Tryptophan-like fluorescence as a fingerprint of dry-weather misconnections into storm drainage system

Hailong Yin1,2, Yue Wang1,2, Yang Yang4, Jingshui Huang5* and Zuxin Xu1,2,3*

Abstract

Background: Inappropriate dry-weather misconnections into storm drainage system are a demanding environmental problem worldwide, which leads to unexpected dry-weather discharge into surface waters. It often costs a large amount of manpower and resources to identify the source of misconnections and estimate its contributions. In this study, we evaluated the possibility of quantifying proportional source contribution in a storm drainage system with dry-weather misconnections from domestic sewage and river water inflow, using rapid and low-cost fluorescence spectroscopy methods. For this purpose, samples of both misconnection sources and outflows of storm drainage system were collected and analyzed in a downtown catchment of Shanghai, China.

Results: Results showed that fluorescent peak intensity of tryptophan-like T1 in domestic sewage (802 ± 126 a.u.) was significantly higher than that in urban river water (57 ± 12 a.u.), while fluorescent peak intensities of tryptophan-like T2 in urban river water (732 ± 304 a.u.) was much higher than that in domestic sewage (261 ± 64 a.u.) due to increased algal activity in the local river and upstream inflow chemistry. However, only peak T2 passed the conservative behavior test in the incubation experiments, which could be used as a fingerprint for quantitatively identifying the misconnections. We further developed a Bayesian fluorescence mass balance model (FMBM) to infer the percentage of dry-weather misconnections into the storm drainage system as a function of fluorescence intensities of peak T2 in the samples of sources and outflow. It was found that the maximum posteriori probability estimate of the percentage of river water intrusion into the storm drains was up to 20.8% in this site, which was validated by the results of on-site investigation.

Conclusion: Our findings implied that in situ fluorescent sensors and Bayesian FMBM for the fingerprint fluorescence peak could be applied to fast track inappropriate dry-weather misconnections into storm drainage system qualitatively and quantitatively with low costs.

Keywords: Storm drainage system, Dry-weather mismatch, River water intrusion, Fluorescence spectroscopy, Bayesian mass balance model

Background

In order to mitigate overflow pollution of combined sewer systems, separate stormwater system has been introduced since the 1970s, which is designed to deliver clean rain or storm water to the surface water system only [1]. However, the stormwater outfalls can become polluted, for example, when foul water outlets from residential or industrial premises are inappropriately connected to storm water system [2–5]. Such phenomenon leads to the release of untreated sewage into receiving waters, inducing urban water pollution. Additionally, there is also unexpected river water intrusion into storm drains, placing increased burdens on conveyance...
and the following hydraulic overloading of storm pipe network. This may trigger the storm pumps operation on dry-weather days and result in more serious urban river pollution, e.g., river’s black and odorous occurrence [6–8]. Given the mixed surface water and sewage water inputs, there will inevitably be difficulties in specifically distinguishing between misconnection-derived sewage and surface water sources by monitoring and analyzing a storm drain outfall.

Use of chemical tracers is a promising method to detect misconnected source flow contributions into storm drains. Specifically more and more researchers have shed light on chemical or biological marker species to identify and quantify sewage source discharge. For example, many researchers demonstrated that PPCPs, artificial sweeteners and inorganic ions such as potassium, chloride and sodium, could serve as promising markers of sewage contamination [9–16]. However, no published literatures were found concerning the use of markers in identifying surface water intrusion into the storm pipes. Additionally, current use of above chemical or biological markers is still less easily automated for higher throughputs especially in a catchment-scale assessment.

Compared with the chemically and biologically based analytical methods, fluorescence spectroscopy has the advantages of high sensitivity, simple measurement, fast analysis, low sample size request and no secondary pollution, which provides both qualitative and quantitative information of dissolved organic matter (DOM) in natural and engineered systems [17–20]. Three-dimensional excitation–emission matrix (EEM) fluorescence has been widely used to quantitatively characterize DOM and to follow its dynamics in different aquatic environments, e.g., in rivers [21–25], estuaries or seawaters [26–28], lakes [29–32], wastewater effluents [33] and groundwater [34]. Such studies concluded that fluorescence characteristics of DOM varied in different water bodies. Thus, by examining them, it is possible to distinguish different sources of DOM and determine the presence of different sources in water bodies of concern.

In urbanized areas, while an increasing number of studies had characterized and traced the sources of DOM using fluorescence spectroscopy to explore the influences of urban landscapes and environmental factors on the urban waterbodies, few related studies focused on the DOM fluorescence characterization of sewer overflows, a fact that may be ascribed to the complexity of wet-weather and dry-weather pollutants. Chen et al. [35] investigated the changes of DOM characteristics and yields in the wet-weather overflow in response to the anthropogenic influences of paved runoff and sanitary sewage inputs. However, no published literatures were found regarding the dry-weather misconnections into storm drainage system, a phenomenon arising from the complexity of mixed non-stormwater inputs.

The aim of this study was to qualitatively and quantitatively identify contributions from different sources of dry-weather misconnections into storm drainage system using fluorescence detection. The derived outcomes are expected to inform a methodology that can be directly applied to fast diagnose the in situ dry-weather misconnections, including the quantification of surface water intrusion into the storm drains.

