductivity. Then, six parameters (R_{RDI}, R_{RII}, K_{RDI}, K_{RII}, N_{RDI}, and N_{RII}) related to the wet weather were used in the model. These six parameters were calibrated either by fitting the observed wet weather flow [Equation (5)] or by fitting the wet weather conductivity [Equation(6)], which will produce two parameter sets to be compared to crosscheck the model applicability in different rainfall scenarios. Moreover, the parameter set calibrated by Equation (6) can be substituted into Equation (5) to calculate wet weather flow and vice versa. The goodness-of-fit was evaluated using the Nash–Sutcliffe (NS) efficiency coefficient [Equation (S3)[. The above data analysis methodology is summarized in Figure 1.

Nonlinear fitting method based on iterative reweighted least squares algorithm (Holland and Welsch 1977; Dumouchel and O'Brien 1991) was performed to optimize the parameters in the model. The weights of each iteration were recalculated based on the residual of the observed data of the previous iteration. This recalculation mitigated the influence of outliers. Iterations continued until the weights converged. Given that RDI represents the direct inflow of runoff, RII represents infiltration, in which rainwater enters into the system through soil. Inflow and infiltration can be separated, because they have different flow rates in physics. During fitting, the constraints of parameters were determined as $R_{RDI}>R_{RII}$, $K_{RDI}<K_{RII}$, $N_{RDI}<N_{RII}$ to guarantee the flow rate difference.

The uncertainty of the model and 90% confidence interval of parameters were evaluated with generalized likelihood uncertainty estimation (GLUE) method (Beven and Binley 1992). GLUE adopts the concept of equifinality of models, parameters, and variables. This methodology consists of three steps (Jensen 2003). First, the prior distribution, which is typically discrete or continuous uniform distribution, of the parameters was determined. Second is the stochastic simulation based on the parameters defined previously, with Monte Carlo method to evaluate a random sample of the parameter sets. Third, the simulation and the corresponding

parameter sets were rated according to which they fit the observed data for every single simulation. If the simulated results were close to the observed values, the simulation and this parameter set were accepted as having a given likelihood (NS value in our case). Otherwise, this simulation and parameter set would be rejected. These three steps were repeated, until iterations reached the set point (5000 iterations in our case). The cumulative distribution function (CDF) was calculated for all accepted parameters, and the value corresponding to 5% and 95% of the CDF will be the 90% confidence interval of the parameters. The model uncertainty was revealed through the CDF of the accepted simulation results.

3 Results

3.1 Flow and conductivity variations during dry weather

Time series of dry weather flow rate and conductivity profiles were analyzed using spectral analysis and FFT. The dry weather data before the large rainfall event were analyzed as an example. Spectral analysis was used for all seven days of the dry weather flow to obtain an overall fitting. Two major frequencies were extracted from the flow and conductivity monitoring data [Figures 2(a)–(d)]. Frequencies of 1/day and 2/day reflected periods of 1 day and 0.5 days, and these values correspond to water usage habits with morning and evening peaks, as shown by Butler et al. (1995) and Gernaey et al. (2011). The flow amplitude at a frequency of 1/day was 62.80 m^3/h, while the amplitude at 2/day was 37.69 m^3/h [Figures 2(a) and (b), Table S1]. Thus, the diurnal period was more notable than the semidiurnal period. Conductivity also exhibited similar periodic characteristics [Figures 2(c) and (d)]. Increasing the number of Fourier series yielded higher NS values. For example, the NS values were 0.67, 0.86, and 0.93 using second-order, fifth-order, and tenth-order Fourier series, respectively. The fifth-order Fourier series were adopted in our study, because it can meet the

requirement to regenerate the dry weather flow and conductivity (Tables S1 and S2). Figures 2(a) and (c) show the fitting results of the flow and conductivity data, respectively.

3.2 RDII determination during a light rainfall event

A small rainfall event was used to calibrate the model and estimate RDI and RII. In Figure 3(a), the mean dry weather flow before the rainfall event was 44 m^3/h, while the maximum wet weather flow reached 131.4 m^3/h during the rainfall event. The model parameters were determined based on the observed flow data using Equation (5) (Table 2). The results indicate that the fraction of rainfall volume converted to inflow was 22.7% (R_{RDI}) and the fraction of rainfall volume converted to infiltration was 13.2% (R_{RII}). Therefore, approximately 35.9% of the rainfall volume entered the sanitary sewerage pipes as RDII. During inflow, 6.8 cascade reservoirs were used, and each had a storage coefficient of 5.2. During infiltration, the soil was divided into 9.6 vertical cascade reservoirs, and each reservoir had a storage coefficient of 6.8. The large value of N resulted in slower infiltration than the inflow process as expected. The total RDI, total RII, and total RDII volumes were 1135.2 m^3, 660.7 m^3, and 1795.9 m^3, respectively (Table 3). Conductivity simulation based on Equation (6) is shown in Figure 3(b), and the parameter values are listed in Table 2. The parameters calculated based on the conductivity and flow data were similar (Table 2). Figures 3(c) and (d) show the RDI and RII flows determined based on flow and conductivity, respectively. The RDI and RII flows calculated based on the conductivity data exhibited good agreement with the result based on flow data [Figures 3(e)-2(f)], confirming the suitability of conductivity as a water quality parameter for determination of RDI and RII flows. Using the parameters determined by conductivity to predict RDII of other two light rainfall events, the validation results are shown in Figure S2. Conductivity of the other two events were good fitted, indicating the parameter set can be applied to assess RDII of small rainfall events.

3.3 RDII determination during a medium rainfall event

A medium rainfall event created SSOs in this sanitary sewer system due to high water levels. Figure 4(a) shows that the flow increased on the second day and decreased on the third day after this event. However, the water level remained high until the third day. These phenomena indicated that some of the water was discharged as SSOs in the sewer system. Model parameters were determined based on the flow data for the medium rainfall event to test the adaptability of flow-based model. The model parameters and 90% confidence interval are listed in Table 2. The NS efficiency coefficient between the observed and simulated flow data was only 0.029 (Table 2), suggesting that the flow data were difficult to fit. Figure 4(b) shows that the simulated flow was lower than the observed flow. Thus, the flow-based model does not produce a good prediction for the observed flow data in this scenario because of backwater. Moreover, the calculated RDI and RII flows shown in Figure 4(c) are not reliable estimates of the true inflow and infiltration flows. Substituting the calculated RDI and RII flows into Equation (6) shows the modeled conductivity profile [Figure 4(d)]. The flow-based model underestimated RDII but overestimated the conductivity. The results indicate that the model based on the flow data cannot effectively estimate the inflow and infiltration process for this medium rainfall event.

When the conductivity data was used to estimate the RDI and RII flows for the medium rainfall event, the observed and simulated conductivity data (Figure 4(e)) fit well with an NS efficiency coefficient of 0.54 (Table 2). The fractions of the rainfall volume entering into the sewer system as inflow and infiltration are 10.1% (R_{RDI}) and 8.2% (R_{RII}), respectively (Table 2). R_{RDI} and R_{RII} were less than those associated with the light rainfall event probably because RDII exceeded the capacity of the system. The inflow and infiltration profiles are shown in Figure 4(f). Infiltration is considerably slower than inflow rate. The total RDII volume

was approximately 4172.4 m^3, with an RDI and RII volumes of 2302.8 m^3 and 1869.6 m^3, respectively.

