Environmental Science & Technology

Ecotoxicology and Human Environmental Health

Experimental investigation and modelling of the transformation of illicit drugs in a pilot-scale sewer system

Jiaying Li, Jianfa Gao, Phong K. Thai, Adam Shypanski, Ludwika Nieradzik, Jochen F. Mueller, Zhiguo Yuan, and Guangming Jiang

Environ. Sci. Technol., **Just Accepted Manuscript** • DOI: 10.1021/acs.est.8b06169 • Publication Date (Web): 11 Mar 2019

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

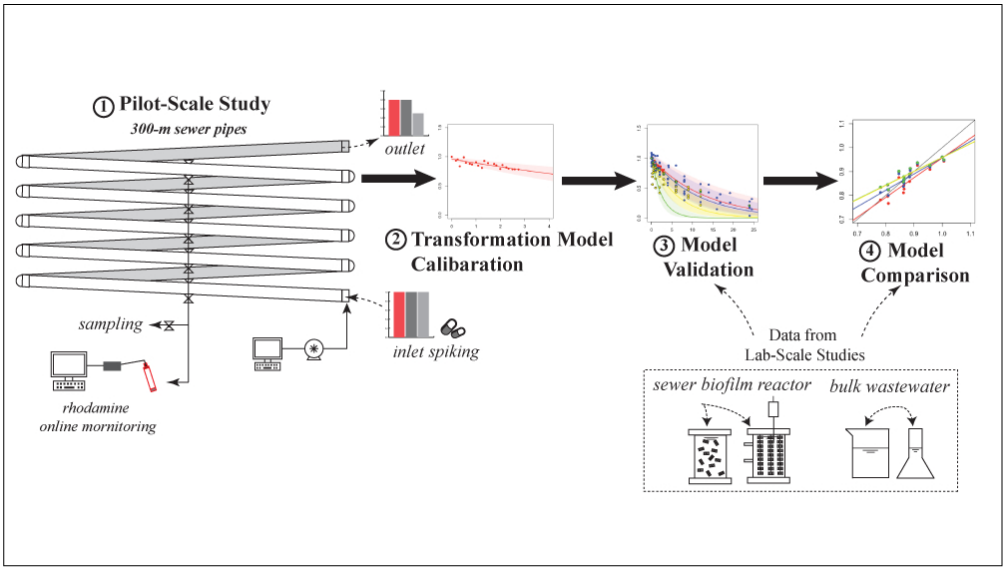


Table of Contents

Experimental investigation and modelling of the transformation of illicit drugs in a pilot-scale sewer system

Jiaying Li[a], Jianfa Gao[b], Phong K. Thai[b], Adam Shypanski[a], Ludwika Nieradzik[a], Jochen F. Mueller[b], Zhiguo Yuan[a], Guangming Jiang[a,c,]*

[a]Advanced Water Management Centre, The University of Queensland, St Lucia, QLD 4072, Australia

[b]Queensland Alliance for Environmental Health Sciences, The University of Queensland, Brisbane, QLD 4102, Australia

[c]School of Civil, Mining and Environmental Engineering, University of Wollongong, Wollongong, NSW 2522, Australia

KEYWORDS

Wastewater-based epidemiology, illicit drugs, biotransformation, pilot sewer system, model validation

ABSTRACT

In-sewer stability of illicit drug biomarkers has been evaluated by several reactor-based studies but less has been done in sewer pipes. Experiments conducted in sewer pipes have advantages over lab-scale reactors in providing more realistic biomarker stability due to the flow and biological dynamics. This study assessed the transportation and transformation of seven illicit drug biomarker compounds in a pilot-scale rising main and a gravity sewer pipe. Biomarkers presented diverse stability patterns in the pilot sewers, based on which a drug transformation model was calibrated. This model was subsequently validated using transformation datasets from literature, aiming to demonstrate the predictability of the pilot-based transformation coefficients under varying sewer conditions. Furthermore, transformation coefficients for five investigated biomarkers were generated from four studies and their prediction capabilities under the pilot sewer conditions were jointly assessed using performance statistics. The transformation model was successful in simulating the in-sewer stability for most illicit drugs. However, further study is required to delineate the sources and pathways for those compounds with potential formations to be simulated in the transformation model. Overall, the transformation model calibrated using the pilot-sewer data is a credible tool for the application of wastewater-based epidemiology.

INTRODUCTION

Wastewater-based epidemiology (WBE) has been developed rapidly over the last decade in a bid to achieve objective and timely assessment of community health and consumption behaviours via analysing trace levels of substances (termed as

biomarkers) in wastewater, including illicit drugs, pharmaceuticals and new psychoactive substances.[1] Back-estimating the catchment-wide usage of illicit drugs is an emerging area of WBE, which is demonstrated to be a useful complementary tool to conventional drug monitoring approaches.[2-5] In order to improve the accuracy of back-estimation, researches have been widely conducted to address the uncertainties associated with sampling method and chemical analysis,[5-9] while a comprehensive understanding of biomarkers stability in real sewers is still ongoing.[10-17] Biomarkers are subjected to physiochemical and biological processes during their transport in rising main and gravity sewer pipes, where the hydraulic retention time (HRT) may last for hours.[18] Neglecting the biomarkers transformation (e.g., the degradation or formation in sewers) will lead to an under- or over-estimation of drug consumption in a catchment.[4, 10, 12, 13, 15, 19] This uncertainty varies depending on the stability of biomarkers and the characteristics of sewer systems such as HRT distributions, which was suggested to be negligible for the stable biomarkers[8] but significantly increased for the unstable biomarkers with >40% median mass losses in the catchments.[15]