**Materials and methods**

**Site description**

The study site is a typical high-density urbanized area (approximate 270 capital/ha) in Shanghai’s downtown area, with two semiconductor manufacturing enterprises in the south-west (Fig. 1). Surrounded by three tidal rivers (i.e., the Puhuitang, Shang’aotang and Caohejing river), this area is served by separate sewer and storm drainage system that was built in 1980s, with a coverage of 374 ha. There is only one terminal outfall for the storm drainage system, where a storm pump station has been constructed to convey the surface road runoff in wet-weather days to the receiving waters. However, non-storm water flows including domestic sewage as well as surface water also find their ways into storm drains.

On-site investigation showed that there were potential 99 misconnections between river and storm drains in this catchment [6]. On dry-weather days, when surface water level is higher than the water level of terminal outfall (i.e., during high tide of the surrounding tidal rivers), river water intrusion into storm drains occurs and may even trigger storm pumps operation, delivering untreated sewage into the nearby rivers.

**On-site sampling methods**

Sampling activities were conducted during 16th November, 2011 to 31st January, 2013. For the potential misconnected entries as well as the dry-weather catchment outflow, sampling methods were described as below.

**Sampling of domestic sewage**

Sewage samples were collected from the sewer outlet of one residential community, which was inappropriately connected to the storm drain (see Fig. 1). An automatic sampler (ISCO 6712C, Teledyne, Lincoln, Nebraska, USA) was used to collect the water sample by choosing the “1-h equal interval sampling” program, and 24 h was set as a sampling cycle. The mixture of the 24 water samples was used for each sampling activity. All together 18 sampling activities were conducted and correspondingly 18 composite sewage samples were obtained.
Sampling of river water

The river water sampling sites were located in the three tidal rivers surrounding this catchment, i.e., Puhuitang, Shang’aotang, and Caohedgeing Rivers, where six sampling sites were selected (see Fig. 1). At each sampling site, water samples were collected 0.5 m below the water surface in the midstream of the watercourse. For each sampling event, one river water sample was collected at each monitoring station. Totally 15 sampling activities were performed and 90 river water samples were obtained.

Sampling of catchment outflow

For the catchment outflow, the antecedent dry weather period of 48 h was chosen before the start of each sampling campaign. Therefore, the contribution of potential storm water runoff to the catchment outflow could be excluded. Water samples were extracted from the terminal wet-well using an ISCO 6712 automatic field sampler. Like the sewage samples, outflow samples were collected every hour by choosing the “1-h equal interval sampling” program, and 24 samples were obtained within each day. Finally, the mixture of 24 water samples was used for the experiment analysis for each sampling activity. Totally 12 sampling activities were performed.

Sample analytical methods

All of the collected water samples were stored in the pre-washed and dried sampling bottles. After being transported to the laboratory, the samples were immediately filtered through a 0.45-μm microporous membrane. The filtered samples were stored at 4 °C in the
dark before analysis. The measurements were completed within 48 h.

**Fluorescence EEM measurement**

Fluorescence detection was carried out using Hitachi three-dimensional fluorescence spectrophotometer F-4600. Each sample was placed into a 1-cm quartz cuvette for detection and blank corrected. The scanning parameters were set as follows: the photomultiplier tube (PMT) voltage was 700 V, and the scanning speed was 12,000 nm min⁻¹. The ranges of excitation wavelength (Ex) and emission wavelength (Em) were 200–450 nm and 250–550 nm, respectively, with a step size of 5 nm. Some protocols followed in this study included (1) warming up fluorometer lamps for 1 h before beginning measurements; (2) turning on the air conditioner in the instrument room to maintain a constant room temperature of 20 °C; (3) taking samples out of the refrigerator in advance and warming them up until room temperature for measurement.

**UV–visible absorbance spectroscopy measurement**

UV–visible absorbance spectrum was measured from 200 nm to 800 nm with a spectrophotometer (UV-2550, Shimadzu, Japan) in a 1-cm quartz cuvette. Milli-Q water was used for blank correction.

**EEM correction procedures**

Several correction procedures were implemented to adjust the EEM data. First, the excitation and emission data were corrected for instrument-specific response. Scans of blank and Milli-Q water EEMs were collected daily by the lab to check the instrument performance stability. The spectral correction factors were supplied by the instrument manufacturers. Second, the EEM response of Milli-Q water was subtracted from sample EEMs. Third, the UV–visible absorption spectra were used to correct the EEM data for inner filter effects. Correction factors were generated from sample-specific ‘best-average’ absorption scan [36].

**Peak-picking analysis of EEM**

Peak-picking is a viable analysis technique used for the development and use of a real-time tool, which can be directly tied to custom sensors available today [19, 22, 33]. Commonly there are four kinds of fluorescence peaks observed in freshwater aquatic samples [19, 37, 38]. These have been classified as follows: peak A—fulvic-like ($\lambda_{ex/em} = 237–260/400–500$ nm); peak C—humic-like ($\lambda_{ex/em} = 300–370/400–500$ nm); peak $T_1$—tryptophan-like ($\lambda_{ex/em} = 275/340–350$ nm) and peak $T_2$—tryptophan-like ($\lambda_{ex/em} = 225–237/340–381$ nm); peak $B_1$—tyrosine-like ($\lambda_{ex/em} = 275/300–305$ nm) and peak $B_2$—tyrosine-like ($\lambda_{ex/em} = 225–237/309–321$ nm) [19, 22, 37].