The actual total wastewater flow was estimated based on the RDII flows and the DWF pattern [Figure 4(g)]. Significant difference was found between the simulated and observed flows mainly during the period with higher water levels [Figure 4(h)]. The flow difference was likely caused by storage in upstream sewerage pipes and overflow discharged from sewerage pipes upstream of the monitoring location. The amounts of overflow can be calculated by integrating the difference between the simulated and observed flows, which produced a value of 1207.9 m^3 for this event as a result of SSOs (Table 3). The result indicates that the conductivity-based model can identify inflow and infiltration flows during medium rainfall events with SSOs. The validation results when the parameters of medium rainfall event were used to model conductivity and RDII of other medium rainfall events are shown in Figure S3. The good fitting results show that the model and parameter set were applicable to predict RDII of medium rainfall events.

3.4 RDII determination during a heavy rainfall event

The proposed conductivity-based model was also implemented during a heavy rainfall event. This case demonstrated an extreme scenario when flow rate decreased abnormally after the rainfall [Figure 5(a)]. This scenario was more complex (Figure S4) than the previous two scenarios. The model simulation results based on the dynamic variations of conductivity are shown in Figure 5(b). The NS value between the observed and simulated conductivities was 0.59 (Table S1), indicating a good fit. The ratio of rainfall entering into the sanitary system as RDI (R_{RDI}) was 8.1%, while the ratio of rainfall transferring into RII (R_{RII}) was 3.5%. The estimated RDI and RII flows are shown in Figure 5(c). The total volumes of RDI and RII were 11534.4 m^3 and 4984.0 m^3, giving a total RDII volume of 16518.4 m^3 during the heavy rainfall event. The integration of

the difference between the observed and simulated flow data [Figure 5(d)] yielded an estimated SSO volume of 39732.2 m^3 during the rainfall event (Table 3). The results indicate that the conductivity model is suitable for estimating inflow and infiltration during extreme and abnormal flow scenarios associated with heavy rainfall events. By contrast, the flow-based model is obviously not suitable. The validation of large rainfall events are shown in Figure S5. In addition, the method was applied to larger scales from pump station to WWTP, and acceptable results were obtained (Figure S6). In larger scale, the method was still applicable according to the variation mode of conductivity.

4 Discussion

4.1 Choice of timespan for determining dry weather flow and conductivity patterns

For the proposed approach, determining a proper time scale is important to represent the characteristics of dry weather flow and conductivity profiles. Normally, the quantity and quality of domestic wastewater have two peaks in the morning and evening, respectively, with late-night and mid-day minima (Butler et al. 1995, Gernaey et al. 2011). However, distinct variations in quality and quantity are found in different sewer system catchments, and significant seasonal variations are observed. Simple periodic functions with periodicities of one day, such as low-order Fourier series, have been proposed in previous studies to describe the diurnal variations in dry weather flows and pollutant concentrations (Rodríguez et al. 2013). Generally, dry weather period for the studied sewer normally lasts for 1–weeks between two rainfall events (Figure S7). The influence of the different time scales of dry weather flow should be evaluated. The parameter distribution of dry weather [Equation (1)] was conducted based on different timescales, namely, 1, 2, 3, and 7 days, based on spectral analysis (Figure S8). The amplitudes, initial phases, and mean values of flow and conductivity at

frequencies of 1/day to 4/day are shown in Figure S9. The results indicate that variations in the given parameters decrease as the timescale increases. Moreover, the predictions of dry weather flow and conductivity based on 7 days were more accurate than those based only on individual days.

4.2 Parameter sensitivity and model uncertainty analysis

The uncertainty of a model is associated with three different sources, namely, the measured data uncertainty, parameter uncertainty, and model structure uncertainty (Saltelli et al. 2008). The assumption of cascading linear reservoirs for inflow and infiltration processes will introduce uncertainty into the model. The criteria for assessing simulation results should be determined based on the uncertainty analysis. Parameter sensitivity and uncertainty were analyzed using GLUE (Beven and Binley 1992). The parameters were divided into accepted and rejected categories according to NS values higher or lower than a threshold value. In this case, the NS threshold was chosen as 0.5 for conductivity and 0.3 for flow to obtain a better fitted result. The lower and upper bounds of the output range were obtained by resampling the parameters in the accepted sets. The medium rainfall event was chosen to evaluate the model uncertainty and parameter sensitivity. Figure 6 shows the 90% confidence regions of the simulated conductivity and calculated RDII, with the distribution of the six parameters in the wet weather model shown in Figure S10. Parameters R_{RDI} and R_{RII} exhibited approximately normal distributions, N_{RDI} and N_{RII} exhibited skewed distributions, and K_{RDI} and K_{RII} exhibited approximately uniform distributions. Two-sample Kolmogorov-Smirnov (KS) test (Wang et al. 2003) was used for the accepted and rejected parameter sets to evaluate the sensitivities of the parameters. The test results are shown in Table S3. Global sensitivity can be sorted according to the S value of KS test results. The findings show that N_{RDI} and N_{RII} had the highest sensitivity, while K_{RDI} and K_{RII} showed the lowest sensitivity. These characteristics indicate that N_{RDI} and N_{RII} should be paid more attention, such as giving more reasonable prior

distribution according to its physical meaning when calibrating parameters. More intensive online monitoring data and higher data quality can also reduce the uncertainty and increase the sensitivity of the parameters.

4.3 Implication of the proposed method

Understanding inflow and infiltration is important for sewer operations. Inflow and infiltration can be used as potential indicators to evaluate sewer conditions, such as possible cross-connections and pipe leakage/failures. Several infiltration models, such as Horton's Method (Horton 1933, 1945), Green-Ampt equation (Green and Ampt 1911), and the Curve Number Method (Rossman 2015) embedded in SWMM, can be used to determine the total volume of rainwater that infiltrate into groundwater. Karpf and Krebs (2011) assumed that the infiltration rate is a ratio of the difference between the groundwater level and the water level in pipes based on Darcy's law, and SWMM adopted an unsaturated upper zone and a saturated lower zone to describe the amount of groundwater entering a sewer system. However, these methods are difficult to apply to our monitoring area because of the lack of detailed infrastructure data. By contrast, the model proposed in this study first successfully decomposes the dynamic profile of RDII separately based on the conductivity monitoring. Compared with these methods, the proposed method use conductivity to evaluate the RDII process based on the instantaneous unit pollutograph. This process is advantageous, especially when backwater exists and the water level in the sewer system is high. Moreover, conductivity monitoring is simpler and easier than measuring COD, NH_4^+-N, or PO_4^{3-}-P in the sewer system.

5 Conclusions

A method for estimating inflow and infiltration flows in sewers was proposed and demonstrated using real-life applications. The method was based on conductivity data and an instantaneous unit pollutograph model. The

proposed method was successfully implemented to identify the dynamic inflow and infiltration flows associated with rainfall entering a sewer system. The proposed method has distinct advantages over traditional flow-based methods for estimating inflow and infiltration, especially when the events produce backwater, overflow, and abnormal flow data in sewer systems. The method uses very simple conductivity sensors. Thus, the proposed approach can be easily implemented for real-life applications.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (No. 51678337), Major Science and Technology Program for Water Pollution Control and Treatment of China (No. 2014ZX07305001) and the Tsinghua University Initiative Scientific Research Program (No. 2014z21028). Zhiguo Yuan acknowledges the support by the Australian Research Council-funded Industry Linkage project LP 160101040, with City of Gold Coast, Queensland Urban Utilities, and South Australian Water Corporation as collaborating partners.

Reference

Bareš, V., Stránský, D. and Sýkora, P. (2009) Sewer infiltration/inflow: long-term monitoring based on diurnal variation of pollutant mass flux. Water Science & Technology 60(1), 1-7.

Belhadj, N., Joannis, C. and Raimbault, G. (1995) Modelling of rainfall induced infiltration into separate sewerage. Water Science and Technology 32(1), 161-168.

