

Article

Fine-Scale Dissolved Organic Matter Fluorescence Fingerprints Reveal First-Flush Transition Dynamics in Urban Drainage Overflows

Hao Chen ^{1,2} , Yu Li ^{1,2,*}, Pengyi Cui ³ , Ting Zhang ^{1,2}, Jing Li ^{1,2}, Yaqin Tan ^{1,2} and Yali Guo ^{1,2}

¹ National Engineering Research Center of Eco-Environment in the Yangtze River Economic Belt, China Three Gorges Corporation, Wuhan 430014, China; chenh.87722@gmail.com (H.C.); zhang_ting1@ctg.com.cn (T.Z.); li_jing28@ctg.com.cn (J.L.); tan_yaqin@ctg.com.cn (Y.T.); guo_yali@ctg.com.cn (Y.G.)

² Shanghai Investigation, Design & Research Institute Co., Ltd., Shanghai 200335, China

³ School of Environment and Architecture, University of Shanghai for Science and Technology, Shanghai 200093, China; pycui@usst.edu.cn

* Correspondence: li_yu4@ctg.com.cn

Abstract

Urban drainage overflows can release a large fraction of event-scale pollutants during the early stage, yet current control remains largely driven by hydraulic signals rather than pollutant-release dynamics. This study created a dissolved organic matter (DOM)-based fluorescence fingerprint method to precisely identify the shift from pollutant flushing to dilution or ongoing input, helping determine the timing of first-flush transitions and potential interception. Fourteen wet-weather overflow events from seven drainage systems in Shanghai and Changzhou were investigated using excitation–emission matrix fluorescence spectroscopy, combined with non-negative matrix factorization, random forest feature screening, principal component analysis, mass–volume (M(V)) curve analysis, and Pettitt change-point detection. Five macro-scale fluorescence fingerprints were resolved, representing protein-like, fulvic-like, and humic-like components. Protein-like fingerprints dominated rapid event-scale variations, while fulvic-like and humic-like fingerprints reflected continuous surface-derived input and stable background contribution, respectively. Peak-shift trajectories revealed three fluorescence-evolution modes: directional red-shift migration, peak-position stability, and weak, non-directional variability, reflecting different source-release dynamics and DOM compositional adjustments during overflow. Random forest screening identified 20 high-importance fine-scale fluorescence fingerprints, with 90% concentrated in protein-like regions linked to sewage-derived and labile DOM. Compared with macro-scale fingerprints and conventional water quality indicators, fine-scale fluorescence fingerprints showed clearer stage separation, stronger consistency with M(V)-based cumulative response patterns, and more distinct first-flush interception timing. This timing marked the transition from early concentrated pollutant release to dilution or sustained input, whereas macro-scale fingerprints indicated broader transition intervals and conventional indicators showed delayed responses. These findings highlight the potential of fine-scale fluorescence fingerprints to support future fluorescence-assisted overflow control by improving transition identification and targeted interception decisions.



Received: 9 June 2026

Revised: 23 July 2026

Accepted: 24 July 2026

Published: 28 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

Keywords: urban drainage overflow; dissolved organic matter (DOM); EEM fluorescence spectroscopy; fine-scale fluorescence fingerprints; first-flush transitions; first-flush interception timing

1. Introduction

Wet-weather overflows from urban drainage systems are important pathways for the rapid discharge of rainfall-induced pollutants into receiving water bodies [1]. In many events, pollutants accumulated on urban surfaces, within sewer networks, and in drainage facilities are rapidly mobilized during the early overflow stage. This produces the first-flush phenomenon, in which pollutant mass is discharged disproportionately during the initial portion of the overflow volume [2]. Precisely recognizing the shift from concentrated pollutant flushing to either dilution or continued input is essential for enhancing overflow pollution management.

In current engineering practice, overflow interception and pumping-station operation are still mainly regulated by hydraulic parameters, such as water level, and by empirical operational periods. These indicators are effective for water-volume regulation and flood-control safety, but they do not directly reflect pollutant-release dynamics. As a result, the timing of interception may not correspond to the period of concentrated pollutant release, leading to a mismatch between hydraulic regulation and pollution control. Therefore, the key challenge is not only to determine whether a first-flush effect occurs after an event, but also to identify the process boundary at which early concentrated pollutant release shifts to dilution or continuous input. This transition boundary can also serve as a reference point for more targeted overflow interception decisions.

Existing first-flush assessment methods, including mass–volume ($M(V)$) curves, first-flush ratios, event mean concentration (EMC) analysis, and cumulative load comparisons, have been widely used to quantify the distribution of event-scale pollutant loads relative to runoff or overflow volume [3]. In $M(V)$ -based assessment, a curve above the 1:1 mass–volume line indicates that pollutant mass accumulates faster than runoff volume during the early event stage, while criteria such as the discharge of 80% of the pollutant mass within the first 30% of the event volume have been used to identify a pronounced first flush [2]. For example, Maniquiz-Redillas et al. [4] systematically characterized the first-flush effect of urban rainfall runoff based on $M(V)$ curves, clarifying the quantitative features of concentrated initial pollutant release. Luo et al. [5] reported temporal variations in rainfall pollution output from river-entry runoff outlets using EMC analysis and cumulative load methods.

However, these methods generally rely on post-event reconstruction of flow and pollutant concentration data, and therefore mainly provide retrospective descriptions of cumulative load behavior at the event scale. They are less suitable for capturing rapid transition signals during overflow processes or for directly supporting potential interception-timing identification. The formation and evolution of overflow pollution are highly dynamic and heterogeneous. Rainfall intensity and duration, antecedent dry period, catchment land use, surface pollutant accumulation, pipe-network deposits, domestic sewage mixing, and drainage-system operation can jointly affect pollutant release. Under these conditions, conventional water quality indicators usually reflect cumulative concentration responses rather than the earlier source-switching and mixing processes that drive first-flush transitions. Therefore, additional process-sensitive indicators are needed to capture short-term changes in pollutant release during overflow events.

Dissolved organic matter (DOM) provides a useful tool for characterizing overflow pollution dynamics because it is closely linked to organic pollutant sources, microbial activity, sediment disturbance, and surface runoff inputs. Changes in DOM concentration and composition can reflect source conversion and variations in pollutant-release intensity during rainfall events. Three-dimensional excitation–emission matrix (EEM) fluorescence spectroscopy has been widely used for DOM source apportionment and component identification due to its rapid measurement, high sensitivity, and rich spectral information [6].

Existing studies have identified typical fluorescent components, such as protein-like and humic-like substances, using EEM and have confirmed that DOM fluorescence characteristics can effectively reflect contributions from different sources, including domestic sewage input, surface scouring, and terrestrial background [7,8]. Some studies have also shown that protein-like components respond well to changes in the intensity of urban rainfall runoff and overflow pollution [9,10]. These findings suggest that fluorescence-based DOM characterization has strong potential for interpreting source evolution during urban overflow processes.

However, most existing fluorescence-based studies have focused on component intensity, component ratios, or overall differences before and after events. These approaches are useful for source interpretation, but they may not fully capture high-frequency spectral changes associated with rapid source release, sewage–stormwater mixing, and dilution during the early overflow stage. Macro-scale fluorescence fingerprints derived from decomposition methods such as parallel factor analysis (PARAFAC) or non-negative matrix factorization (NMF) integrate fluorescence signals over relatively broad spectral regions and thus mainly represent component-level DOM compositional dynamics [11,12]. However, such indicators can barely describe short-lived, localized spectral responses and peak-position changes, which may be smoothed during component integration. Recent studies have introduced source-identification indices derived from fine-scale fluorescence features, showing that source-specific DOM information can be retained in subtle spectral variations at individual excitation–emission coordinates beyond macro-scale fluorescence fingerprints and conventional indices [13]. This suggests that fine-scale fluorescence information may provide a more sensitive basis for resolving differences in DOM sources and changes in processes.

In highly dynamic overflow processes, first-flush transitions may initially appear as fine-scale fluorescence responses before becoming evident in broader macro-scale features. Peak-shift trajectories can indicate changes in DOM source-release behavior and compositional shifts during overflow. Meanwhile, pixel-level, fine-scale fluorescence intensities at specific excitation–emission coordinates may capture responses linked to sewage-derived and labile DOM. By retaining this information, fine-scale fluorescence fingerprints may be more sensitive to the transition from early, concentrated pollutant release to dilution or sustained input than macro-scale features are. However, it is unclear whether such fingerprints can effectively identify first-flush transitions in overflows, which EEM regions are most sensitive to short-term source-release dynamics, and whether they can provide clearer transition identification and potential interception timing than macro-scale fingerprints and conventional water quality indicators.

This study developed a fine-scale fluorescence-fingerprinting strategy to identify first-flush transitions in urban drainage overflows. Fourteen wet-weather overflow events from seven drainage systems in Shanghai and Changzhou, China, were analyzed using EEM fluorescence spectroscopy. NMF was used to extract macro-scale fluorescence fingerprints representing component-level DOM compositional dynamics and peak-shift trajectories, whereas Random Forest was used to identify fine-scale fluorescence fingerprints retaining localized spectral responses. The two fingerprint scales, together with conventional water quality indicators, were then compared using principal component analysis (PCA), M(V) curves, and Pettitt change-point analysis to evaluate stage separation, consistency with cumulative first-flush behavior and first-flush transition timing. We hypothesized that fine-scale fluorescence fingerprints would capture rapid source-release transitions more clearly and consistently than macro-scale fingerprints and conventional water quality indicators.

2. Materials and Methods

2.1. Study Area

This study was conducted in Shanghai and Changzhou, China (Figure 1). Shanghai is characterized by extensive impervious surfaces, drainage networks constructed over different periods, and complex pumping-station operations, making rainfall-induced overflow a recurrent wet-season problem. Changzhou represents another densely developed urban area where drainage overflows occur during rainfall events. The two cities were selected to capture overflow processes under common urban drainage and land-use conditions in the Yangtze River Delta.