**Lab incubation experiment setup**

Besides on-site sampling to detect fluorescence components of the study site, in order to determine the conservative behavior or not of the detected DOM fluorescence, lab experiments were also conducted. In this study site, the recorded water level at terminal wet-well fluctuated around 2.3 m on dry-weather days, and the pre-set alarm water level to start storm pumps was 2.6 m in order to prevent pipe surcharging [39]. Under this circumstance, the dry-weather flows accommodated within the storm pipes occupied 85% storage capacity of the whole storm pipe network or higher, based on our developed hydrodynamic model [39]. Therefore, the storm pipes were represented as almost completely filled tubes, which were prone to development of anoxic or even anaerobic conditions [40, 41]. Additionally, featuring a backwater state, the water flowing velocity within the storm drains was very low, with estimated data below 0.1 m/s and approximate travel time of 24 h [39].

In order to simulate the actual environment in the sewer, a continuous darkroom anaerobic condition for more than 24 h was used in the laboratory incubation experiment. Specifically, sewage or river water samples were, respectively, injected into a series of 250-mL borosilicate glass bottles, which were sealed with parafilm and wrapped with foil to represent anoxic and dark environment. For each experiment, the duration was equal or greater than 24 h. The fluorescence was measured before the experiment and at 24-h interval during the incubation experiment.

The experiments were performed both in the spring and winter seasons to determine the effect of environment temperature and associated microbial activities on the conservative behavior of detected DOM components. In detail, the incubation experiments were conducted for the collected sewage samples in May, 2012 (i.e., the period of spring season) and January, 2013 (i.e., the period of winter season). The experiment temperatures were controlled close to the environment temperatures in the two seasons, namely with 25 °C for the incubation experiment in May and 5 °C for the one in January.

**Fluorescence mass balance model**

Hudson et al. pointed out that fluorescence intensities of the peaks in the specific regions of the EEM could correspond to the concentrations of the compounds...
represented by the peaks [38]. Goldman et al. further found a linear response in fluorescence peak intensity of individual peaks (e.g., peaks A, T and C) for the mixture of different source flows [22]. In accordance with these criteria, for a fluorescent component exhibiting conservative behavior, i.e., no obvious chemical or biological degradation within the storm drains, the fluorescence mass balance equation can be written as:

\[
F_{\text{river}} \times Q_{\text{river}} + F_{\text{sewage}} \times Q_{\text{sewage}} = F_{\text{outflow}} \times Q_{\text{outflow}},
\]

where \( F_{\text{river}} \) is the fluorescence intensity of the peak for river water samples; \( Q_{\text{river}} \) is the flow rate of river water intrusion into the storm drains; \( F_{\text{sewage}} \) is the fluorescence intensity of the peak for domestic sewage samples; \( Q_{\text{sewage}} \) is the flow rate of domestic sewage misconnected into the storm drains; \( F_{\text{outflow}} \) is the fluorescence intensity of the peak for catchment outflows; \( Q_{\text{outflow}} \) is the flow rate of the storm drain outfall on dry-weather days.

Assuming that the percentage share of river water intrusion into the storm drains on dry-weather days is \( r \) (%), the following formula could be derived from the above mass balance model as:

\[
F_{\text{river}} \times Q_{\text{outflow}} \times r + F_{\text{sewage}} \times Q_{\text{outflow}} \times (1 - r) = F_{\text{outflow}} \times Q_{\text{outflow}}.
\]

Divided by \( Q_{\text{outflow}} \), the model of fluorescence intensities and the percentage share of river water intrusion into the storm drains can be generated as:

\[
F_{\text{sewage}} + r \times (F_{\text{river}} - F_{\text{sewage}}) = F_{\text{outflow}}.
\]

A Bayesian method was employed here to infer \( r \), which is capable of accounting for the measurement errors of fluorescence intensities and estimate the uncertainty of the percentage share [42]. A widely used Markov Chain Monte Carlo (MCMC) approach was integrated with the model using DREAM [43]. Normal distribution tests were performed with the measurements of \( F_{\text{river}}, F_{\text{sewage}} \) and \( F_{\text{outflow}} \) to obtain the means and standard deviations of their distributions, respectively. \( F_{\text{river}}, F_{\text{sewage}} \) and \( r \) were defined as the input parameters in the model. The prior information of \( F_{\text{river}}, F_{\text{sewage}} \) was provided by the estimates of their normal distribution parameters, respectively. A uniform probability density function (PDF) between 0 and 1 was considered for the prior distribution of \( r \). \( F_{\text{outflow}} \) is defined as the output of the model. The model parameter inferences were based on the log-likelihood function:

\[
\log L = \frac{-M}{2} \log (2\pi) - \sum_{i=1}^{M} \log \sigma_i - \frac{1}{2} \sum_{i=1}^{M} \frac{1}{\sigma_i} (F_{\text{obs}} - F_{\text{sim}})^2,
\]

where \( i \) and \( M \) donate the \( i \)th measurement and the number of measurements, respectively; \( F_{\text{obs}} \) and \( F_{\text{sim}} \) are observed and simulated \( F_{\text{outflow}} \), respectively; \( \sigma \) denotes the standard deviation of the Gaussian distribution of observed \( F_{\text{outflow}} \).

### Results and discussion

The EEMs from samples collected from the domestic raw sewage and urban river water have been evaluated and show 3 peaks (\( T_1, T_2, B_1 \)) with the strongest fluorescence intensities (Fig. 2). More information concerning the EEM spectra of these samples collected is listed in Additional file 1: Figs. S1 and S2 for domestic sewage and river water, respectively.