Bennett, D. (1999) Using flow prediction technologies to control sanitary sewer overflows: project 97-CTS-8, Water Environment Research Foundation.

Beven, K. and Binley, A. (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrological Processes 6(3), 279-298.

Boukhemacha, M.A., Gogu, C.R., Serpescu, I., Gaitanaru, D. and Bica, I. (2015) A hydrogeological conceptual approach to study urban groundwater flow in Bucharest city, Romania. Hydrogeology Journal 23(3), 437-450.

Butler, D., Friedler, E. and Gatt, K. (1995) Characterising the quantity and quality of domestic wastewater inflows. Water Science and Technology 31(7), 13-24.

De Bénédittis, J. and Bertrand-Krajewski, J. (2005) Infiltration in sewer systems: comparison of measurement methods. Water Science & Technology 52(3), 219-227.

Dumouchel, W. and O'Brien, F. (1991) Integrating a robust option into a multiple regression computing environment. Institute for Mathematics and Its Applications 36, 41.

Ellis, J.B. and Bertrand-Krajewski, J.L. (2010) Assessing Infiltration and Exfiltration on the Performance of Urban Sewer Systems (APUSS), IWA Pub.

Franz, T. (2006) Spatial classification methods for efficient infiltration measurements and transfer of measuring results. Dissertation, Dresden University of Technology, Institute for Urban Water management.

Gernaey, K.V., Flores-Alsina, X., Rosen, C., Benedetti, L. and Jeppsson, U. (2011) Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach. Environmental Modelling & Software 26(11), 1255-1267.

Green, W.H. and Ampt, G. (1911) Studies on Soil Phyics. The Journal of Agricultural Science 4(01), 1-24.

Gustafsson, L., Hernebring, C. and Hammarlund, H. (1999) Continuous Modelling of Inflow/Infiltration in Sewers

with MouseNAM-10 Years of Experience. Third DHI Software Conference.

Hastie, T., & Loader, C. (1993). Local regression: Automatic kernel carpentry. Statistical Science, 120-129.

Holland, P.W. and Welsch, R.E. (1977) Robust regression using iteratively reweighted least-squares. Communications in Statistics-theory and Methods 6(9), 813-827.

Horton, R.E. (1933) The role of infiltration in the hydrologic cycle. Eos, Transactions American Geophysical Union 14(1), 446-460.

Horton, R.E. (1945) Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Geological Society of America Bulletin 56(3), 275-370.

Houhou, J., Lartiges, B., France-Lanord, C., Guilmette, C., Poix, S. and Mustin, C. (2010) Isotopic tracing of clear water sources in an urban sewer: A combined water and dissolved sulfate stable isotope approach. Water Research 44(1), 256-266.

Jensen, J. B. (2003). Parameter and Uncertainty Estimation in Groundwater Modelling. Aalborg : The Hydraulics and Coastal Engineering Group, Dept. of Civil Engineering, Aalborg University. (Series Paper; No. 23).

Karpf, C. and Krebs, P. (2011) Quantification of groundwater infiltration and surface water inflows in urban sewer networks based on a multiple model approach. Water Research 45(10), 3129-3136.

Kracht, O. and Gujer, W. (2005) Quantification of infiltration into sewers based on time series of pollutant loads. Water Science and Technology 52(3), 209-218.

Kracht, O., Gresch, M. and Gujer, W. (2007) A stable isotope approach for the quantification of sewer infiltration. Environmental Science & Technology 41(16), 5839-5845.

Lai, F.-h. (2008) Review of sewer design criteria and RDII prediction methods, p. 30, United States Environmental Protection Agency, Washington DC.

Laurenson, E.M., Mein, R.G. and Nathan, R. (1990) RORB--version 3, Runoff Routing Program: User Manual, Monash University Department of Civil Engineering.

Mao, L., Li, Y., Hao, W., Mei, X., Bralts, V.F., Li, H., Guo, R. and Lei, T. (2016) An approximate point source method for soil infiltration process measurement. Geoderma 264, Part A, 10-16.

Mattsson, J., Mattsson, A., Davidsson, F., Hedström, A., Österlund, H. and Viklander, M. (2016) Normalization of Wastewater Quality to Estimate Infiltration/Inflow and Mass Flows of Metals. Journal of Environmental Engineering, 04016050.

Mikalson, D.T. (2011) Development of Analytical Probabilistic Models for the Estimation of Rainfall Derived Inflow/Infiltration Frequency, MASc Thesis, University of Toronto.

Nash, J. (1957) The form of the instantaneous unit hydrograph. International Association of Scientific Hydrology, Publ 3, 114-121.

Pate, H. and Rahman, A. (2010) Design flood estimation using Monte Carlo Simulation and RORB Model: stochastic nature of RORB model parameters, p. 4692. Proceedings of World Environmental and Water Resources Congress, held in Providence, Rhode Island, 16-20 May, 2010.

Rodríguez, J.P., McIntyre, N., Díaz-Granados, M., Achleitner, S., Hochedlinger, M. and Maksimović, Č. (2013) Generating time-series of dry weather loads to sewers. Environmental Modelling & Software 43, 133-143.

Rossman, L.A. (2015) Storm Water Management Model Reference Manual Volume I – Hydrology, U.S.

Environmental Protection Agency.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008) Global sensitivity analysis: the primer, John Wiley & Sons.

Shelton, J.M., Kim, L., Fang, J., Ray, C. and Yan, T. (2011) Assessing the severity of rainfall-derived infiltration and inflow and sewer deterioration based on the flux stability of sewage markers. Environmental Science & Technology 45(20), 8683-8690.

Staufer, P., Scheidegger, A. and Rieckermann, J. (2012) Assessing the performance of sewer rehabilitation on the reduction of infiltration and inflow. Water Research 46(16), 5185-5196.

Vallabhaneni, S., Lai, F.-h., Chan, C., Burgess, E.H. and Field, R. (2008) SSOAP—a USEPA toolbox for SSO analysis and control planning.

Vaes, G., Willems, P. and Berlamont, J. (2005) Filtering method for infiltration flow quantification, Proceedings 10th International Conference on Urban Drainage, p.7

Wang, J., Tsang, W.W. and Marsaglia, G. (2003) Evaluating Kolmogorov's distribution. Journal of Statistical Software 8(18).

Weiss, G., Brombach, H. and Haller, B. (2002) Infiltration and inflow in combined sewer systems: long-term analysis. Water Science and Technology 45(7), 11-19.

Zhang, Z. (2005) Flow data, inflow/infiltration ratio, and autoregressive error models. Journal of Environmental Engineering 131(3), 343-349.

Tables and Figures

Figure 1. The diagram of the model integration.

Figure 2. Spectrum analyses of flow and conductivity data. (a) Dry weather flow and model fit using second-order and fifth-order Fourier series; (b) Spectrum analysis of flow data; (c) Dry weather conductivity and model fit using second-order and fifth-order Fourier series; (d) Spectrum analysis of conductivity data.

Figure 3. Estimating RDII for the light rainfall event. (a) Simulated and observed flows; (b) Simulated and observed conductivity profiles; (c) Inflow and infiltration flows calculated based on flow data; (d) Inflow and infiltration flows calculated based on conductivity data; (e) Comparison of RDI flow calculated from flow and conductivity data; (f) Comparison of RII flow calculated from flow and conductivity data.

Figure 4. Estimating RDII using flow data ((a)- (d)) and conductivity data ((e) - (h)) for the medium rainfall event. (a) Observed flow and water level in the manhole; (b)Observed and Simulated flow data by fitting flow data; (c)RDII calculated based on flow data; (d)Observed and simulated conductivity data by fitting flow data; (e)Observed and simulated conductivity data by fitting conductivity data; (f) RDII calculated based on conductivity data. (g)Observed and Simulated flow data by fitting conductivity data; (h) Dynamic water level and flow difference (Δflow) between the observed and simulated flow by fitting conductivity data.