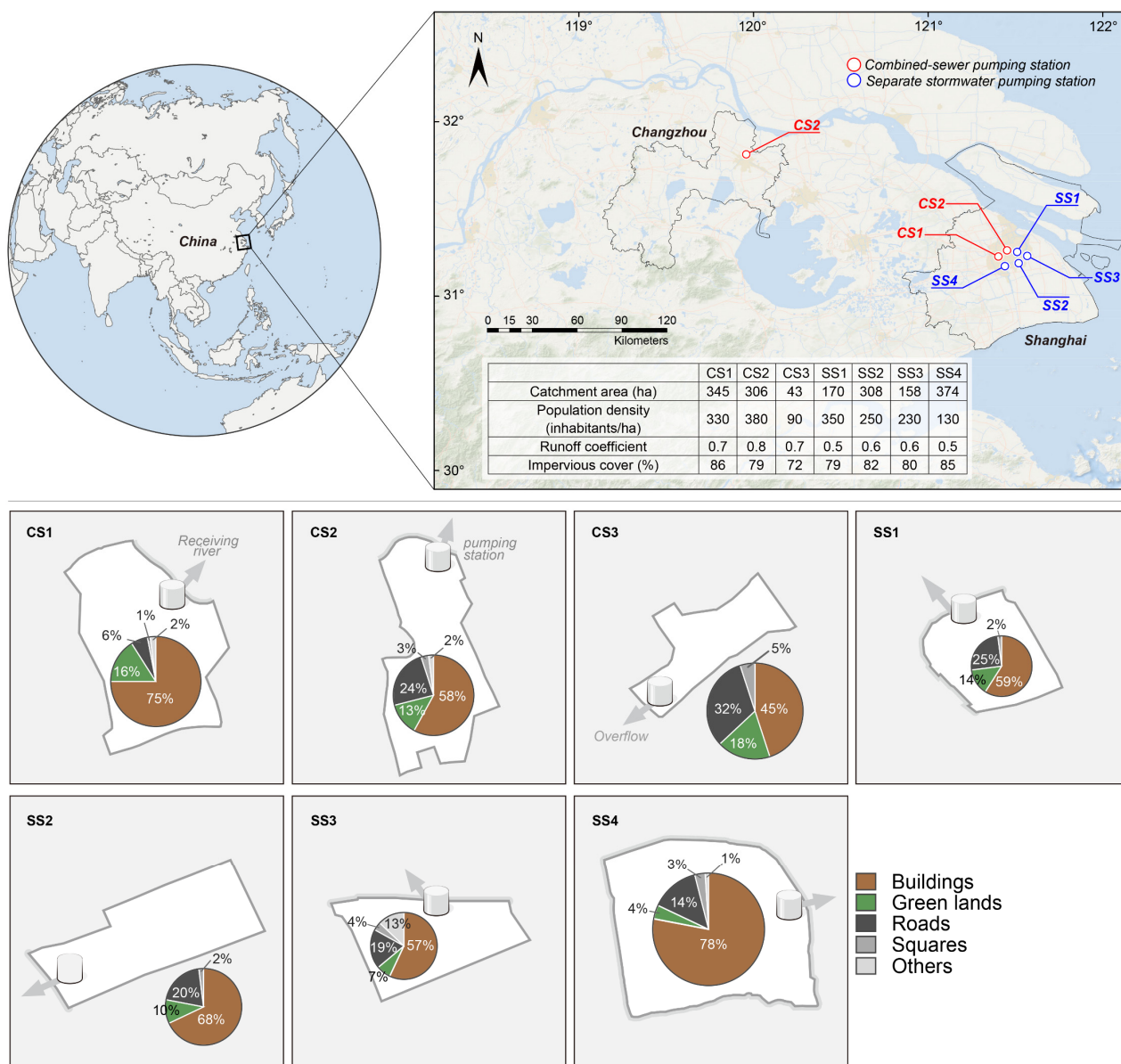


Figure 1. Locations and main characteristics of the 7 monitoring sites and their contributing catchments in Shanghai and Changzhou, China. CS1–CS3 represent combined-sewer pumping stations, whereas SS1–SS4 represent separate-stormwater pumping stations. The table panel summarizes the main site characteristics.

Seven monitoring sites were included, comprising six sites in Shanghai and one in Changzhou. CS1 and CS2 were combined-sewer pumping stations in Shanghai, whereas CS3 was a combined-sewer gravity overflow point in Changzhou. SS1–SS4 were separate

stormwater pumping stations in Shanghai. The contributing catchments consisted mainly of residential areas and urban roads, with smaller proportions of commercial land and urban green space, representing common land-use and land-surface conditions in densely developed areas of the Yangtze River Delta. Their relatively high imperviousness favors the antecedent accumulation of road-deposited sediments, atmospheric deposition, and domestic-source pollutants, which can be rapidly mobilized by surface runoff and drainage flows during the initial rainfall stage. All sampling locations, except CS3 (sampled from a gravity overflow well), were at the pump sumps. The locations of the two study cities, the seven monitoring sites, their contributing catchments, drainage system types, and the main site characteristics are shown in Figure 1. Pumping-station stormwater drainage capacities and rainfall intensities for the 14 overflow events are shown in Table 1.

Table 1. Pumping-station stormwater drainage capacities and rainfall intensities for the 14 overflow events.

Sampling Sites	Rainwater Drainage Capacity (m ³ /s)	Overflow Event	Rainfall (mm/h)
CS1	19.98	CS1-1	7.7
		CS1-2	3.4
CS2	22.495	CS2-1	5.0
		CS2-2	11.7
CS3	4.7	CS3-1	40.0
		CS3-2	1.4
SS1	16.8	SS1-1	4.8
		SS1-2	6.2
SS2	16.08	SS2-1	3.1
		SS2-2	6.2
SS3	14.2	SS3-1	3.7
		SS3-2	6.2
SS4	13.8	SS4-1	3.9
		SS4-2	6.4

2.2. Sample Collection

Field monitoring and sample collection were conducted during the rainy season from May to July 2015. Wet-weather overflow did not occur during every rainfall event because its occurrence depended on rainfall depth, intensity, duration, and temporal pattern, as well as the available conveyance and storage capacity of the drainage network, pumping capacity, and the wet-weather flow that could be intercepted and conveyed to the wastewater treatment system. Among the overflow events occurring during the monitoring period, two qualifying events were selected at each site using consistent criteria: an antecedent dry period of at least 7 days, complete rainfall and pumping-operation records, a clearly defined overflow process, and a continuous valid sampling sequence covering the monitored event. Events without overflow or with incomplete operational or sampling records were excluded. The selected events were denoted by “−1” and “−2”, resulting in 14 independent wet-weather overflow events spanning different drainage-system types, operating conditions, rainfall intensities, and overflow responses.

Overflow samples were collected using an ISCO 6712 automatic sampler (Teledyne ISCO, Lincoln, NE, USA), with 1 L collected as an individual sample at each sampling interval. The sampler was manually initiated in synchrony with the control-room operation that started the stormwater pumps, after which sampling proceeded automatically until

pumping ceased. At the gravity-overflow well, the sampler was initiated when overflow was first observed and continued operating until the overflow ended.

Before each monitoring event, the likely overflow response was initially assessed based on the forecast rainfall depth, intensity, and duration, together with the antecedent dry period, historical water-level variations, pumping-station operation records, and previous monitoring experience. During the event, real-time information on rainfall over the contributing catchment, pumping-station water level, pump start–stop status, and the number of operating stormwater pumps was obtained from the pump and gate control system. The discharge was estimated from the number of operating pumps and their rated capacities, and these data were used to dynamically assess the overflow stage and adjust the sampling intervals. To capture rapid changes in pollutant concentrations and fluorescence characteristics during the first-flush period, shorter sampling intervals were used during the initial stage, when the water level was rising rapidly, pumps were being successively activated, and the discharge was changing markedly [14]. The intervals were appropriately extended during the middle and late stages as rainfall, water level, and pumping conditions gradually became more stable, thereby balancing the capture of first-flush dynamics, coverage of the complete event, and field sampling feasibility. Specifically, 7 events, including CS1-1, SS1-2, SS2-1, SS2-2, SS3-1, SS4-1, and SS4-2, were sampled at 2 min intervals during the initial overflow stage, with 12–20 samples collected per event. For 4 events, including CS2-1, CS2-2, CS3-1, SS1-1, and SS3-2, the initial sampling interval was 5–15 min, with 8–18 samples collected per event. For CS1-2 and CS3-2, the sampling intervals were 10–20 min and 15–60 min, respectively, yielding 9 and 5 samples. During the middle and late stages, the sampling intervals were adjusted based on observed water-level and overflow dynamics and were generally extended to 10–40 min, with a maximum interval of 60 min for individual events. This sampling strategy yielded 211 overflow samples in total.

Pumping-station operation and rainfall-process data were recorded synchronously during sampling, including pump start and stop times, flow rate, water level, rainfall intensity, cumulative rainfall, and antecedent dry period. Supervisory Control and Data Acquisition (SCADA) data for pumping-station operation in Shanghai were provided by Shanghai Drainage Company, whereas rainfall data in Changzhou were obtained using an ISCO 674 rain gauge (Teledyne ISCO, Lincoln, NE, USA) deployed on site. After collection, all samples were immediately transported to the laboratory, stored at 4 °C, and pretreated within 24 h. Conventional water quality analyses and three-dimensional EEM fluorescence measurements were completed within 3 days.

2.3. Conventional Water Quality Analysis and EEM Fluorescence Spectroscopy

Dissolved organic carbon (DOC), ammonia nitrogen ($\text{NH}_3\text{-N}$), and total suspended solids (TSS) were selected as conventional water quality indicators. DOC was measured using a VCPN total organic carbon analyzer (Shimadzu Corporation, Kyoto, Japan), with a detection range of 5–50 mg C/L. $\text{NH}_3\text{-N}$ was determined by Nessler's reagent spectrophotometry according to HJ 535-2009 [15], and TSS was measured using the gravimetric method according to GB 11901-89 [16].

EEM fluorescence spectra were measured using a FluoroMax-4 fluorescence spectrophotometer (HORIBA Scientific, Edison, NJ, USA). Before analysis, samples were filtered through 0.45 μm polyvinylidene fluoride (PVDF) membranes (Merck Millipore, Darmstadt, Germany) to remove particulate interference. Sample pH was adjusted to 2.0 ± 0.2 using 0.1 mol/L HCl or NaOH, and KCl was added to maintain an ionic strength of 10 mmol/L (Sinopharm Chemical Reagent Co., Ltd., Shanghai, China) [17], ensuring consistent measurement conditions among samples. To reduce inner-filter effects, DOC concentrations

were controlled below 5 mg C/L, and UV_{254} absorbance was kept below 0.05 cm^{-1} [18]. The excitation wavelength range was 250–450 nm, and the emission wavelength range was 300–550 nm, both with a 5 nm scanning interval. Ultrapure water was used as the blank throughout the measurements. Each sample was scanned in triplicate, and the averaged EEM spectrum was used for subsequent fluorescence fingerprint extraction and analysis.

2.4. Data Analysis

The overall analytical workflow, from wet-weather overflow event monitoring and sampling, EEM determination and preprocessing to fluorescence fingerprint extraction (NMF decomposition and Random Forest feature screening), comparative transition analysis (PCA, M(V)-based evaluation, and Pettitt change-point analysis), and first-flush transition identification, is summarized in Figure 2.