#### EEM characteristics of domestic sewage

Figure 2a shows that basically there are two main fluorophores in the spectrum of domestic sewage. The first fluorescence intensity peak is at Ex/Em of 275/350 nm, which corresponds to tryptophan-like peak \( T_1 \) components. The second fluorescence intensity peak is at Ex/Em of 225/350 nm, which is associated with tryptophan \( T_2 \) fluorescent components. The detected tryptophan-like fluorescent components are mainly protein-like substances, which may be derived from the human activities such as human excreta, food residues and cooking oils [44]. Based on Fig. 2b, the averaged fluorescent peak intensity of tryptophan-like \( T_1 \) and tryptophan-like \( T_2 \) material are 802 a.u. and 261 a.u., respectively, with an average \( T_1/T_2 \) ratio of 3.1. Tryptophan-like \( T_1 \) materials are related to soluble microbial products, and tryptophan-like \( T_2 \) materials are associated with proteins of aromatic structures [45].

In addition, a distinct narrow band occurs at \( \lambda_{\text{ex}} = 330–375 \) nm for \( \lambda_{\text{em}} = 410–450 \) nm. This is can be explained by the fluorescent whitening agents from the grey water such as dishes and clothes washing, which also indicates the contribution of the domestic sewage to some extent. Considering domestic sewage is transported through pipelines and has no direct contact with soil, the possibility of humus from the soils is small. Therefore, the humic-like fluorescent components in domestic sewage are mainly derived from the humic acids originated in tap water supply [46]. From this perspective, the intensity of humic-like substances is much less pronounced than the protein-like ones in the domestic raw sewage samples.

#### EEM characteristics of river water

As indicated in Fig. 2c, there are two distinct fluorescence peaks in the EEMs of the river water samples, with the Ex/Em at 230/340 nm and 275/305 nm, respectively. The peak at EEM of 230/340 nm corresponds to tryptophan-like \( T_2 \) materials, and the peak at EEM of 275/305 nm
corresponds to tyrosine-like B₁ materials. Both the tryptophan-like T₂ and tyrosine-like B₁ belong to protein-like materials.

As for the humic or fulvic substance, a fluorophore was also generated at with the Ex/Em at 250/435–455 nm. However, its fluorescence intensity was not as high as that of tryptophan-like T₂ and tyrosine-like B₁ substances in this case. Previous studies by Hudson et al. [38] showed that in clean rivers, peaks C₁, C₂ and A predominated. This is because DOM originating from clean river water is dominated by natural organic matter from plant material, whereas sewage-derived DOM is dominated by organic matter originating from microbial activity. By contrast, our study showed that tryptophan-like or tyrosine-like substances, instead of humic-like or fulvic-like substances, were the indication of polluted urban river water. Therefore, with increasing urbanization and anthropogenic activities, the fluorescent signature of urban waters changed from being humic-rich to protein-rich with peaks T and B.

Figure 2d further shows the measured fluorescent peak intensities of urban river water at six sampling stations. Among the six sites, the S1 monitoring site is located in Shang’aotang River, where the sluice gates are installed (see Fig. 1). As a result, Shang’aotang River is separated from other rivers including Puhuitang River and Caohejing River, which are polluted by the dry-weather as well as wet-weather outflow from the storm drains of this catchment. By contrast, no storm drain outfalls are directly connected to the Shang’aotang River. Therefore, water quality at S1 station was better than that of the other five sampling stations. However, the fluorescence intensities of tryptophan-like T₂ and tyrosine-like B₁ at the S1 station were not considerably different from those at other five stations. This showed that fluorescence property of urban rivers was determined by materials more than sewage-derived DOMs in this catchment.
Differences of EEM characteristics between domestic sewage and urban river water

Our study showed that domestic sewage featured most strongly tryptophan-like peaks T₁ and T₂, whereas urban river water featured tryptophan-like peak T₂ and tyrosine-like peak B₁. Both domestic sewage and urban river water had the characteristics of tryptophan-like T₂ fluorescence; however, it was interesting that tryptophan-like peak T₂ in the urban river water was more intense than that in the untreated domestic sewage.

Figure 3 shows that organic matter in the investigated urban rivers exhibits seasonally different fluorescence intensities over the sampling period. Tryptophan-like peak T₂ as well as tyrosine-like peak B₁ was observed to be more intensive in winter as compared to that in summer. This was related to elevated river water levels due to stormwater runoff discharge into rivers and the resulting dilution effect in summer to some extent. However, despite being diluted by stormwater runoff, fluorescence intensities in peaks T and B of river water samples were still higher than those of sewage samples. The explanation is that peaks T and B are related to microbial activity [47] and may be transported into a system (allochthonous) or be created by microbial and biological activity within a system (autochthonous). Intensive tryptophan-like T₂ and tyrosine-like B₁ are related to biological activity particularly in areas of high primary productivity, that is, surface waters with phytoplankton or algal activity. Under this circumstance, tryptophan-like T₂ and tyrosine-like B₁ fluorescence may be present as ‘free’ molecules or else bound in proteins or humic structures of algae cells and their remnants in river waters.