Figure 5. Estimating RDII using conductivity data for the heavy rainfall event. (a) Modeled and observed flow data; (b) Observed and simulated conductivity data; (c) RDII calculated based on conductivity data; (d) Dynamic water level and flow difference (Δflow) between the observed and simulated flow data.

Figure 6. Model uncertainty analysis of the medium rainfall event. (a) 90% confidence region of simulated

conductivity; (b) 90% confidence region of RDII.

Table 1. Three typical rainfall scenarios used to evaluate the model.

Table 2. Optimized parameters and 90% confidence interval based on flow and conductivity under different rainfall events.

Table 3. Volume (m^3) of RDII and SSOs in the three rainfall events

Figure S1. The map of monitoring site and the sewer system.

Figure S2. Validation of small rainfall events.

Figure S3. Validation of medium rainfall events.

Figure S4. The backflow process in the sewer system. (a) Level and velocity. (b) Flow variation at the monitoring site and conductivity profiles at an upstream and a downstream location.

Figure S5. Validation of large rainfall events.

Figure S6. Fitting result of conductivity in WWTP and pump station.

Figure S7. Daily rainfall data recorded from 2013 to 2015.

Figure S8. Four different ways to group the dry weather flow and conductivity to identify parameters. (A) Dividing of the seven days of dry weather flow and conductivity into seven parts, each comprising one day; (B) Grouping of the seven days of dry weather flow and conductivity every two days; (C) Grouping of the

seven days of dry weather flow and conductivity every three days. (D) Treating the seven days' data as a whole.

Figure S9. Parameter identification of dry weather flow and conductivity on different timescales with spectral analysis. (a) Flow amplitude at frequencies of 1 d^{-1} to 4 d^{-1}; (b) Initial phase of flow at frequencies of 1 d^{-1} to 4 d^{-1}; (c) Mean value of flow; (d) Amplitude of conductivity at frequencies of 1 d^{-1} to 4 d^{-1}; (e) Initial phase of conductivity at frequencies of 1 d^{-1} to 4 d^{-1}; (f) Mean value of conductivity.

Figure S10. Distributions of the six parameters of the medium rainfall event with GLUE method.

Table S1. Results of spectrum analysis of dry weather flow

Table S2. Results of spectrum analysis of dry weather conductivity

Table S3. KS test for the sensitivity analysis of model parameters

Quantifying rainfall-derived inflow and infiltration in sanitary sewer systems based on conductivity monitoring

Mingkai Zhang[a], Yanchen Liu[a]*, Xun Cheng[a], David Z. Zhu[b], Hanchang Shi[a], Zhiguo Yuan[a,c]

a. State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China, 100084

b. Department of Civil and Environmental Engineering, University of Alberta, T6G2W2, Edmonton, Alberta, Canada

c. Advanced Wastewater Management Centre (AWMC), The University of Queensland, St. Lucia, Queensland 4072, Australia

*Email: **liuyc@mail.tsinghua.edu.cn**. Tel: +86-10-62796953; Fax: +86-10-62771472

This file includes:

Figure 1 to 6

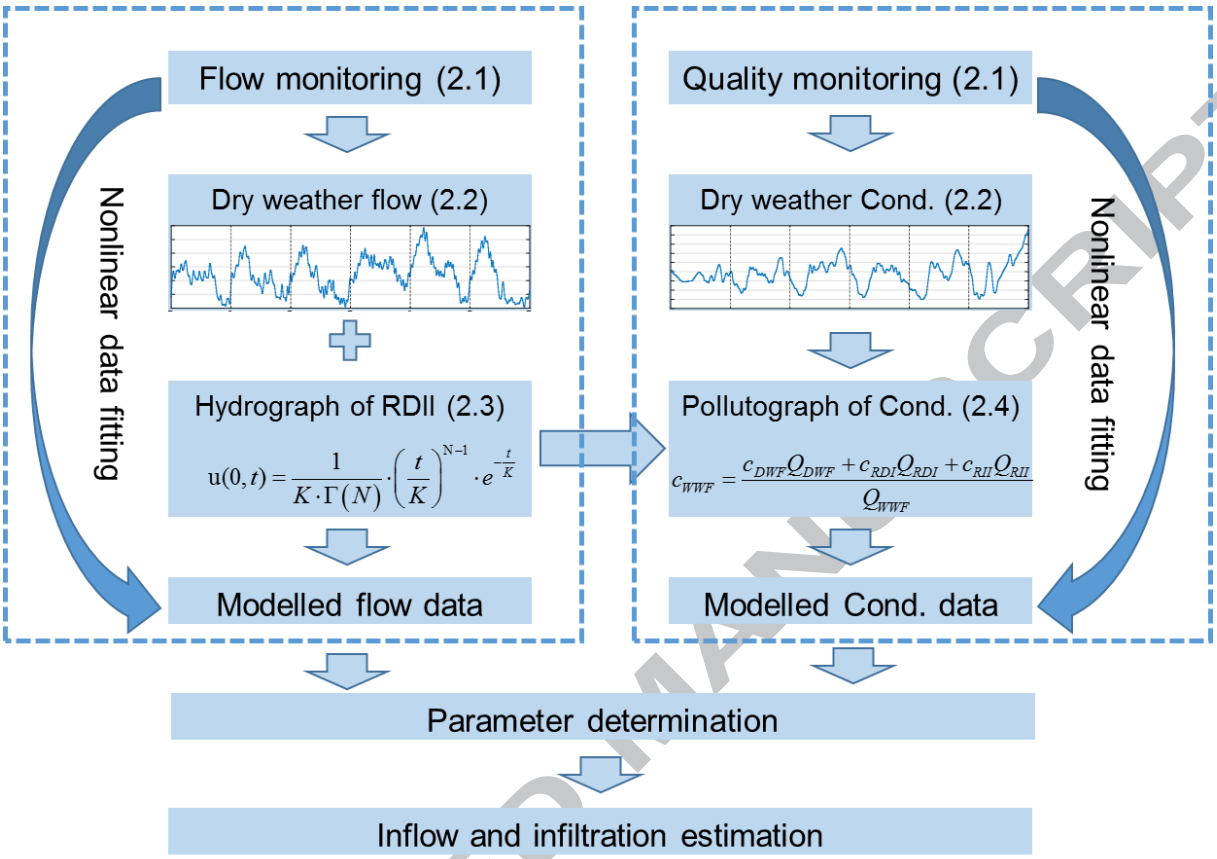
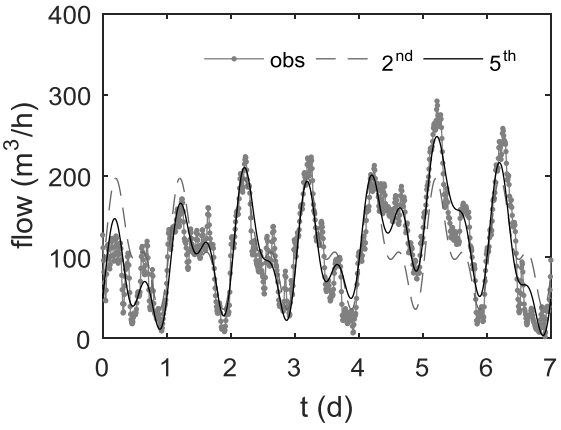
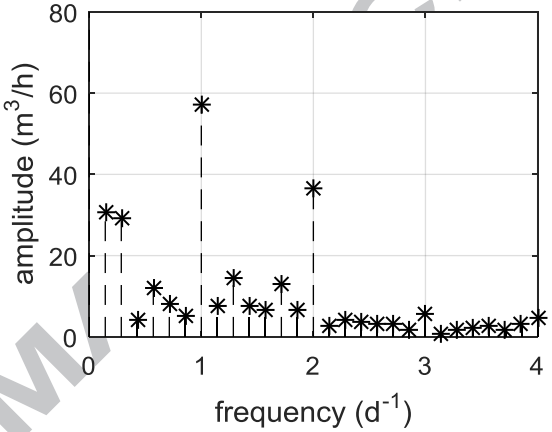