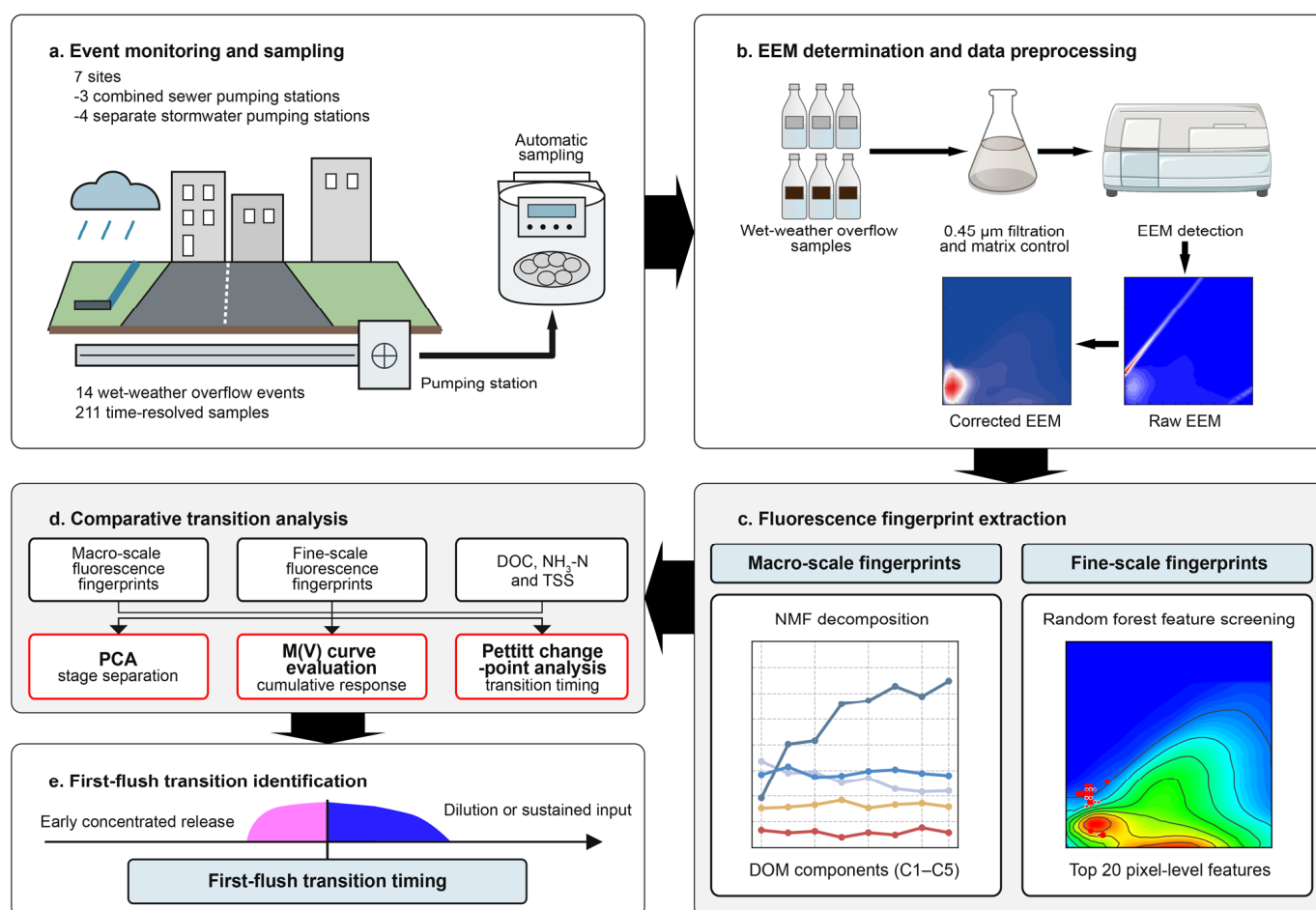


Figure 2. Overall workflow for comparing fine-scale fluorescence fingerprints, macro-scale fluorescence fingerprints, and conventional water quality indicators, including wet-weather overflow event monitoring and sampling, EEM determination and pre-processing to fluorescence fingerprint extraction (NMF decomposition and Random Forest feature screening), comparative transition analysis (PCA, M(V)-based evaluation, and Pettitt change-point analysis), and first-flush transition identification.

Raw EEM fluorescence data were preprocessed to remove spectral interferences and improve comparability among samples. First- and second-order Rayleigh scattering was removed using Delaunay triangulation interpolation [19,20], and Raman scattering was corrected using the ultrapure-water blank. Corrected fluorescence intensities were normalized to the Raman peak area to reduce instrumental and concentration-related variability.

NMF was applied to extract macro-scale fluorescence fingerprints, which represent component-level DOM signals integrated over broad spectral regions and capture event-scale fluorescence patterns and their temporal dynamics through a flexible non-negative representation. The component number was determined by comparing solutions with different preset ranks based on reconstruction residuals, spectral separation, and correspondence with established DOM fluorescence regions [21]. Fewer than five components merged distinct protein-like and humic/fulvic-like signals, whereas additional components produced little residual improvement and mainly subdivided existing spectral regions. The five-component solution was therefore retained as the most parsimonious and interpretable representation [22,23]. Fingerprint assignments and potential sources are summarized in Table 2.

Table 2. Five macro-scale fluorescence fingerprints and their potential source assignments [23,24].

Fluorescent Component	$\lambda_{\text{ex/em}}$ (nm)	Source
Tyrosine-like component (C1)	<250/~310	Domestic sewage, microbial metabolites
Tryptophan-like component (C2)	~275/~340/360	Domestic sewage, biodegradable organics
Anthropogenic humic/fulvic-like component (C3)	~250, (300)/~380–420	Surface runoff, road leaching
Terrestrial humic-like component (C4)	~325–355/~440	Soil leaching, plant residues
Refractory humic-like component (C5)	<250, (380)/~502	Soil organic matter, background input

Fine-scale fluorescence fingerprints were defined as pixel-level fluorescence intensities at specific excitation–emission coordinates selected using Random Forest. Unlike macro-scale fingerprints, they retain localized spectral responses associated with event progression. The sampling sequence within each overflow event was used as the response variable, and fluorescence intensities across the valid EEM region were used as predictors. Random Forest was selected because it can handle high-dimensional EEM data, capture nonlinear relationships and interactions among spectral variables, and rank individual excitation–emission coordinates according to their contributions to decision-tree splitting [24]. This allows informative regions on peak ridges, slopes, and valleys to be identified while preserving the original wavelength coordinates. In contrast, correlation-based screening evaluates coordinates independently, mutual information generally does not account for joint predictor interactions, and linear projection methods generate latent variables rather than retaining specific wavelength coordinates [25,26]. Random Forest was therefore considered more suitable for identifying localized and interpretable spectral features associated with event progression. The top 20 ranked coordinates were retained as the core fine-scale fluorescence fingerprints for first-flush transition analysis [13].

PCA, Pettitt change-point analysis, and M(V) curves were used to compare fine-scale and macro-scale fluorescence fingerprints with conventional water quality indicators. PCA was used to compare their ability to distinguish overflow stages. M(V) curves [27] were constructed to evaluate their consistency with cumulative first-flush behavior. Cumulative pollutant-load and fluorescence-response ratios were calculated relative to cumulative overflow-volume ratios for conventional indicators and fluorescence fingerprints, respectively. A first-flush response was identified when either curve lay above the 1:1 reference line during the early overflow stage. M(V) analysis was limited to events with complete cumulative overflow-volume records; CS1-2, CS3-1, and CS3-2 were excluded because of insufficient volume-process data.

The Pettitt nonparametric test [28] was applied to the time series of the three indicator groups to identify statistically significant change points separating the early concentrated-

release stage from subsequent dilution or sustained input. For each event, the dominant change point among indicators with $p < 0.05$ was defined as the first-flush transition timing. The resulting transition timings were then compared across the three indicator groups with reference to M(V) curve deviations from the 1:1 line and the corresponding stage-change characteristics.

Data analysis and visualization were performed using Python 3.7.1 (Python Software Foundation, Wilmington, DE, USA), Matplotlib 3.5.3 (Matplotlib Development Team), and OriginPro 2025 (OriginLab Corporation, Northampton, MA, USA).

3. Results

3.1. Dynamic Characteristics of Macro-Scale Fluorescence Fingerprints

Using NMF, the fluorescence dataset from 14 wet-weather overflow events was decomposed into five component-level macro-scale fluorescence fingerprints representing broader patterns in DOM composition (Table 2). C1 and C2 were assigned to tyrosine-like and tryptophan-like protein-like fingerprints, respectively. C3 was assigned to an anthropogenic humic/fulvic-like fingerprint. C4 and C5 were assigned to terrestrial humic-like fingerprints.

In Figure 3, the temporal patterns of these macro-scale fingerprints varied among monitoring sites, reflecting differences in source contributions and dominant release processes. In CS1 and SS2 events, protein-like fingerprints C1 and C2 increased rapidly during the early overflow stage and then decreased continuously as the event progressed, with overall reductions exceeding 50%. In contrast, C3 showed smoother temporal variation at these sites and lacked the sharp early peaks observed for C1 and C2. SS3 showed a different pattern, with C1 and C2 remaining relatively stable and without clear initial peaks. Across the monitored events, C4 and C5 generally showed low intensities and limited fluctuations.

3.2. Peak-Shift Patterns of Macro-Scale Fluorescence Fingerprints

Peak-shift trajectories on the Ex–Em plane differed substantially among wet-weather overflow events (Figure 4). Three representative patterns were identified: directional red-shift migration, peak-position stability, and weak non-directional variability, corresponding to columns A, B, and C, respectively.

Directional red-shift migration was observed in events such as CS1-1, CS1-2, SS2-1, and SS2-2. The C1 peak shifted from approximately Ex/Em $\approx 280/310$ nm at the initial stage to 290–295/315–320 nm, with an overall red-shift amplitude of about 10–15 nm, while C2 showed a weaker emission red shift of about 5–10 nm. These shifts mainly occurred within the first 20–40 min after overflow began and generally coincided with the rapid decrease in C1 from a high initial intensity. By contrast, peak-position stability was observed in CS2-1, CS2-2, CS3-1, and CS3-2, where peak-position variations were generally less than 5 nm. C1 remained within Ex/Em ≈ 275 –285/305–315 nm and C2 near Em ≈ 335 –345 nm, with no clear unidirectional shift, suggesting limited changes in the dominant fluorescent groups and DOM composition. Weak non-directional variability was observed in SS3-1, SS3-2, and SS4-2. In these events, peak variations were mostly within ± 5 nm but lacked a consistent direction, suggesting minor local fluctuations under relatively stable background control rather than a systematic DOM evolution pathway.