In Shanghai area, majority of aquatic plants grow in the spring and summer season and decay in the winter.
season [48]. Decomposition of aquatic plants contributed to the intensified tryptophan-like T\textsubscript{2} and tyrosine-like B\textsubscript{1} substances. Such explanation could be strengthened by recently reported protein-like fluorescence in Taihu Lake, China, where the tryptophan-like components of DOM were significantly higher in winter than in summer and autumn, due to the degradation of phytoplankton [32]. Moreover, at lower water temperature, dinoflagellates or diatoms are the predominant algal species. Correspondingly, there is a good correlation between peaks T and B materials produced by such algal species [49], as proved by this study. In this case, a strong linear relationship between the fluorescence peak intensities of tryptophan-like T\textsubscript{1} and tyrosine-like B\textsubscript{1} decreased over time rapidly within the experimental duration. Fluorescence intensity of peak T\textsubscript{1} and B\textsubscript{1} decreased by 70.8% and 66.1%, respectively, on average after 24 h. In this process, fluorescence intensities of other fluorescent components did not increase correspondingly (see Additional file 1: Fig. S3). This result indicated that the fluorescence intensities of the tryptophan-like T\textsubscript{1} and tyrosine-like B\textsubscript{1} substances converted into non-fluorescent substances with time. As discussed above, the tryptophan-like T\textsubscript{1} and tyrosine-like B\textsubscript{1} component in domestic sewage are considered to be microbial by-products, and therefore changes in their fluorescence intensities are related to microbial metabolism. For this scenario under the lab temperature of 25 °C, active microbial activity dramatically degraded the fluorescent substances of T\textsubscript{1} and B\textsubscript{1}. Reynolds concluded that microbial activity, measured by oxygen depletion in the BOD\textsubscript{5} test, correlated well with the T\textsubscript{1} fluorescence intensity of raw sewage [54]. Therefore, T\textsubscript{1} is presented as a bioavailable substrate. However, for the incubation experiments conducted in January, the fluorescence signature of tryptophan-like T\textsubscript{1} and tyrosine-like B\textsubscript{1} component in domestic sewage almost remained unchanged with time. This can be explained by inhibited microbial activity under lower lab temperature (e.g., the temperature of 5 °C), leading to relatively stable fluorescence peak intensities. Changes of fluorescence signature of tryptophan-like T\textsubscript{2} with experimental duration was also probed for domestic sewage and surface water samples. Figure 4c shows that tryptophan-like T\textsubscript{2} in domestic sewage samples exhibits conservative behavior, independent of environment temperature. This can be explained by that protein-like peak T\textsubscript{2} is associated with proteins of aromatic structures, which are less likely to be degraded as compared to protein-like peak T\textsubscript{1} of soluble biodegradable microbial products [45]. A similar finding was demonstrated in experiments of river water samples, in which fluorescence intensities of tryptophan-like T\textsubscript{1} substances did not change within the experimental duration of 72 h as well (see Fig. 4d). As discussed above, intensified peak T\textsubscript{2} in the river water is influenced by upstream inflow featuring chemical tracer of dispersant MF, a kind of sulfonated naphthalene formaldehyde condensates from dyeing process. This tracer would contribute to the non-biodegradability of fluorescent substances in the river water. Additionally, it was reported that protein-like peak T\textsubscript{2} in river water was given by compounds of aromatic rings like polycyclic aromatic hydrocarbons (PAHs), exhibiting bio-refractory behavior as well [19].
Based on above discussion, tryptophan-like $T_2$ can be used as an ideal marker to quantify percent wastewater and urban river water with inappropriate entry into storm drains on dry-weather days, which features conservative characteristics in both of the misconnected source types. Thus, the uncertainty of fluorescence mass balance model arising from chemical or biological reaction within the storm pipes could be minimized as much as possible.

**Determination of percent surface water and wastewater into the storm drains**

To determine the percent surface water into storm drains, the fluorescence mass balance model was set up with fluorescence intensity measurements of the samples from the river water, domestic sewage and catchment outlet. Information concerning detected EEMs for all of the catchment outflow samples is provided in Additional file 1: Fig. S4. It was found the center positions of the two identified peaks $T_1$ and $B_1$ of the three flow types were significantly different, whereas the center positions of peak $T_2$ among these types were very concentrated. The identified center positions of peak $T_2$ were at Ex/Em 225/350 nm, 230/340 nm and 225/350 nm for wastewater, river water and catchment outflow, respectively. The slight shift of the peak $T_2$ reflected the inevitable difference between the DOM compositions in different types of sources. Comparison of fluorescence peak $T_2$ among domestic sewage, river water and catchment outflow is presented in Fig. 5.

As seen in Fig. 5, for both summer and winter season periods, the fluorescence intensities of peak $T_2$ in the catchment outflows were higher than those in the domestic sewage, due to occurrence of river water intrusion. For the catchment outflow, generally fluorescence intensities of tryptophan-like $T_2$ in winter were higher than that in summer. There are two essential criteria for being selected as a tracer for misconnection source...
apportionment. First, the concentration or level of the tracer should be significantly different between end members. Second, the tracer behaves conservatively during the mixing and transport processes. The significant differences in fluorescence intensities of peak T$_2$ of two sources (Fig. 5) and its conservative behavior (Fig. 4c, d) made it fit the essential criteria for being selected as a tracer. Therefore, fluorescence spectrometry has the promise for indicating the presence of river water intrusion into storm drains as a rapid, reagentless technique that requires little sample preparation.