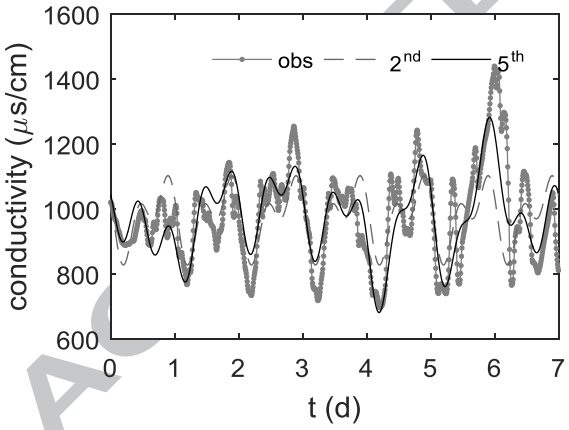
Figure 1. The diagram of the model integration. In Section 2.1, the online monitoring of flow and conductivity are illustrated. In section 2.2, the basic variation pattern of dry weather flow and conductivity is extracted by spectrum analysis. In section 2.3, the hydrograph of RDII is introduced. The wet weather flow is obtained by summing dry weather flow and RDII. By fitting the monitored wet weather flow, the parameters are determined. In section 2.4, the pollutograph of conductivity is generated with material balance equation. By fitting the monitored wet weather conductivity, the parameters are determined. The RDI and RII can be estimated by the calibrated parameters.



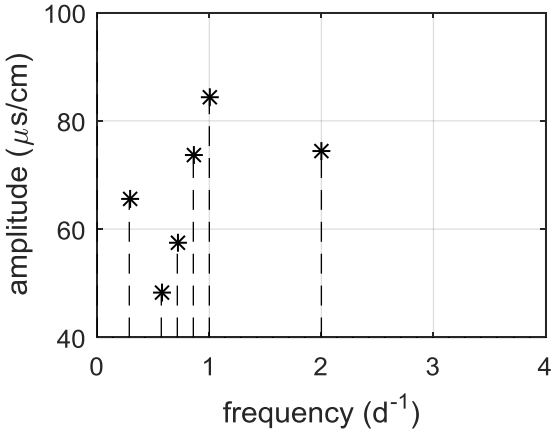
(a)



(b)

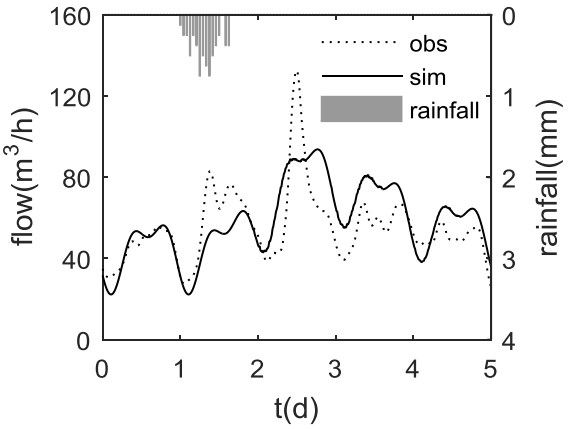


(c)

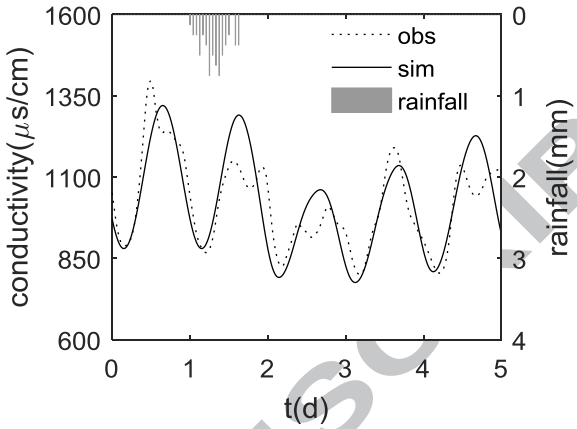


(d)

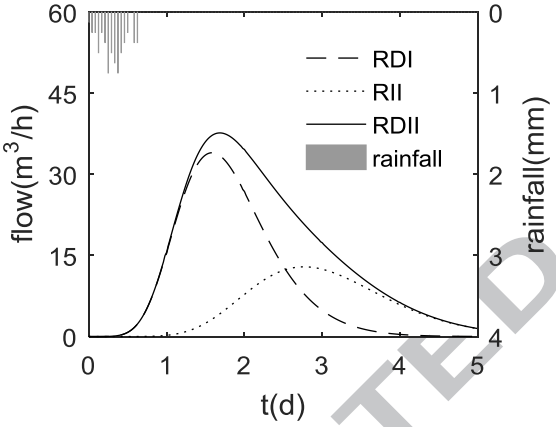
Figure 2. Spectrum analyses of flow and conductivity data. (a) Dry weather flow and model fit using second-order and fifth-order Fourier series; (b) Spectrum analysis of flow data; (c) Dry weather conductivity and model fit using second-order and fifth-order Fourier series; (d) Spectrum analysis of conductivity data.



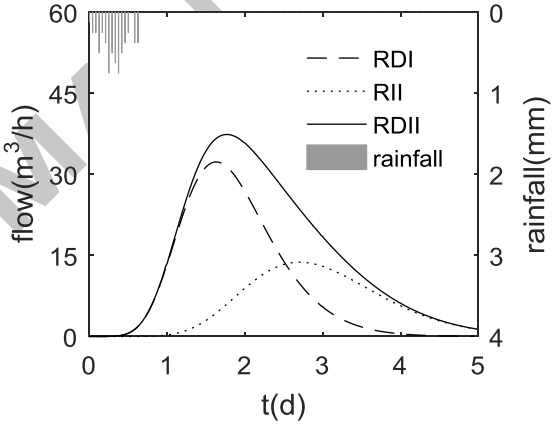
(a)



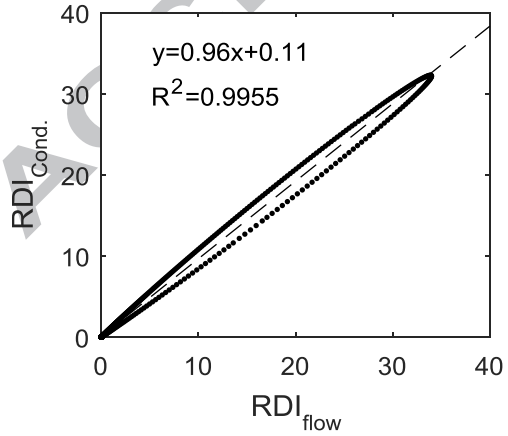
(b)



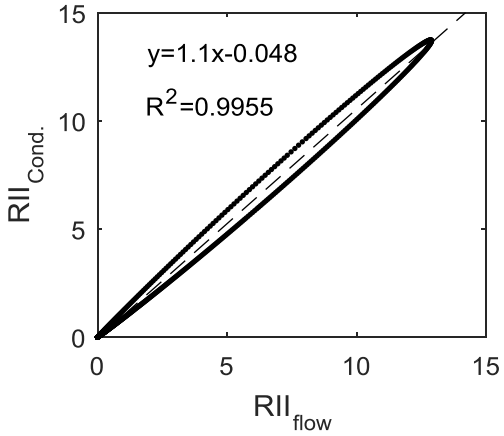
(c)