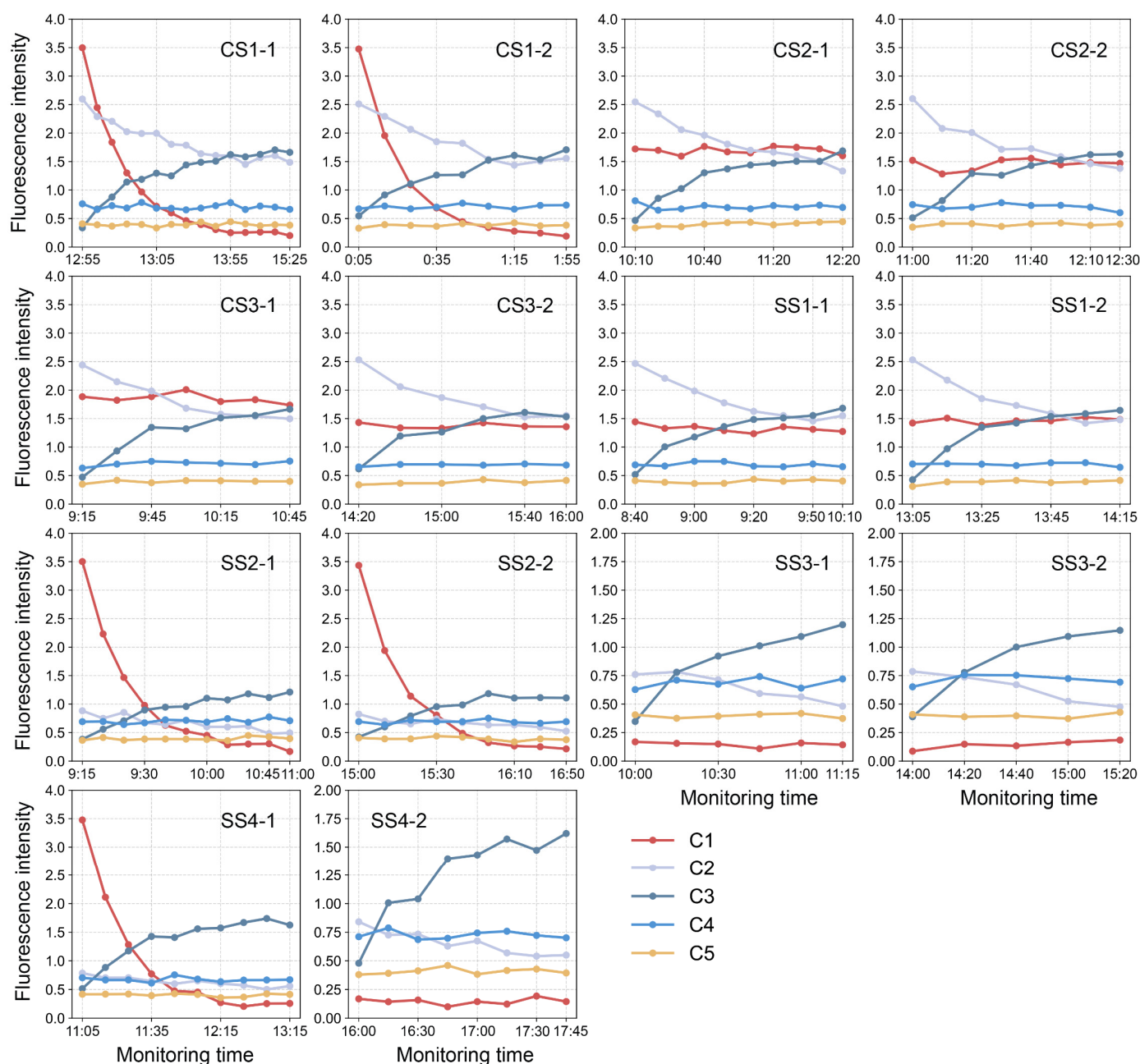


Figure 3. Temporal variations in the intensities of five NMF-derived macro-scale fluorescence fingerprints during 14 wet-weather overflow events. C1: tyrosine-like component; C2: tryptophan-like component; C3: anthropogenic humic/fulvic-like component; C4: terrestrial humic-like component; C5: refractory humic-like component. The corresponding excitation–emission characteristics and potential source assignments are provided in Table 2.

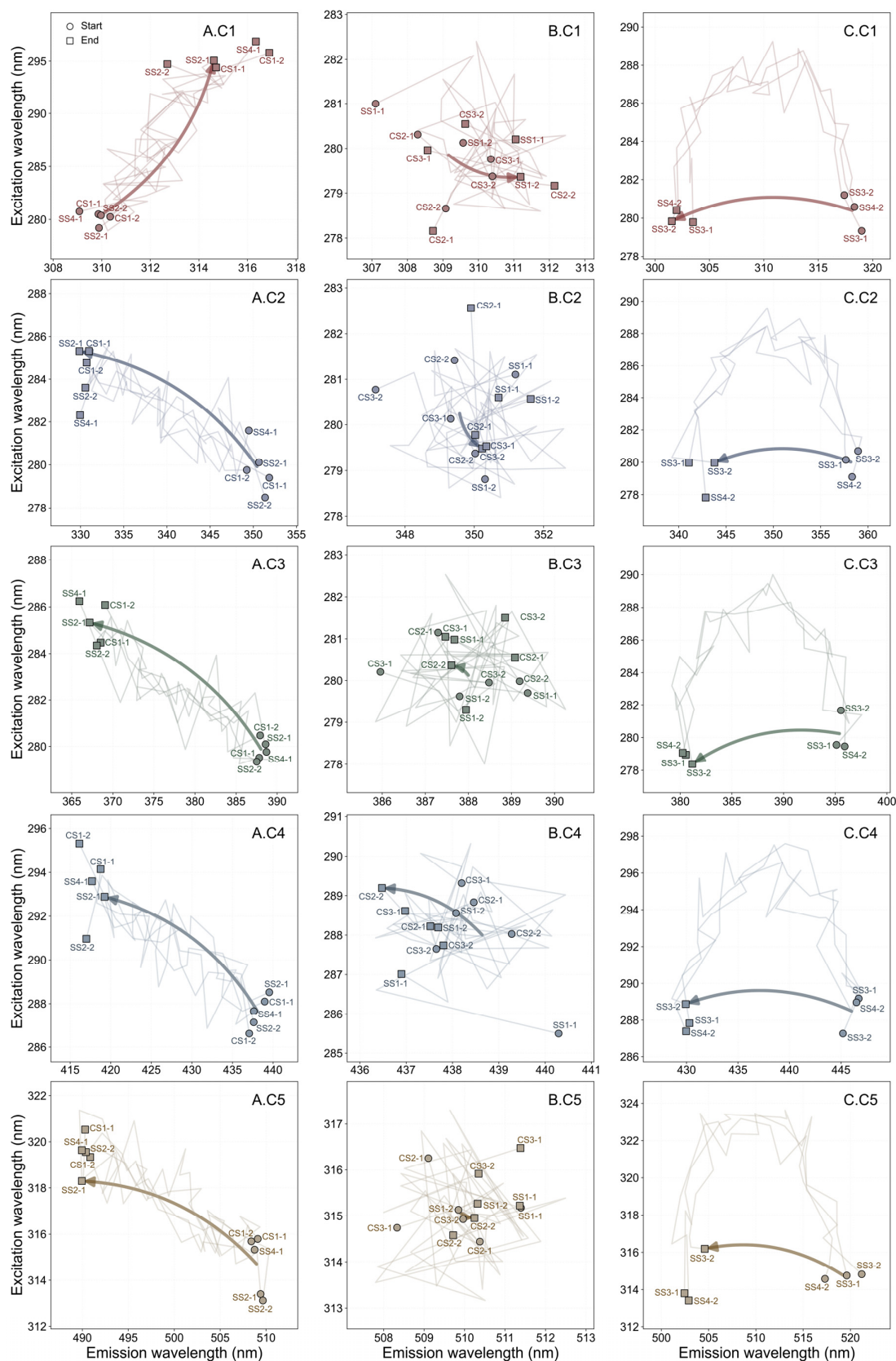


Figure 4. Peak-shift trajectories of macro-scale fluorescence fingerprints on the Ex-Em plane. Columns (A–C) indicate directional peak migration with red-shift characteristics, peak-position stability, and weak non-directional variability, respectively. C1: tyrosine-like component; C2: tryptophan-like component; C3: anthropogenic humic/fulvic-like component; C4: terrestrial humic-like component; C5: refractory humic-like component.

3.3. Spectral Distribution of Fine-Scale Fluorescence Fingerprints

The random forest model identified 20 high-importance pixel-level features as fine-scale fluorescence fingerprints (Figure 5). These fingerprints were not evenly distributed across the EEM space but were concentrated in several specific spectral regions. Eighteen fingerprints, accounting for 90% of the total, were located in the protein-like fluorescence region of Ex 270–285 nm and Em 300–355 nm. This cluster was spatially continuous and mainly centered around Em 310–340 nm, with a slight extension toward approximately Em 350 nm. The remaining two fingerprints were located in the humic/fulvic-like region of Ex 250–255 nm and Em 388–438 nm, near Em values of approximately 390 and 430 nm, respectively. Domestic sewage samples showed strong fluorescence responses in the protein-like region, whereas stormwater-runoff samples showed much weaker responses.

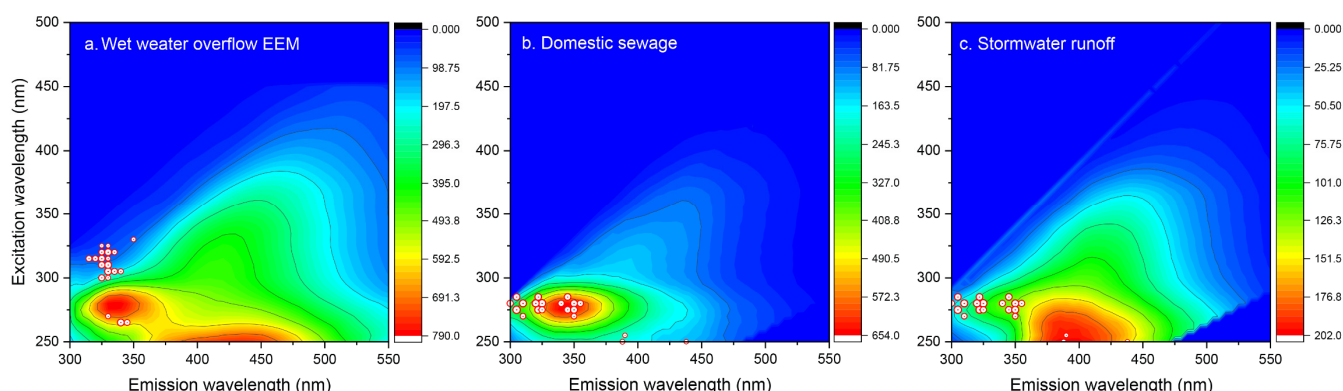


Figure 5. Spectral distribution of high-importance fine-scale fluorescence fingerprints in EEM space.