Using the maximum fluorescence intensities at each peak T$_2$ center of domestic sewage, river water and catchment outflow to get $F_{\text{river}}$, $F_{\text{sewage}}$ and $F_{\text{outflow}}$ in Eq. 3, the Bayesian inference results of the percent river water into the storm drains for the two season periods are shown in Fig. 6. It shows that the maximum a posteriori probability (MAP) estimate of the percent is 12.7% and 20.8% in summer season and winter season, respectively. Considering the center shifts of the peak T$_2$, the Bayesian estimation was also done using the fluorescence intensities of peak T$_2$ at Ex/Em of 230/340 nm for all the three types of samples. The percentage share of river water backflow into the storm drains in summer season and winter season was 9.5% and 22.5%, respectively (Additional file 1: Fig. S8), a difference of −3.2% and 1.7% for summer and winter as compared to the results using maximum fluorescence intensities (Fig. 6). This is within the acceptable range of the source appointment estimate. The small differences can be explained by two reasons. First, the shifts of the chosen peak T$_2$ among the three types were small enough, i.e., one interval of excitation step and two intervals of emission step. Second, the differences in the fluorescence intensities of the chosen peak T$_2$ were large enough among the three types. This further strengthens that peak T$_2$ has great potential for tracking and quantifying dry-weather misconnections into storm pipes.

Theoretically, the amount of river water inflow is determined by the pressure head between surface water and storm drains; the higher the river water level, the larger the amount of river water inflow. Therefore, higher MAP in winter was associated with lower water level of storm drains of non-flood season period. However, uncertainty of Bayesian result in summer was higher than that in winter. This is possibly due to frequent rainfall events and intense precipitation during summer season. Storm water discharges led to significant variations of fluorescent substances in the three surrounding rivers (see Fig. 5), resulting in the increased uncertainty of percent river water intrusion.

The Bayesian inference results were validated with the estimated time-series river water inflow in winter season in this case. Figure 7 shows the measured historical real-time water levels between surrounding rivers and terminal outfall within 1 month of winter season period (i.e., Dec., 2008). Using the mathematical function between river water inflow and pressure head established in this catchment (i.e., $Q_{\text{river}} = 0.104 \times \Delta h^{0.5}$, $Q_{\text{river}}$ is river water inflow and Δh is the sum of real-time water pressure head between river water level and terminal wet-well level within 1 day) [6], the daily river water inflow was accordingly determined, as seen in Fig. 7. It was known from this figure that when storm pumps operated to lower the terminal wet-well level to the lowest alarm level, the significantly increased pressure head drives large amounts of river water inflow in a short period of time. For the non-pumping discharge periods, river water

![Fig. 5](image-url) Comparison of fluorescence peak T$_2$ intensities for the samples among river water, domestic sewage and catchment outflow

![Fig. 6](image-url) Maximum a posteriori probability (MAP) estimate for the percentage share of river water inflow into the storm drains on dry-weather days: a summer season, b winter season
inflow is much smaller as compared to that under the pumping discharge periods. In this figure, the daily river water inflow ranged from 145 to 11,705 m³/day, with an av data of 4261 m³/day.

Our previous study showed that catchment outflow on dry-weather days was approximately 19,350–21,600 m³/day [6]; therefore, maximum likelihood value of river water inflow in winter season was about 4025–4493 m³/day based on MAP estimate for percent surface water intrusion. The daily averaged river water inflow based on Fig. 7 coincided well with the MAP estimate, with a relative error less than 10%. This demonstrates that FMBM is robust in quantifying surface water intrusion into storm drains.

Environmental implications
The use of fluorescence to predict the presence and quantity the dry-weather misconnection into a storm drainage system has significant environment implications. Recent advancements in sensor technology and the development of reliable and specific fluorescence probes have increased our ability to monitor organic matter characteristics in near real-time. The application of real-time data and fluorescence mass balance model would allow water managers to track and quantify the dry-weather misconnection without high costs associated with labor-intensive investigations or complex analytical approaches. For example, establishing a fitted analytical function to determine the river water inflow, have to depend on long-term real-time synchronous flow discharge and pressure head data (river water level versus catchment water level) for each catchment. By contrast, if aided with tryptophan-like T₂ sensor and FMBM introduced in this study, administrative managers would be provided with a useful, fast and cheap alternative way for the investigation of dry-weather misconnections into urban storm drainage system.

The study was conducted on polluted urban rivers, whereas there may be dissimilarities in the organic matter of other surface water bodies or misconnected source types. However, we can develop same approaches to identify and quantify illicit discharges, considering custom sensors have already been constructed to capture the key peaks such as the strong microbial fluorescence signal. Although peak-picking method was used in this study for selecting the tracer peak, other EEM analysis methods such as parallel factor analysis (PARAFAC) would also be useful for the determination of the marker peak and designation of sensors for the same purpose. Additionally, the presented fluorescence approach could be combined with other chemical tracers to present an in-depth insight into illicit discharge investigation of various source types from a wider perspective, and also determine the best times to perform other chemical analyses if necessary.

Conclusion
In the present study, fluorescence spectroscopy method was used to detect and quantify urban river water from the dry-weather flows that inappropriately entered into the storm drains. Our findings are listed as follows.

The fluorescence spectra of two misconnected sources including domestic sewage and urban river water

Fig. 7 Estimated daily river water inflow into storm drains based on real-time water level between river and catchment outfall in winter season period
exhibited substantially different characteristics, with the strongest fluorescence peak centers at an excitation/emission wavelength (Ex/Em) of 275/350 nm (tryptophan-like T1 fluorescent substances) for the domestic sewage and 230/340 nm (tryptophan-like T2 fluorescent substances) for the urban river water, respectively.