(d)

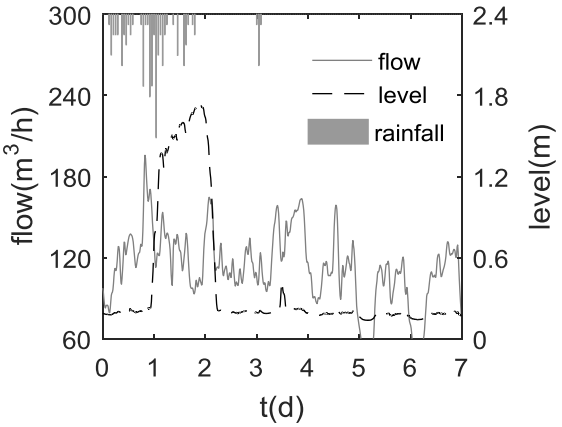


(e)

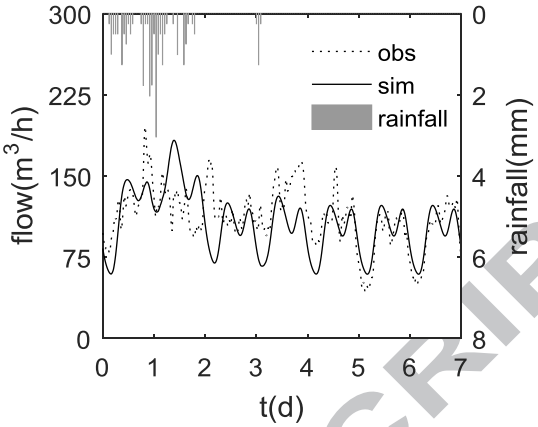


(f)

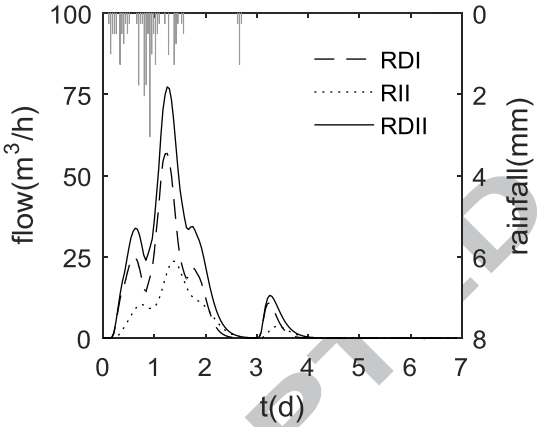
Figure 3. Estimating RDII for the light rainfall event. (a) Simulated and observed flows; (b) Simulated and observed conductivity profiles; (c) Inflow and infiltration flows calculated based on flow data; (d) Inflow and infiltration flows calculated based on conductivity data; (e) Comparison of RDI flow calculated from flow and conductivity data; (f) Comparison of RII flow calculated from flow and conductivity data



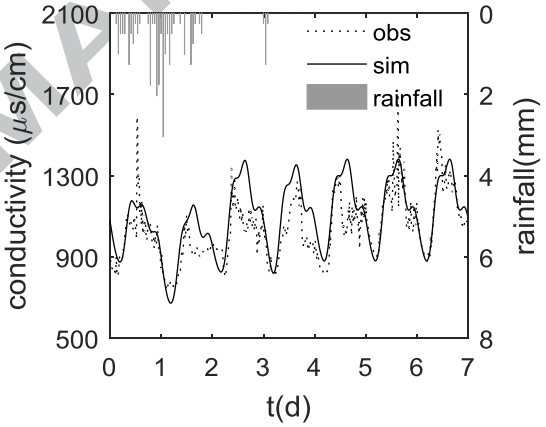
(a)



(b)



(c)



(d)

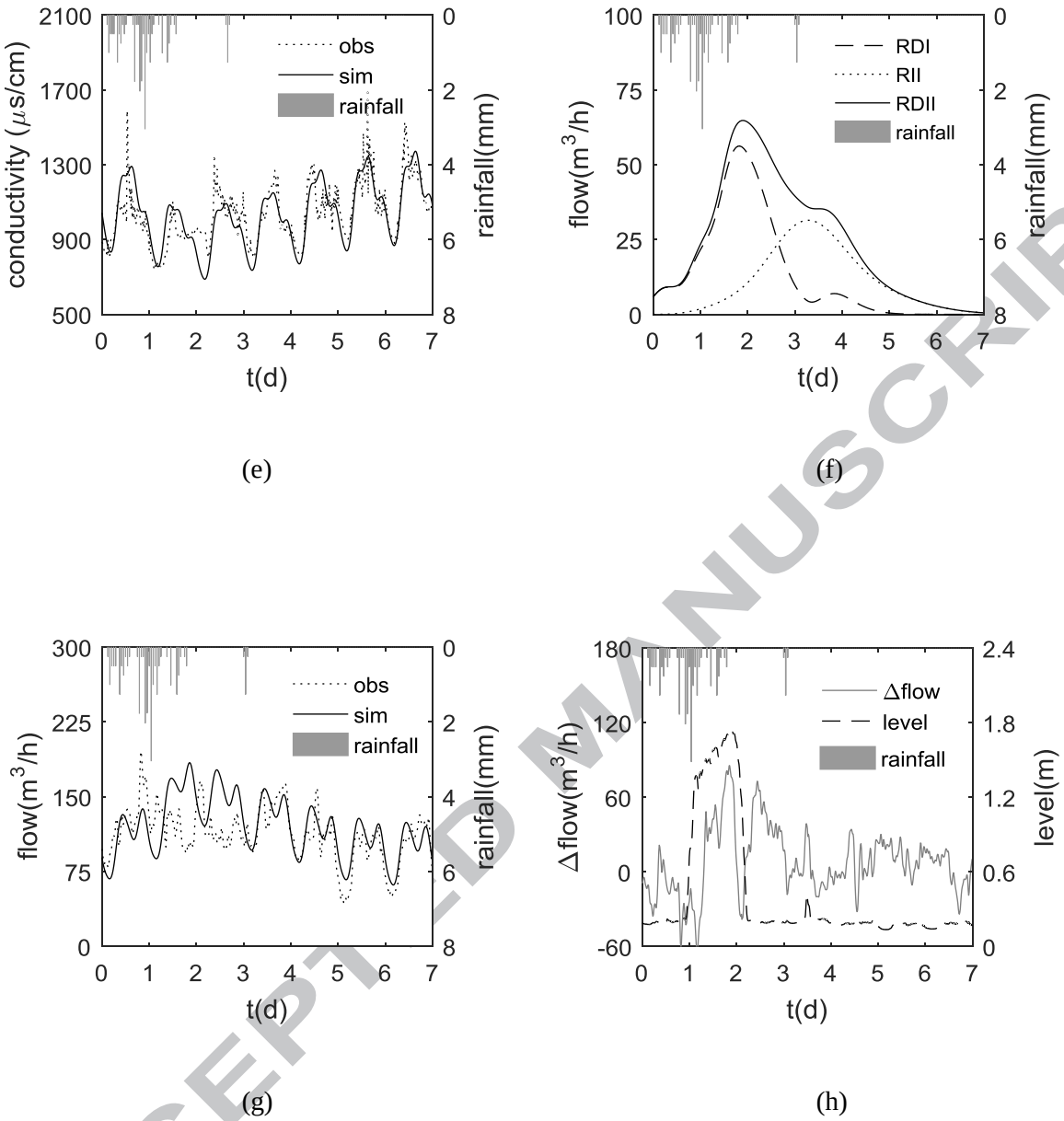


Figure 4. Estimating RDII using flow data ((a) - (d)) and conductivity data ((e) - (h)) for the medium rainfall event. (a) Observed flow and water level in the manhole; (b)Observed and Simulated flow data by fitting flow data; (c)RDII calculated based on flow data; (d)Observed and simulated conductivity data by fitting flow data; (e)Observed and simulated conductivity data by fitting conductivity data; (f) RDII calculated based on conductivity data. (g)Observed and Simulated flow data by fitting conductivity data; (h) Dynamic water level and flow difference (Δflow) between the observed and simulated flow by fitting conductivity data.