3.4. Stage Separation and Cumulative First-Flush Responses

To examine whether different indicators could capture the first-flush process, fine-scale fluorescence fingerprints, macro-scale fluorescence fingerprints, and conventional water quality indicators were compared using PCA and M(V) curve analysis (Figure 6). The fine-scale fluorescence fingerprints consisted of the 20 high-importance pixel-level features selected by the random forest model. The macro-scale fluorescence fingerprints included the five NMF-derived components C1–C5, while the conventional indicators included DOC, $\text{NH}_3\text{-N}$, and TSS.

The PCA results showed that the three indicator groups responded differently to overflow-stage evolution. For macro-scale fluorescence fingerprints, PC1 and PC2 explained 81.4% and 12.3% of the variance, respectively, with a cumulative contribution of 93.7%. The samples showed a directional distribution along the overflow process, but substantial overlap remained among different stages, and the stage boundary was relatively indistinct. Fine-scale fluorescence fingerprints showed a more distinct stage structure. PC1 and PC2 explained 92.5% and 4.7% of the variance, respectively, with a cumulative contribution of 97.2%. Samples were more clearly separated along PC1, and an evident transition interface appeared between the early concentrated pollutant-release stage and the subsequent stage dominated by dilution or sustained input. In contrast, PCA based on conventional water-quality indicators showed a smoother gradient distribution, despite a cumulative contribution of 90.7%, with greater overlap among different stages.

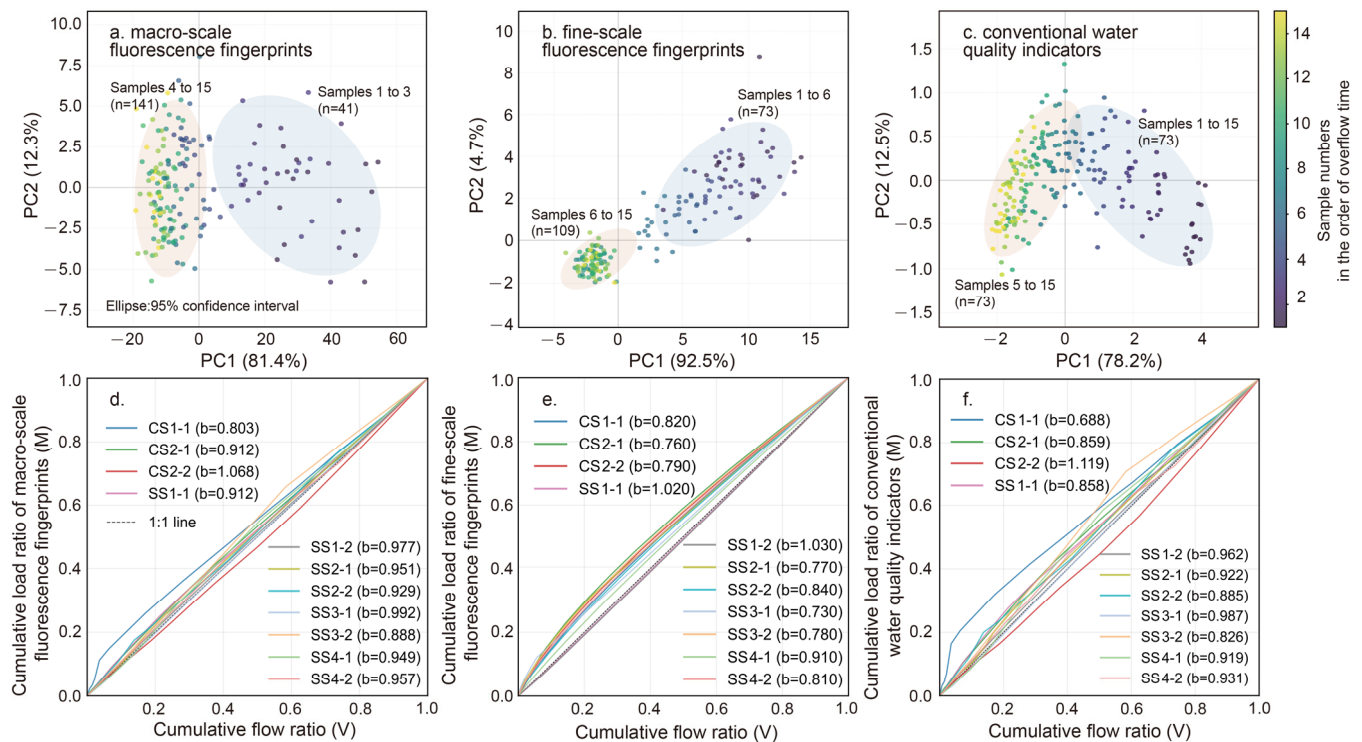


Figure 6. (a–f) PCA ordination and M(V) curves for different indicator groups.

The M(V) curves also showed distinct patterns among the three indicator groups. For conventional water-quality indicators, the curves represented cumulative pollutant-load ratios relative to cumulative overflow-volume ratios. For macro- and fine-scale fluorescence fingerprints, the curves represented cumulative fluorescence-response ratios relative to cumulative overflow-volume ratios and were therefore interpreted as DOM-fluorescence-response-based cumulative curves rather than strict pollutant mass-load curves. Macro-scale fluorescence fingerprints generally deviated from the 1:1 reference line by approximately 0.15–0.35, but the curves changed relatively smoothly and showed weak slope inflections. Fine-scale fluorescence fingerprints showed stronger early-stage deviations from the 1:1 line, reaching approximately 0.2–0.5, together with more pronounced slope changes. Conventional indicators showed smaller deviations, typically within 0.1–0.3, and gentler slope changes. Overall, the fine-scale fluorescence-response curves were more consistent with the pattern of concentrated pollutant release during the early overflow stage.

3.5. First-Flush Transition Timing

Figure 7 summarizes the first-flush transition timing identified by fine-scale fluorescence fingerprints, macro-scale fluorescence fingerprints, and conventional water-quality indicators. Fine-scale fluorescence fingerprints generally identified a more distinct transition timing. The identified timing was consistent with the transition interface observed in PCA space and the inflection point of the fine-scale fluorescence-response M(V) curves. Macro-scale fluorescence fingerprints tended to identify broader transition intervals rather than a single clear timing. Conventional water-quality indicators generally showed later transition signals, lagging those identified by fine-scale fluorescence fingerprints by approximately 2–3 sampling cycles in several events.

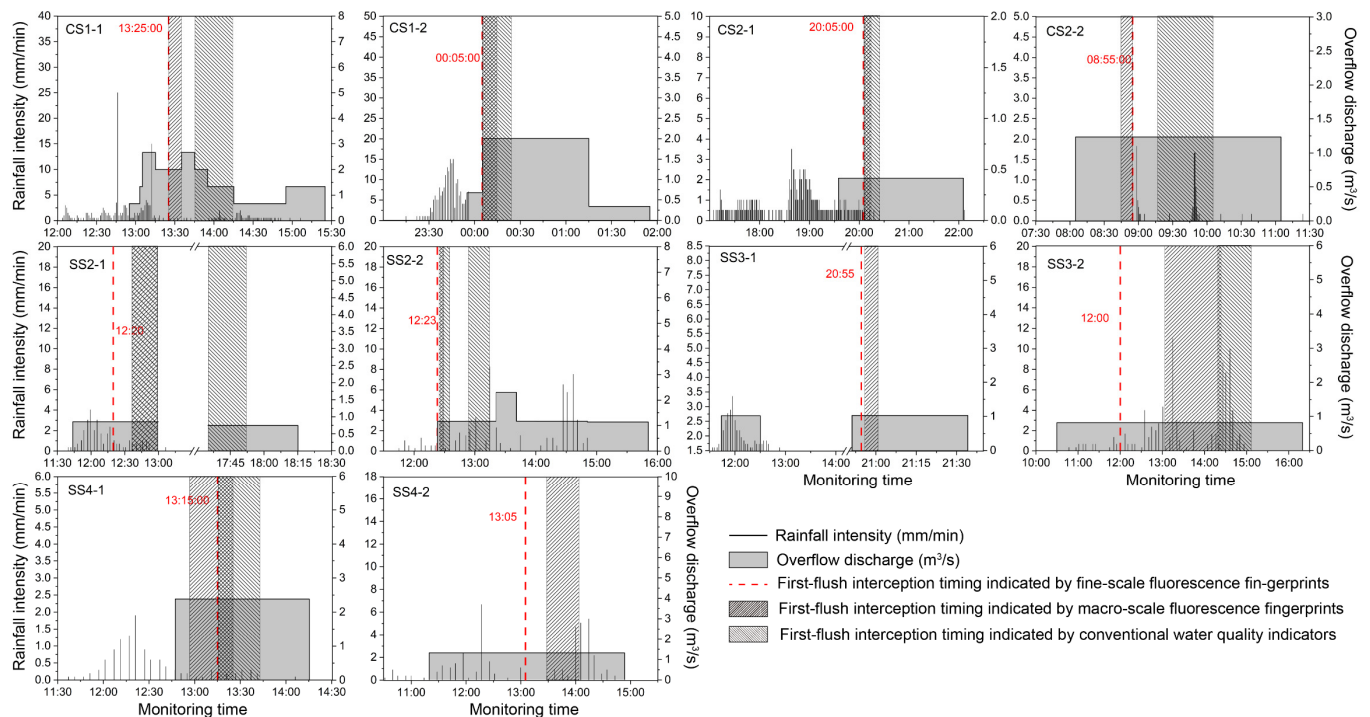


Figure 7. Temporal comparison of first-flush transition timing identified by fine-scale fluorescence fingerprints, macro-scale fluorescence fingerprints, and conventional water quality indicators during wet-weather overflow events.