Results showed that fluorescence peak intensities of tryptophan-like T1 in domestic sewage (802±126 a.u.) was significantly higher than that in urban river water (57±12 a.u.), while fluorescence peak intensities of tryptophan-like T2 in urban river water (732±304 a.u.) was much higher than that in domestic sewage (261±64 a.u.). Seasonal changes in T2 fluorescence intensity observed in the urban river water were associated with the algal activity in the polluted urban river and upstream inflow chemistry.

Lab incubation experiments showed that only peak T2 passed the conservative behavior, which could be used as a fingerprint for quantitatively identifying the misconnections of urban river water intrusion. Using the developed Bayesian fluorescence mass balance model (FMBM), the MAP estimate of the percent river water into the storm drains in a downtown catchment of Shanghai, China (3.74 km²) can be responsible for up to 20.8% in winter. This result has a relative bias less than 10% compared with the averaging values within 1 month of winter season, derived from real-time water pressure head between river and storm drainage system.

Our findings highlighted that the presented Bayesian FMBM approach could be employed to infer the contribution of surface water intrusion into the storm drains. In the future, custom sensors can be installed to real-time analyze the fingerprint fluorescence peak in the storm drainage system. The application of real-time fluorescence data and Bayesian FMBM method should allow urban water manager to track misconnections from both qualitative and quantitative perspective without high costs and implement timely response measures.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12302-020-00336-3.
water recharge system and from non-specific sources. Desalination 273(2):398–404
11. Tran NH, Hu JY, Li JH, Ong SL (2014) Suitability of artificial sweeteners as indicators of raw wastewater contamination in surface water and groundwater. Water Res 48:443–456
12. Tran NH, Li JH, Hu JY, Ong SL (2014) Occurrence and suitability of pharmaceuticals and personal care products as molecular markers for raw wastewater contamination in surface water and groundwater. Environ Sci Pollut Res 21(6):4727–4740
13. Sun Q, Li M, Ma C, Chen XQ, Xie XQ, Yu CP (2016) Seasonal and spatial variations of PPCP occurrence, removal and mass loading in three wastewater treatment plants located in different urbanization areas in Xiamen, China. Environ Pollut. 208(Pt B):371–381
14. Yang YY, Liu WR, Liu YS, Zhao JL, Zhang M, Zhang JN, Jiang YY, Zhang LY, Ying GG (2017) Suitability of pharmaceuticals and personal care products (PPCPs) and artificial sweeteners (ASs) as wastewater indicators in the Pearl River Delta, South China. Sci Total Environ. 590:611–619
15. Xu ZX, Wang LL, Yin HL, Li HZ, Schwlegler BR (2016) Source apportionment of non-storm water entries into storm drains using marker species: modeling approach and verification. Ecol Indic. 61:546–557
16. Yin HL, Xie M, Zhang LY, Huang JS, Xu ZX, Li HZ, Jiang RJ, Wang RC, Zeng X (2019) Identification of sewage markers to indicate sources of contamination: low cost options for disconnected non-storm-water source tracking in stormwater systems. Sci Total Environ 648:125–134
17. Hudson N, Baker A, Ward D, Reynolds DM, Brunsdon C, Carlill-Marquet C, Browning S (2008) Can fluorescence spectrometry be used as a surrogate for the Biological Oxygen Demand (BOD) test in water quality assessment? an example from South West England. Sci Total Environ 389:149–158
18. Henderson RK, Baker A, Murphy KR, Haembla A, Stuetz RM, Khan SJ (2009) Fluorescence as a potential monitoring tool for recycled water systems: a review. Water Res 43:863–881
19. Carstea EM, Baker A, Bioreza M, Reynolds D (2010) Continuous fluorescence excitation–emission matrix monitoring of river organic matter. Water Res 44:5336–5346
20. Yu H, Song Y, Gao H, Liu L, Yao LL, Peng JF (2015) Applying fluorescence spectroscopy and multivariable analysis to characterize structural composition of dissolved organic matter and its correlation with water quality in an urban river. Environ Earth Sci. 73:1563–1571
21. Baker A, Inverarity, R, Charlton M, Richardson S (2003) Detecting river pollution using fluorescence spectrophotometry: case studies from the Ouseburn, NE England. Environ Pollut. 124(1):57–70
22. Goldman JH, Rounds SA, Needoba JA (2012) Applications of fluorescence spectroscopy for predicting percent wastewater in an urban stream. Environ Sci Technol 46:4374–4381
23. Meng F, Huang G, Yang X, Li LQ, Li J, Cao J, Wang ZG, Sun L (2013) Identifying the sources and fate of anthropogenically impacted dissolved organic matter (DOM) in urbanized rivers. Water Res 47:5027–5039
24. Li SJ, Zhang JQ, Guo EL, Zhang F, Ma QY, Mu GY (2017) Dynamics and ecological risk assessment of chromophoric dissolved organic matter in the Yinma River Watershed: rivers, reservoirs, and urban waters. Environ Res 158:245–254
25. Zhao Y, Song KS, Lv LL, Wen ZD, Du J, Shang YX (2018) Relationship changes between CDOM and DOC in the Songhua River affected by highly polluted tributary, Northeast China. Environ Sci Pollut Res 25:2537:1–25382