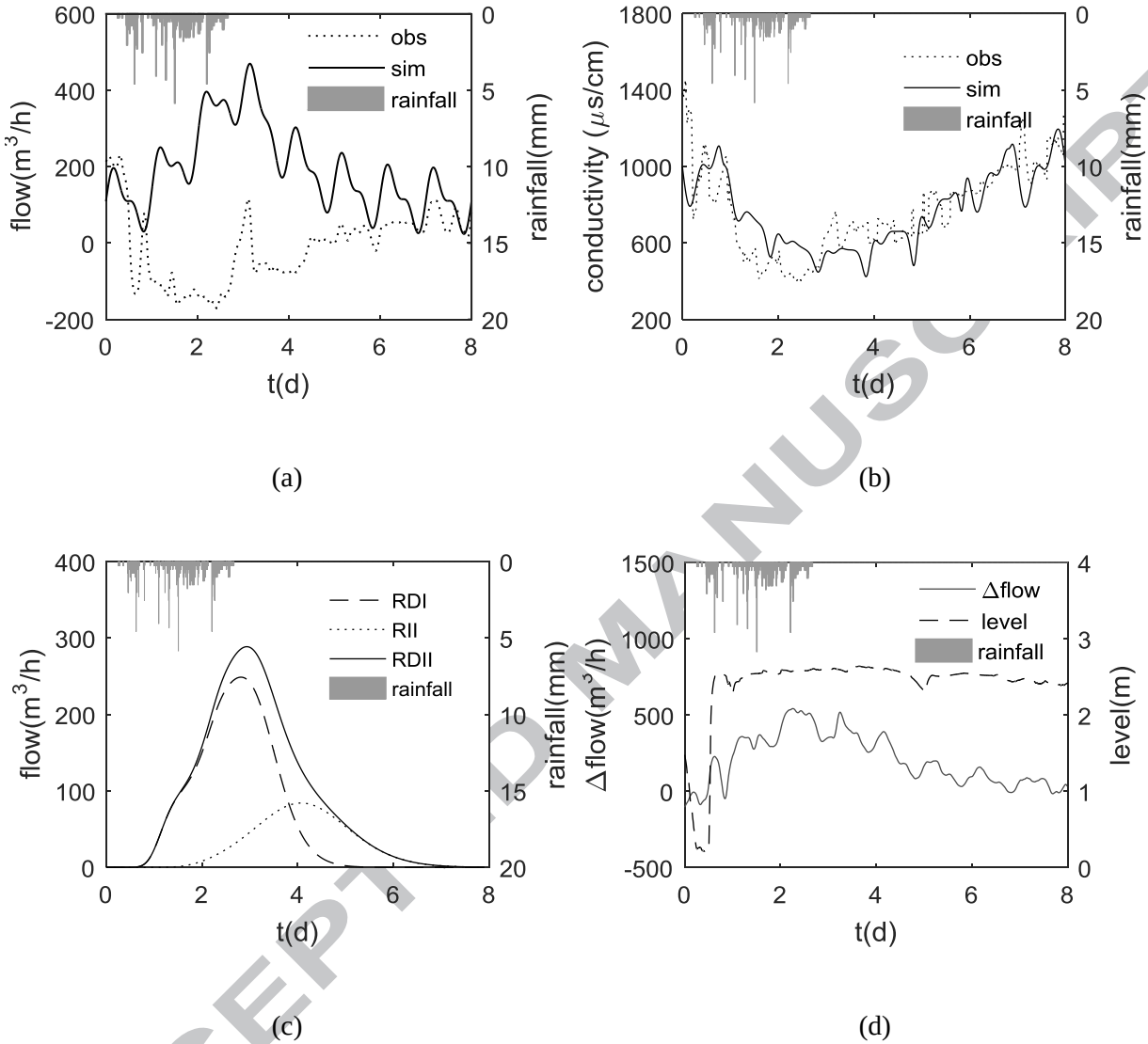
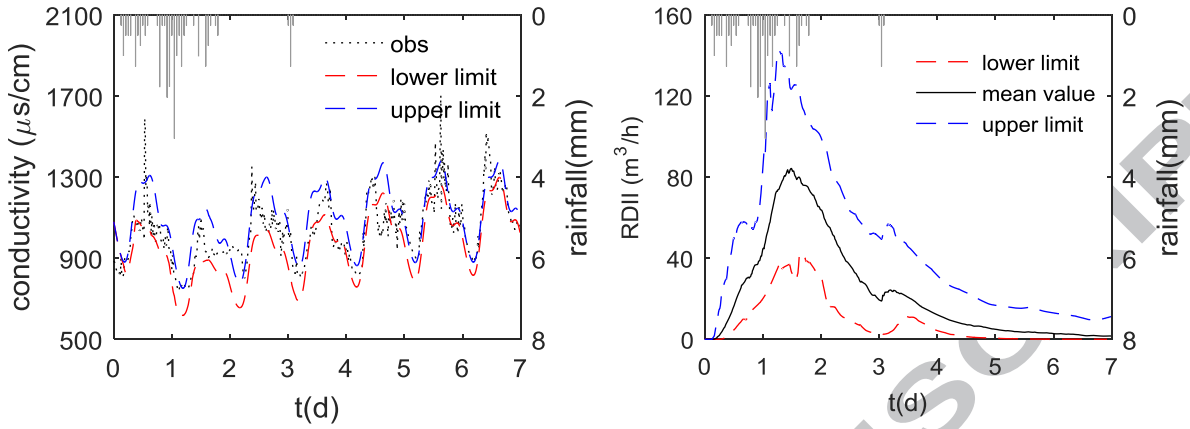


Figure 5. Estimating RDII using conductivity data for the heavy rainfall event. (a) Modeled and observed flow data; (b) Observed and simulated conductivity data; (c) RDII calculated based on conductivity data; (d) Dynamic water level and flow difference (Δflow) between the observed and simulated flow data.