4. Discussion

4.1. DOM Source Evolution and Fine-Scale Spectral Responses

The five macro-scale fluorescence fingerprints represented three principal DOM source categories during overflow: sewage-derived labile organic matter associated with C1 and C2 [12], surface-derived anthropogenic humic/fulvic-like DOM represented by C3 [29], and more aromatic and structurally stable terrestrial humic-like DOM represented by C4 and C5 [30]. The pronounced early increase and subsequent reduction in C1 and C2 in CS1 and SS2, with overall declines exceeding 50%, indicated rapid flushing of protein-like, sewage-derived DOM accumulated within the drainage network prior to rainfall [31]. In comparison, the smoother C3 trajectories suggested a more sustained contribution from surface runoff, while the absence of clear early C1 and C2 peaks at SS3 indicated weaker release of accumulated protein-like pollutants and a greater influence of continuous surface or background inputs. The persistently low and weakly varying C4 and C5 signals further showed that terrestrial humic-like DOM responded less strongly to rainfall disturbance and mainly constituted a stable background component [32]. These patterns reveal a source-dependent sequence in which antecedently accumulated protein-like organic matter is rapidly mobilized during the early stage [33], surface-runoff-derived DOM contributes continuously and increasingly shapes the later-stage composition [34], and terrestrial humic-like DOM remains comparatively stable. Differences among sites indicate that the relative importance of these pathways was controlled by pipe-network pollutant accumulation, surface-runoff contributions, terrestrial background inputs, and drainage-system conditions [35]. Because macro-scale fingerprints integrate fluorescence responses across broad spectral regions, however, localized changes associated with source switching, sewage–stormwater mixing, and early first-flush transitions may be smoothed, requiring further examination of fluorescence peak-shift trajectories and fine-scale spectral features.

DOM compositional changes at fine spectral scales are thought to first appear as shifts in fluorescence peak positions or other localized spectral responses, rather than as

synchronous changes in overall fluorescence intensity [36,37]. The three observed trajectory patterns represented distinct pathways and degrees of DOM source evolution and compositional change. Directional red-shift migration occurred mainly during the early overflow stage and coincided with the decline in protein-like fluorescence intensity, suggesting that rapid flushing, sewage–stormwater mixing, and dilution were accompanied by a shift from labile protein-like DOM toward more conjugated or humified fluorescence signals. Such red shifts are commonly associated with increased aromatic conjugation, apparent molecular size, or degree of humification [38]. In contrast, peak-position stability indicated limited changes in the dominant fluorescent groups and overall DOM composition, whereas weak non-directional variability was more consistent with localized fluctuations under relatively stable background conditions, potentially driven by variations in mixing conditions, temperature, pH, or ionic strength [39]. Directional red-shift migration therefore indicated a stronger early-stage compositional transition, peak-position stability reflected limited DOM restructuring, and weak non-directional variability represented weak disturbance without a consistent direction of evolution. The occurrence of these different patterns suggests that DOM evolution during overflow was jointly shaped by drainage-system conditions, antecedent pollutant accumulation, domestic sewage contributions, and rainfall characteristics. These results show that changes in macro-scale fluorescence intensity alone cannot fully capture the evolution of DOM composition during overflow.

The concentration of 90% of the selected fine-scale fingerprints in the protein-like region further identified where the most process-sensitive information was located. This region includes tyrosine- and tryptophan-like substances, free aromatic amino acids, and microbial protein-like products that are sensitive to inputs from domestic sewage, microbial activity, and hydrological dilution [40]. The contrasting responses of domestic sewage and stormwater-runoff samples support the ability of these wavelengths to track changes in the relative contribution of sewage-derived DOM during overflow. By comparison, the limited number of fingerprints in the humic/fulvic-like region was consistent with the greater persistence and weaker event-scale response of terrestrial humic-like and stable fulvic-like background components [41]. The selected fingerprints therefore narrowed the information relevant to first-flush transitions to a small set of protein-like wavelength regions, whereas stable humic-like signals contributed less to the identification of transitions. Together with the peak-shift patterns, this spectral localization shows that localized fluorescence responses can reveal changes in sewage-derived and labile DOM that are less apparent in integrated macro-scale fingerprints.

4.2. Identification and Interpretation of First-Flush Transitions

The sensitivity of the fine-scale fingerprints was expressed in both overflow-stage separation and cumulative first-flush behavior. PCA characterized the separation between the early concentrated-release stage and the subsequent dilution or sustained-input stage, whereas the M(V) curves evaluated whether the initial overflow volume carried a disproportionate share of pollutant load or fluorescence response. The relatively indistinct PCA boundary and smooth M(V) trajectories of the macro-scale fingerprints can be attributed to the same component-integration effect: aggregation across broader spectral regions preserves the overall direction of DOM evolution but reduces the contrast between successive stages [42,43]. Conventional water-quality indicators showed a different limitation. Because DOC, NH₃-N, and TSS are jointly influenced by pollutant mobilization, dilution, transport, and particulate dynamics, their responses integrate several concurrent processes and therefore change more gradually during the transition [44]. Fine-scale fluorescence fingerprints retained sufficient process contrast to distinguish the early concentrated-release stage in PCA while also reproducing its disproportionate contribution in the M(V) curves [33]. The

agreement between these two analytical perspectives indicates that the observed pattern was not specific to either multivariate stage separation or cumulative response analysis, but reflected the same transition from concentrated pollutant release to subsequent dilution or sustained input.

Pettitt change-point analysis further translated the stage differences identified by PCA and the cumulative first-flush patterns reflected by the $M(V)$ curves into specific transition timings. Fine-scale fluorescence fingerprints resolved the transition more clearly, and the identified timing corresponded with both the PCA transition interface and the inflection of the fine-scale fluorescence-response $M(V)$ curves. The convergence of spectral change, stage separation, and cumulative response around a common boundary strengthens the interpretation that this timing represents a genuine shift in the overflow process rather than an isolated statistical fluctuation. Macro-scale fingerprints produced broader transition intervals because component-level integration smoothed changes associated with rapid source switching and sewage–stormwater mixing. Conventional water-quality indicators generally responded later, suggesting that DOC, $\text{NH}_3\text{-N}$, and TSS mainly reflected changes after localized fluorescence responses had already emerged. These temporal differences reveal a sequential response mechanism: localized spectral changes first appeared in fine-scale fluorescence fingerprints, broader DOM compositional changes were subsequently expressed in macro-scale fingerprints, and the accumulated effects finally became evident in conventional water-quality indicators [7]. The timing identified by fine-scale fluorescence fingerprints therefore marked the process boundary at which early concentrated pollutant release shifted toward dilution or sustained input.

This process-based interpretation extends the conventional understanding of first flush. First flush is commonly assessed through event-scale cumulative load behavior, particularly whether a disproportionate pollutant load is discharged during the initial runoff or overflow volume. The present results also indicate a transition in the contribution of DOM sources and their composition. Early concentrated release was progressively replaced by dilution or sustained input, accompanied by changing sewage-, surface-, and background-derived DOM signals. Fluorescence analysis was therefore extended from source interpretation to process identification: macro-scale component intensities captured the overall evolution of DOM sources, while localized fine-scale responses identified the transition boundary more clearly.

4.3. Practical Implications and Future Validation

The identified transition boundary has direct implications for overflow management. The current operation of combined and separate sewer overflows remains largely governed by hydraulic parameters such as water level, flow rate, and empirical operating periods [45,46]. These parameters are essential for drainage safety and water-volume regulation, but they do not directly indicate when pollutant release is most concentrated. Existing first-flush assessments also rely mainly on post-event statistical reconstruction and are therefore poorly suited to identifying process boundaries during active overflow events [5]. Fine-scale fluorescence fingerprints could complement hydraulic regulation by providing high-frequency information on water quality and pollutant release dynamics. The identified transition timing may therefore provide a process-based reference for more targeted overflow interception and support a gradual shift from hydraulic-condition-based control toward fluorescence-assisted, water-quality-guided management [47,48].

Three-dimensional EEM fluorescence spectroscopy has already shown potential for online DOM identification due to its non-destructive nature, high sensitivity, and rapid response [47]. The present study further indicates that fluorescence information related to first-flush transitions is not uniformly distributed across the full EEM spectrum, but

is concentrated in several key wavelength regions. This provides a possible basis for future targeted monitoring schemes based on selected fine-scale fluorescence fingerprints. Commercially available in situ fluorescence sensors can continuously monitor selected DOM-related signals at one or a few fixed excitation–emission wavelength pairs. In practical wastewater monitoring, such sensors may be used alongside other commercial water-quality technologies. Online UV–Vis spectrometers are increasingly applied for continuous wastewater monitoring because absorbance-based measurements, such as SAC_{254} , can be used directly or, after site-specific calibration, to estimate parameters such as COD and DOC. Lower-cost surrogate sensors, including electrical conductivity, oxidation–reduction potential, and turbidity probes, can provide complementary information on dilution and source mixing, redox-condition changes, and suspended-pollutant mobilization during overflow events. Their relatively low cost and suitability for continuous deployment make them useful for distributed monitoring and event detection and, when combined with optical measurements, for distinguishing different processes underlying water-quality changes. Fluorescence sensing may provide more process-specific information on DOM composition and first-flush transitions within such a monitoring system. However, the established applications of in situ fluorescence sensors remain largely confined to environmental water-quality monitoring, including rivers, reservoirs, groundwater, and sewage-impacted receiving waters, whereas long-term deployment within sewer networks or CSO facilities is still limited [49–51]. Compared with conventional CSO monitoring based on rainfall, water level, and flow, fluorescence sensors can provide additional high-frequency information on water quality and pollutant-release dynamics. In future applications, fluorescence sensing may therefore be most useful as part of a broader multi-parameter monitoring framework that combines hydraulic measurements, online UV–Vis spectroscopy, and lower-cost surrogate sensors, rather than as a standalone monitoring approach. This added capability comes with additional equipment and maintenance costs, as well as the need for site-specific calibration, regular sensor cleaning, and correction for interference from turbidity, complex water matrices, and sensor fouling. At present, fluorescence sensing is therefore better suited as a water-quality complement to conventional hydraulic monitoring than as a direct replacement. The key wavelength regions identified in this study may support the future development of simplified monitoring systems. However, the present findings demonstrate the feasibility of the proposed framework for event-scale process identification under the monitored conditions, rather than its universal transferability across drainage systems or its readiness for real-time operational application.