26. Coble PG (1996) Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Mar Chem 51(4):325–346
27. Baker A, Spencer G (2004) Characterization of dissolved organic matter from source to sea using fluorescence and absorbance spectroscopy. Sci Total Environ 333(1–3):217–232
28. Osburn CL, Handels LT, Mikam MP, Paerl HW, Montgomery MT (2012) Fluorescence tracking of dissolved and particulate organic matter quality in a river-dominated estuary. Environ Sci Technol 46(16):8628–8636
29. Zhang YL, Yin Y, Feng LQ, Zhu GW, Shi ZQ, Liu XH, Zhang YZ (2011) Characterizing chromophoric dissolved organic matter in Lake Tiansimu and its catchment basin using excitation-emission matrix fluorescence and parallel factor analysis. Water Res 45(16):5110–5122
30. Yao X, Zhang YL, Zhu GW, Qin BQ, Feng LQ, Cai LL, Gao GA (2011) Resolving the variability of CDOM fluorescence in Lake Taihu and its tributaries. Chemosphere 82:145–155
31. Song FH, Wu FC, Feng WF, Liu SS, He J, Li TT, Zhang J, Wu AM, Amaranthsavaranada D, Xing BS, Bai YC (2019) Depth-dependent variations of dissolved organic matter composition and humification in a plateau lake using fluorescence spectroscopy. Chemosphere 225:507–516
32. Lu WW, Yao X, Shao KQ, Zhang BH, Gao G (2019) Unraveling the sources and fluorescence compositions of dissolved and particulate organic matter (DOM and POM) in Lake Taihu, China. Environ Sci Pollut Res. 26:4027–4040
33. Shon HK, Vigneswaran S, Snyder SA (2006) Effluent organic matter (EOM) in wastewater: constituents, effects, and treatment. Crit Rev Env Sci Tec. 36:327–374
34. Baker A, Lamont-black J (2001) Fluorescence of dissolved organic matter as a natural tracer of ground water. Groundwater 39(5):745–750
35. Chen H, Liao ZL, Gu YX, Xie XQ, Li HZ, Zhang J (2017) Anthropogenic influences of paved runoff and sanitary sewage on the dissolved organic matter quality of wet weather overflows: an excitation–emission matrix parallel factor analysis assessment. Environ Sci Technol 51:1157–1167
36. Murphy KR, Butler KD, Spencer RG, Stedmon CA, Boehme JR, Aiken GR (2010) Measurement of dissolved organic matter fluorescence in aquatic environments: an interlaboratory comparison. Environ Sci Technol 44(24):9405–9412
37. Li WT, Xu ZX, Li AM, Wu W, Zhou Q, Wang JN (2013) HPLC/HPLC–FLD with multi-excitation/ emission scan for EEM interpretation and dissolved organic matter analysis. Water Res 47(3):1246–1256
38. Hudson N, Baker A, Reynolds D (2007) Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review. River Res Appl. 23:631–649
39. Ye Z (2012) Study on the correction of dry-weather pollutants inappropriate entry into storm drainage based on pipe hydrodynamic model. Master Thesis, Tongji University, Shanghai (in Chinese)
40. Guisasiola A, Haas D, Keller J, Yuan Z (2008) Methane formation in sewer systems. Water Res 42(6–7):1421–1430
41. Jelic A, Rodriguez-Mozaz S, Barcelo D, Gutierrez O (2015) Impact of in-sewer transformation on 43 pharmaceuticals in a pressurized sewer under anaerobic conditions. Water Res 68:98–108
42. Sharifi S, Haghshenas MM, Dekissa T, Green P, Hare W, Massoudieh A (2014) Storm water pollution source identification in Washington, DC, using Bayesian chemical mass balance modeling. J Environ Eng 140:04015015
43. Vrugt JA (2016) Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation. Environ Modell Softw. 75:273–316
44. Chen MF, Wu J, Lv YL, Chen QJ (2008) Fluorescence properties of municipal wastewater. Acta Optica Sinica 28(4):712–718
45. Chen W, Westerhoff F, Leenheer JA, Booksh K (2003) Fluorescence excitation–emission matrix regional integration to quantify spectra for dissolved organic matter. Environ Sci Technol 37(24):5701–5710
46. Murphy KR, Hambly A, Singh S, Henderson RK, Baker A, Stuetz R, Khan SJ (2011) Organic matter fluorescence in municipal water recycling schemes: toward a unified PARAFAC model. Environ Sci Technol 45(7):2909–2916
47. Parlanti E, Worz K, Geoffroy L, Lamotte M (2000) Dissolved organic matter (DOM and POM) in Lake Taihu, China. Environ Sci Pollut Res 21(6):4727–4740
48. Ye Z (2012) Study on the correction of dry-weather pollutants inap-propriate entry into storm drainage based on pipe hydrodynamic model. Master Thesis, Tongji University, Shanghai (in Chinese)
51. Li WT, Xu ZX, Wu Q, Li Y, Shuang CD, Li AM (2015) Characterization of fluorescent-dissolved organic matter and identification of specific fluorophores in textile effluents. Environ Sci Pollut Res 22(6):4183–4189
52. Liu B, Wu J, Cheng C, Tang JK, Khan MFS, Shen J (2019) Identification of textile wastewater in water bodies by fluorescence excitation emission matrix-parallel factor analysis and high-performance size exclusion chromatography. Chemosphere 216:617–623
53. Wu Q, Li WT, Yu WH, Li Y, Li AM (2016) Removal of fluorescent dissolved organic matter in biologically treated textile wastewater by ozonation-biological aerated filter. J Taiwan Inst Chem E. 59:359–364
54. Reynolds DM (2002) The differentiation of biodegradable and non-biodegradable dissolved organic matter in wastewaters using fluorescence spectroscopy. J Chem Technol Biot. 77(8):965–972

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.