(a) (b)

Figure 6. Model uncertainty analysis of the medium rainfall event. (a) 90% confidence region of simulated conductivity; (b) 90% confidence region of RDII

Quantifying rainfall-derived inflow and infiltration in sanitary sewer systems based on conductivity monitoring

Mingkai Zhang[a], Yanchen Liu[a]*, Xun Cheng[a], David Z. Zhu[b], Hanchang Shi[a], Zhiguo Yuan[a,c]

a. State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China, 100084

b. Department of Civil and Environmental Engineering, University of Alberta, T6G2W2, Edmonton, Alberta, Canada

c. Advanced Wastewater Management Centre (AWMC), The University of Queensland, St. Lucia, Queensland 4072, Australia

*Email: **liuyc@mail.tsinghua.edu.cn**. Tel: +86-10-62796953; Fax: +86-10-62771472

This file includes:

1. Table 1 – Table 3

Table 1. Three typical rainfall scenarios used to evaluate the model

Event	Total rainfall (mm)	Duration (h)	Peak Intensity (mm/hour)	Classification
1	6.25	15	0.75	light

2	28.5	92	3.04	medium
3	178	59	10.92	heavy

Table 2. Optimized parameters and their 90% confidence interval based on flow and conductivity under different rainfall events

rainfall type	calibration data	R_{RDI}	R_{RII}	K_{RDI}	N_{RDI}	K_{RII}	N_{RII}	NS_{flow}	NS_{Cond}
light	flow	22.70% (20.1%,23.8%)	0.132 (10.1%,15.4%)	5.2 (4.04,5.48)	6.8 (6.06,7.71)	6.8 (6.28,7.95)	9.6 (8.49,10.90)	0.32	0.62
light	Cond.	0.221 (20.3%,23.8%)	0.138 (12.3%,15.8%)	5.3 (4.14,5.75)	6.9 (6.07,7.80)	6.7 (6.21,7.92)	9.5 (8.36,10.89)	0.28	0.65
medium	flow	0.062 (4.1%,7.2%)	0.031 (1.1%,4.4%)	2.1 (1.06,2.86)	2.9 (2.05,3.80)	2.5 (1.19,3.91)	4.2 (3.12,4.91)	0.029	0.42

		0.101	0.082	3.8	6.1	5.4	11.2		
medium	Cond.							-0.09	0.54
		(8.2%,11.7%)	(6.2%,9.7%)	(3.08,4.84)	(5.08,6.85)	(4.19,6.86)	(10.11,11.90)		
		0.081	0.035	4.2	6.4	5.8	10.6		
heavy	Cond.							-7.07	0.59
		(6.2%,9.8%)	(2.1%,4.8%)	(3.13,5.78)	(5.09,6.88)	(5.10,6.89)	(9.15,11.83)		

Calibration data (flow or conductivity) was used to determine the model parameters with nonlinear fitting. The 90% confidence interval of the parameters was evaluated with generalized likelihood uncertainty estimation (GLUE) method. RRDI and RRII was the ratio of rainfall entering into the system as RDI or RII. KRDI and KRII was the storage coefficient of RDI and RII process. NRDI and NRII was the number of the reservoirs of the RDI and RII process in the model.

Table 3. Volume (m3) of RDII and SSOs in the three rainfall events

Volume	RDI	RII	RDII	DWF	TOTAL	SSO	Transported
light	1135.2	660.67	1795.9	8425.9	10221.8	0	10221.8
medium	2302.8	1869.6	4172.4	18439.6	22612.0	1207.9	21404.1
heavy	11534.4	4984.0	16518.4	21220.0	37738.4	39732.2	-1993.8*

* minus means inverse flow. In this case, the wastewater flow upstream, contributing to SSO.

The total volume of RDI can be calculated with the equation

$$V_{RDI} = \int_0^t Q_{RDI}(t)\, dt = V_{rain} R_{RDI}$$

Where QRDI(t) is the RDI hydrograph, Vrain is the total rainfall volume. In the same way, the volume of RII can be calculated. Volume of RDII was the sum of RDI and RII. The volume of DWF was the integral of dry weather flow. TOTAL means the sum of DWF and RDII. The volume of SSO is calculated by the integral of difference between modelled and observed flow data, as the equation:

$$V_{SSO} = \int_0^t (flow_{mod} - flow_{obs})\, dt$$

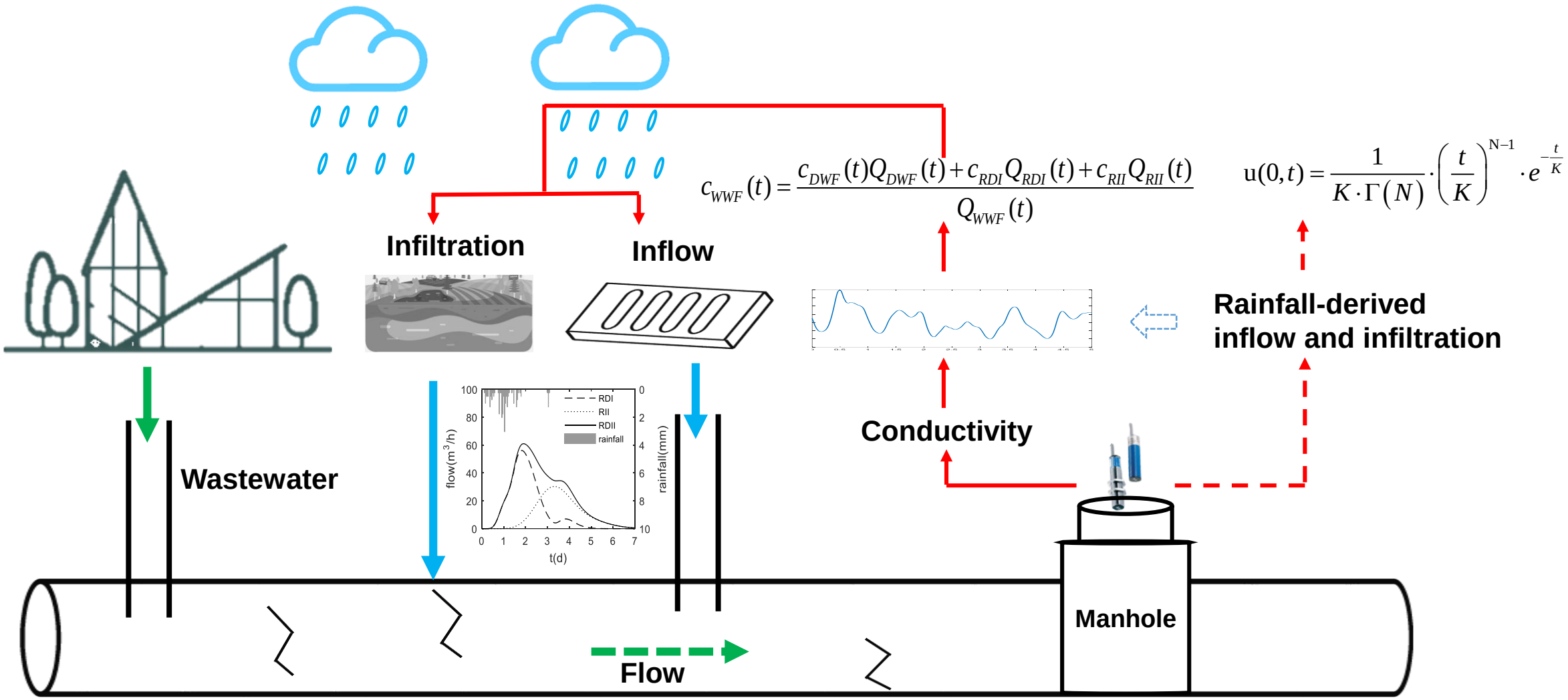
The transported volume was the integral of the observed flow data at the monitoring site. In the medium rainfall event, 21404.1 m3 wastewater was transported downstream. In the heavy rainfall event, 1993.8 m3 wastewater flow upstream as part of the SSO.

> A novel pollutograph method was proposed for estimating rain-induced inflow and infiltration

> The inflow and infiltration was decomposed separately based on the conductivity monitoring

> The method was successfully applied to a real-life case study for three typical rainfall events

> The method exhibits distinct advantages in estimating RDII and overflow simultaneously



$$c_{WWF}(t) = \frac{c_{DWF}(t)Q_{DWF}(t) + c_{RDI}Q_{RDI}(t) + c_{RII}Q_{RII}(t)}{Q_{WWF}(t)}$$

$$u(0,t) = \frac{1}{K \cdot \Gamma(N)} \cdot \left(\frac{t}{K}\right)^{N-1} \cdot e^{-\frac{t}{K}}$$