The present results demonstrate the feasibility of the proposed framework for event-scale process identification under the monitored conditions, but they do not establish universal transferability across drainage systems or readiness for real-time operational use. Overflow DOM composition varies with catchment land use, drainage-system configuration, pipe-network conditions, rainfall characteristics, antecedent dry period, and the relative contributions of domestic sewage and surface runoff. The fingerprints and transition criteria identified here therefore require validation across different drainage systems and rainfall scenarios. Matrix-related interference also remains a technical challenge in drainage environments. Turbidity and temperature effects can be partly addressed through sensor-specific corrections based on concurrent measurements [49,52], while absorbance- or scattering-based methods can mitigate inner-filter effects [49,53]. When multiwavelength or EEM data are available, multivariate calibration and spectral decomposition may further help separate overlapping fluorescence signals from matrix-related variation [54]. However, these approaches are generally instrument- and site-specific and may become less reliable under highly turbid conditions or severe sensor fouling. Accordingly, real-time

field validation, cross-site testing, and long-term performance assessment are still required before the proposed approach can be used for operational overflow monitoring [55].

Future work should advance the proposed framework through staged field validation. The robustness of the selected wavelength regions and transition criteria should first be tested across different catchments, drainage systems, rainfall scenarios, and overflow conditions. Real-time field monitoring is then needed to determine whether these signals remain stable under continuous operation and whether the identified transition timing can be reproduced under changing hydraulic and water-quality conditions. Fluorescence signal-correction methods should also be improved for complex overflow matrices, particularly under high-turbidity and suspended-solid conditions. In parallel, simplified monitoring approaches based on selected fine-scale fluorescence regions should be evaluated against full EEM measurements to determine whether they retain sufficient sensitivity to identify first-flush transitions. With such validation, fine-scale fluorescence fingerprints may provide a practical water-quality signal for targeted overflow interception and pollution-process-oriented drainage management [56,57].

5. Conclusions

This study developed and evaluated a fine-scale fluorescence-fingerprinting strategy for identifying first-flush transitions in urban drainage overflows. The main conclusions corresponding to the study objectives are as follows:

1. To characterize component-level DOM dynamics, NMF resolved five macro-scale fluorescence fingerprints representing protein-like, fulvic-like, and humic-like signals. Protein-like fingerprints dominated rapid event-scale variations and reflected the early release of sewage-derived and labile DOM, whereas humic-like fingerprints remained relatively stable as background components.
2. To further interpret short-term DOM compositional changes, peak-shift analysis identified three fluorescence-evolution patterns: directional red-shift migration, peak-position stability, and weak non-directional variability. These patterns reflected different degrees of DOM source release and compositional adjustment during overflow events.
3. To retain localized spectral responses, Random Forest screening identified 20 high-importance fine-scale fluorescence fingerprints, 90% of which were concentrated in protein-like spectral regions associated with sewage-derived and labile DOM. These fingerprints provided higher-resolution information on rapid changes in pollutant release than the component-level fingerprints.
4. To evaluate their ability to identify first-flush transitions, fine-scale fluorescence fingerprints were compared with macro-scale fluorescence fingerprints and conventional water-quality indicators using PCA, Pettitt change-point analysis, and M(V) curves. Fine-scale fingerprints showed clearer stage separation, stronger consistency with cumulative first-flush patterns, and more distinct transition timing. These results support the study hypothesis that localized fluorescence responses can identify rapid source-release transitions more clearly and consistently than broader component-level signals or conventional concentration indicators.

Author Contributions: H.C.: Conceptualization, experimental design, experimentation, data analysis, writing—original draft, and writing—review and editing. Y.L.: Conceptualization, data processing and analysis, experimental index determination, validation, and writing—review and editing. P.C.: Conceptualization, project administration, data collation and statistical analysis, and manuscript revision. T.Z.: Investigation, experimental implementation and operation, data curation. J.L.: Data curation, manuscript polishing and revision. Y.T.: Methodology, data collation and statistical analysis.

Y.G.: Project administration, resources, methodology establishment, experimental supervision, and validation. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Scientific Research Project of China Three Gorges Corporation (grant number GCZX25-3-044) and the Natural Science Foundation of Shanghai (grant number 24ZR1450500).

Data Availability Statement: Data are contained within the present article. The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: During the preparation of this manuscript/study, the author(s) used ChatGPT-V5.4 for the purposes of language editing and Python 3.7.1 along with Origin 2025 for the purposes of data visualization and graphics. These tools were not used to generate data, perform analyses, or draw scientific conclusions. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: Authors Hao Chen, Yu Li, Ting Zhang, Jing Li, Yaqin Tan, and Yali Guo are employed by Shanghai Investigation, Design & Research Institute Co., Ltd., a second-tier subsidiary of China Three Gorges Corporation, and are also affiliated with the National Engineering Research Center of Eco-Environment in the Yangtze River Economic Belt, a research institution directly administered by China Three Gorges Corporation. The authors declare that this study received funding from the China Three Gorges Corporation. The funder was not involved in the study design, data collection, analysis, interpretation, manuscript writing, or the decision to submit the article for publication.

References

1. Botturi, A.; Ozbayram, E.G.; Tondera, K.; Gilbert, N.I.; Rouault, P.; Caradot, N.; Gutierrez, O.; Daneshgar, S.; Frison, N.; Akyol, Ç.; et al. Combined Sewer Overflows: A Critical Review on Best Practice and Innovative Solutions to Mitigate Impacts on Environment and Human Health. *Crit. Rev. Environ. Sci. Technol.* **2021**, *51*, 1585–1618. [[CrossRef](#)]
2. Bertrand-Krajewski, J.-L.; Chebbo, G.; Saget, A. Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon. *Water Res.* **1998**, *32*, 2341–2356. [[CrossRef](#)]
3. Wambecq, T.; Kroll, S.; Van Assel, J.; Delgado, R. CSO Generator-A Parsimonious Wastewater Quality Model for Combined Sewer Overflows. *Water* **2023**, *15*, 3424. [[CrossRef](#)]
4. Maniquiz-Redillas, M.; Robles, M.E.; Cruz, G.; Reyes, N.J.; Kim, L.-H. First flush stormwater runoff in urban catchments: A bibliometric and comprehensive review. *Hydrology* **2022**, *9*, 63. [[CrossRef](#)]
5. Luo, H.B.; Luo, L.; Huang, G.; Liu, P. Total pollution effect of urban surface runoff. *J. Environ. Sci.* **2009**, *21*, 1186–1193. [[CrossRef](#)] [[PubMed](#)]
6. Chu, J.; Luo, C.; Wang, H.; Liao, Z. Insights into the Associations between Op-tical and Molecular Signatures of Dissolved Organic Matter from Urban Stormwater Runoff. *ACS ES&T Water* **2022**, *2*, 339–348. [[CrossRef](#)]
7. Yuan, K.; Wan, Q.; Chai, B.; Lei, X.; Kang, A.; Chen, J.; Chen, X.; Shi, H.; He, L.; Li, M. Characterizing the effects of stormwater runoff on dissolved organic matter in an urban river (Jiujiang, Jiangxi province, China) using spectral analysis. *Environ. Sci. Pollut. Res.* **2023**, *30*, 50649–50660. [[CrossRef](#)] [[PubMed](#)]
8. Lin, Z.Y.; Lim, J.Y.; Loy, A.C.M.; Park, J.Y.; Oh, J.-M. Unraveling the Behavior of Stormwater Runoff at Interception Facilities: An Explicit Focus on Dissolved Organic Matter with Analytical Measurement. *ACS ES&T Water* **2024**, *5*, 398–412. [[CrossRef](#)]
9. Du, X.; Zhu, Y.; Han, Q.; Yu, Z. The influence of traffic density on heavy metals distribution in urban road runoff in Beijing, China. *Environ. Sci. Pollut. Res.* **2019**, *26*, 886–895. [[CrossRef](#)] [[PubMed](#)]
10. Chen, H.; Ji, X.N.; Li, J.; Li, Y.; Zhang, T.; Zhou, W. Microbial metabolism-mediated transformation of dissolved organic matter: Linking antecedent dry-weather degradation to first-flush effects in urban wet-weather overflows. *J. Environ. Chem. Eng.* **2026**, *14*, 120845. [[CrossRef](#)]
11. Qian, C.; Wang, L.-F.; Chen, W.; Wang, Y.-S.; Liu, X.-Y.; Jiang, H.; Yu, H.-Q. Fluorescence Approach for the Determination of Fluorescent Dissolved Organic Matter. *Anal. Chem.* **2017**, *89*, 4264–4271. [[CrossRef](#)] [[PubMed](#)]
12. Boguta, P.; Pieczywek, P.M.; Sokołowska, Z. A Comparative Study of the Application of Fluorescence Excitation-Emission Matrices Combined with Parallel Factor Analysis and Nonnegative Matrix Factorization in the Analysis of Zn Complexation by Humic Acids. *Sensors* **2016**, *16*, 1760. [[CrossRef](#)] [[PubMed](#)]

13. Liu, D.; Nie, L.; Xi, B.; Gao, H.; Yang, F.; Yu, H. A novel-approach for identifying sources of fluvial DOM using fluorescence spectroscopy and machine learning model. *NPJ Clean. Water* **2024**, *7*, 79. [[CrossRef](#)]
14. Lee, J.H.; Bang, K.W.; Ketchum, L.H.; Choe, J.S.; Yu, M.J. First flush analysis of urban storm runoff. *Sci. Total Environ.* **2002**, *293*, 163–175. [[CrossRef](#)] [[PubMed](#)]
15. Ministry of Environmental Protection. *Water Quality—Determination of Ammonia Nitrogen—Nessler’s Reagent Spectrophotometry: HJ 535-2009*; China Environmental Science Press: Beijing, China, 2009.
16. State Environmental Protection Administration. *Water Quality—Determination of Suspended Solids—Gravimetric Method: GB 11901-89*; China Environmental Science Press: Beijing, China, 1989.
17. Westerhoff, P.; Chen, W.; Esparza, M. Fluorescence analysis of a standard fulvic acid and tertiary treated wastewater. *J. Environ. Qual.* **2001**, *30*, 2037–2046. [[CrossRef](#)] [[PubMed](#)]
18. Yang, L.; Chang, S.; Shin, H.; Hur, J. Tracking evolution of stream DOM source during storm events using end member mixing analysis based on DOM quality. *J. Hydrol.* **2015**, *523*, 33–34. [[CrossRef](#)]
19. Lawson, C.L. Software for C¹ Surface Interpolation. In *Mathematical Software III*; Rice, J.R., Ed.; Academic Press: New York, NY, USA, 1977; pp. 161–194. [[CrossRef](#)]
20. Zepp, R.G.; Sheldon, W.M.; Moran, M.A. Dissolved organic fluorophores in southeastern US coastal waters: Correction method for eliminating Rayleigh and Raman scattering peaks in excitation–emission matrices. *Mar. Chem.* **2004**, *89*, 15–36. [[CrossRef](#)]
21. Cai, Y.; Gu, H.; Kenney, T. Rank selection for non-negative matrix factorization. *Stat. Med.* **2023**, *42*, 5676–5693. [[CrossRef](#)] [[PubMed](#)]
22. Yu, S.; Zhang, Y.; Liu, W.; Zhao, N.; Xiao, X.; Yin, G. Component Recognition with Three-Dimensional Fluorescence Spectra Based on Non-Negative Matrix Factorization. *J. Chemom.* **2011**, *25*, 586–591. [[CrossRef](#)]
23. Li, X.; Zhou, Z.; Wang, X. Rapid Factorization of Single EEM for Dissolved Organic Matter Analysis. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2026**, *345*, 126752. [[CrossRef](#)]
24. Genuer, R.; Poggi, J.M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [[CrossRef](#)]
25. Rossi, F.; Lendasse, A.; François, D.; Wertz, V.; Verleysen, M. Mutual information for the selection of relevant variables in spectrometric nonlinear modelling. *Chemom. Intell. Lab. Syst.* **2006**, *80*, 215–226. [[CrossRef](#)]
26. Xiaobo, Z.; Jiewen, Z.; Povey, M.J.W.; Holmes, M.; Hanpin, M. Variables selection methods in near-infrared spectroscopy. *Anal. Chim. Acta* **2010**, *667*, 14–32. [[CrossRef](#)] [[PubMed](#)]
27. Wang, H.Y.; Liu, M.; Liu, Q.M.; Hou, L.J.; Ou, D.L. First flush effect of urban storm runoff and runoff pollution management. *Adv. Water Sci.* **2006**, *17*, 181–185. [[CrossRef](#)]
28. Pettitt, A.N. A nonparametric approach to the change-point problem. *Appl. Stat.* **1979**, *28*, 126–136. [[CrossRef](#)] [[PubMed](#)]
29. An, Z.Q.; Qiao, X.C.; Liu, Y.; Lu, Y.; Li, X.; Wang, H. Influencing factors and control suggestions of overflow pollution in combined sewer areas. *J. Environ. Eng. Technol.* **2024**, *14*, 986–994. [[CrossRef](#)]
30. Yuan, D.; Pang, Y.; Zhai, L.; Yan, C.; Kou, Y. Exploring the combination characteristics of dissolved organic matter with erythromycin in a soil infiltration system. *Environ. Technol.* **2024**, *16*, 3263–3275. [[CrossRef](#)] [[PubMed](#)]
31. Lusk, M.G.; Garzon, P.S.; Muni-Morgan, A. Nitrogen forms and dissolved organic matter optical properties in bulk rainfall, canopy throughfall, and stormwater in a subtropical urban catchment. *Sci. Total Environ.* **2023**, *896*, 165243. [[CrossRef](#)] [[PubMed](#)]
32. Hosen, J.D.; McDonough, O.T.; Febria, C.M.; Palmer, M.A. Dissolved organic matter quality and bioavailability changes across an urbanization gradient in headwater streams. *Environ. Sci. Technol.* **2014**, *48*, 7817–7824. [[CrossRef](#)] [[PubMed](#)]
33. Li, X.L.; Zhou, Z.W.; Jia, H.F.; Li, Z.; Yang, Z.; Cai, Z.; Zhou, H.; Shi, X. Mechanisms of accumulation–transport–discharge and source apportionment of combined sewer overflow pollution. *Water* **2026**, *18*, 573. [[CrossRef](#)]
34. Cui, B.; Gao, H.-J.; Zheng, Z.-P.; Yu, H.-B. Applying EEM Coupled with 2D-COS to Character Structural Composition of DOM and Its Spatial Variations in an Urban River. *J. Ecol. Rural Environ.* **2021**, *37*, 369–377. [[CrossRef](#)]
35. Shimotori; Koichi; Watanabe, K.; Hama, T. Fluorescence characteristics of humic-like fluorescent dissolved organic matter produced by various taxa of marine bacteria. *Aquat. Microb. Ecol.* **2012**, *65*, 249–260. [[CrossRef](#)]
36. Li, W.; Li, X.; Han, C.X.; Gao, L.; Wu, H.; Li, M. A new view into three-dimensional excitation-emission matrix fluorescence spectroscopy for dissolved organic matter. *Sci. Total Environ.* **2023**, *855*, 158963. [[CrossRef](#)] [[PubMed](#)]
37. Lloyd, J.B.F.; Evett, I.W. Prediction of peak wavelengths and intensities in synchronously excited fluorescence emission spectra. *Anal. Chem.* **1977**, *49*, 1710–1715. [[CrossRef](#)]
38. Song, F.H.; Wu, F.C.; Feng, W.Y.; Tang, Z.; Giesy, J.P.; Guo, F.; Shi, D.; Liu, X.; Qin, N.; Xing, B.; et al. Fluorescence regional integration and differential fluorescence spectroscopy for analysis of structural characteristics and proton binding properties of fulvic acid sub-fractions. *J. Environ. Sci.* **2018**, *74*, 116–125. [[CrossRef](#)] [[PubMed](#)]
39. Gabor, R.S.; Baker, A.; McKnight, D.M.; Miller, M.P. Fluorescence indices and their interpretation. In *Handbook of Environmental Chemistry*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 303–338. [[CrossRef](#)]

40. Trubetskaya, O.E.; Richard, C.; Trubetskoj, O.A. High amounts of free aromatic amino acids in the protein-like fluorescence of water-dissolved organic matter. *Environ. Chem. Lett.* **2016**, *14*, 495–500. [\[CrossRef\]](#)
41. Coble, P.G. Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. *Mar. Chem.* **1996**, *51*, 325–346. [\[CrossRef\]](#)
42. Murphy, K.R.; Stedmon, C.A.; Graeber, D.; Bro, R. Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Anal. Methods* **2013**, *5*, 6557–6566. [\[CrossRef\]](#)
43. Jensen, D.M.R.; Sandoval, S.; Aubin, J.B.; Bertrand-Krajewski, J.-L.; Xuyong, L.; Mikkelsen, P.S.; Vezzaro, L. Classifying pollutant flush signals in stormwater using functional data analysis on TSS M(V) curves. *Water Res.* **2022**, *217*, 118394. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Chen, H.; Guo, Y.; Fang, N.; Ye, J.; Tang, J.; Zhu, Y.; Li, J.; Li, Y.; Zhang, T.; Zhou, W. Illicit Sewage Connection Fuels CH₄ and CO₂ Emissions in Storm Sewer: Tracking from Sources to Mechanical Influences. *ACS ES&T Eng.* **2025**, *5*, 2613–2625. [\[CrossRef\]](#)
45. Ao, F.Y.; Ye, C.; Xu, Y.L.; Wu, Z.; Xu, Z.; Chu, W. Advancing illicit connection diagnosis of urban stormwater pipes: Comprehensive analysis with EEM fluorescence spectroscopy. *Water Res.* **2024**, *258*, 121793. [\[CrossRef\]](#) [\[PubMed\]](#)
46. van der Werf, J.A.; Kapelan, Z.; Langeveld, J. Towards the long term implementation of real time control of combined sewer systems: A review of performance and influencing factors. *Water Sci. Technol.* **2022**, *85*, 1295–1320. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Weerasinghe, M.; Jayathilaka, K.; Vithanage, M. Sensors for detection and monitoring of contaminants in wastewater. *Curr. Opin. Environ. Sci. Health* **2025**, *45*, 100609. [\[CrossRef\]](#)
48. van der Werf, J.A.; Kapelan, Z.; Langeveld, J. Realtime control of combined sewer systems: Risks associated with uncertainties. *J. Hydrol.* **2023**, *617*, 128900. [\[CrossRef\]](#)
49. Booth, A.; Fleck, J.; Pellerin, B.A.; Hansen, A.; Etheridge, A.; Foster, G.M.; Graham, J.L.; Bergamaschi, B.A.; Carpenter, K.D.; Downing, B.D.; et al. *Field Techniques for Fluorescence Measurements Targeting Dissolved Organic Matter, Hydrocarbons, and Wastewater in Environmental Waters: Principles and Guidelines for Instrument Selection, Operation and Maintenance, Quality Assurance, and Data Reporting*; U.S. Geological Survey Techniques and Methods; Book 1; US Geological Survey: Reston, VA, USA, 2023; Chapter D11. [\[CrossRef\]](#)
50. Mladenov, N.; Biggs, T.; Ford, K.; Garcia, S.; Yuan, Y.; Grant, A.; Piazza, E.; Rivera, E.; Pinongcos, F.; Keely, S.P.; et al. Evaluation of Real-Time Fluorescence Sensors and Benchtop Fluorescence for Tracking and Predicting Sewage Contamination in the Tijuana River Estuary at the US–Mexico Border. *Sci. Total Environ.* **2024**, *950*, 175137. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Huang, X.; Fu, X.; Zhao, Z.; Yin, H. The Telltale Fluorescence Fingerprints of Sewer Flows for Interpreting the Low Influent Concentration in Wastewater Treatment Plant. *J. Environ. Manag.* **2024**, *349*, 119517. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Downing, B.D.; Pellerin, B.A.; Bergamaschi, B.A.; Saraceno, J.F.; Kraus, T.E.C. Seeing the light: The effects of particles, dissolved materials, and temperature on in situ measurements of DOM fluorescence in rivers and streams. *Limnol. Oceanogr. Methods* **2012**, *10*, 767–775. [\[CrossRef\]](#)
53. Du, M.; Chen, W.; Qian, C.; Chen, Z.; Liu, Y.; Hu, S.; Li, H. Using Rayleigh scattering to correct the inner filter effect of the fluorescence excitation–emission matrix. *Anal. Chem.* **2023**, *95*, 12273–12283. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Ishii, S.K.L.; Boyer, T.H. Behavior of reoccurring PARAFAC components in fluorescent dissolved organic matter in natural and engineered systems: A critical review. *Environ. Sci. Technol.* **2012**, *46*, 2006–2017. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Du, X.; Chi, Z.; Chen, M.; Yu, Z.; Zhu, Y. Characteristics of dissolved organic matter in urban road runoff under different traffic densities in Beijing, China. *Environ. Sci. Pollut. Res.* **2022**, *29*, 90520–90529. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Ly, Q.V.; Truong, V.H.; Ji, B.X.; Nguyen, X.C.; Cho, K.H.; Ngo, H.H.; Zhang, Z. Exploring potential machine learning application based on big data for prediction of wastewater quality from different full-scale wastewater treatment plants. *Sci. Total Environ.* **2022**, *832*, 154930. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Lund, N.S.V.; Falk, A.K.V.; Borup, M.; Madsen, H.; Mikkelsen, P.S. Model predictive control of urban drainage systems: A review and perspective towards smart realtime water management. *Crit. Rev. Environ. Sci. Technol.* **2018**, *48*, 279–339. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.