Depending on the different experimental scales and conditions utilized, research on biomarker stability can be divided into four categories: 1) *in-water study* using clean/sterile bottles or other containers where abiotic processes such as chemical hydrolysis occur in water;[14, 20, 21] 2) *in-wastewater study* that is conducted in raw wastewater where suspended biomass and certain microbial activities contribute to

the transformation process;[12, 14, 21-25] 3) *sewer reactor study* employing lab reactors with intact/suspended biofilms or activated sludge to mimic the biologically active sewer environments;[10-12, 14, 16, 19] 4) *real sewer pipe study* using the sewer pipes with the same (or similar) size and operational conditions as the real sewer networks.[10, 15, 17, 26, 27]

Among the *real sewer pipe studies*, two of them assessed the change of biomarkers using 24-h composite samples, however, with limited understanding of the concurrent in-pipe hydraulics or biological activities.[15, 26] Two other studies spiked biomarkers in a real rising main pipe and evaluated their variations from the pipe upstream to a downstream sampling point, coupled with the measurements of flow dynamics and biological activities in the pipe. However, due to the poor accessibility of real sewer pipes and the narrow HRT windows, the obtained data points were insufficient for kinetics evaluation.[10, 27] In comparison to the inherent limitations of static lab-reactors and complex real sewers[28], a pilot-scale sewer can be more beneficial for the study of biomarker stability in sewer pipes by providing multiple sampling points, online monitoring, controllable flow, and other environmental factors. This has been demonstrated in a recent *real sewer pipe study* investigating the fate of pharmaceutical biomarkers.[17]

The temporal transformation data obtained in pilot-scale sewers can be very valuable for the modelling of biomarker stability. Transformation modelling is a useful approach with which to utilize information on biomarker stability in the application

of WBE in reality. So far, three *sewer reactor studies* investigated the transformation modelling of illicit drugs,[10, 14, 16] but only one study validated the estimated transformation coefficients in a real rising main pipe.[10] It is thus imperative to calibrate the illicit drug transformation model using dynamic data from sewer pipes (e.g. pilot-scale sewers) instead of lab reactors. More importantly, the transformation coefficients derived from different studies need to be systematically evaluated for their transferability across diverse sewer conditions.

The first objective of this study is to measure and model the stability of illicit drug biomarkers in the pilot-scale sewer pipes. Experiments were conducted in a pilot-scale sewer system, including a rising main pipe and a gravity sewer pipe with online monitoring and control system. A drug transformation model was calibrated with the pilot-sewer data and was subsequently validated using literature data under varying sewer conditions. The second objective is to systematically evaluate the transformation coefficients generated by previous studies, through the comparison of the prediction capabilities using performance statistics. Collectively speaking, this work intends to advance WBE through not only providing valuable data on illicit drug transformation in sewer pipes, but also enhancing the generalizability and applicability of the transformation model to the application of WBE.

MATERIALS AND METHODS

Compounds for Investigation

This study investigated the parent compounds of major illicit drug biomarkers by spiking them into the pilot-scale sewer pipes, including cocaine (COC), ketamine (KET), 3,4-methylenedioxymethamphetamine (MDMA), morphine (MOR), and methadone (MTD). According to the analysis results of wastewater samples at the experiment site, the spiking concentrations (2.5-8 ppb) for most biomarkers were higher than their native residues in raw wastewater (<0.5 ppb). The native methamphetamine (METH) was investigated without being spiked because of its relatively high background concentrations (around 1 ppb). The native benzoylecgonine (BE) was evaluated as a specific metabolite of COC without being spiked. Other metabolites of the parent compounds were not spiked in separate tests. The flow tracer rhodamine was added into the spiking mixture solution. Rhodamine signal was measured by a portable Cyclops®-7 Submersible Rhodamine Sensor coupled with a Cyclops® Explorer.

Tests in the pilot-scale sewer system

Experiments were conducted in a pilot sewer system located at the Luggage Point Wastewater Treatment Plant, Queensland, Australia. This study employed two 300-m long sewer pipes, one rising main and one gravity pipe, on a controllable platform equipped with programmable logic controller, pumps, meters, sensors, etc.[17, 29] The experimental procedures in this study were the same as reported previously.[17] The layout and parameters of the sewer system are illustrated in Supplementary Information S1.1.

The ***Rising main pipe*** is completely filled with wastewater and anaerobic biofilm were cultivates on the pipe's inner surface with a thickness of 1~2 mm.[29] The internal diameter is 100 mm, leading to a biofilm-area-to-wastewater-volume (A/V) ratio of 40 m^{-1}. This rising main pipe is constructed to spiral up from the ground layer (inlet) to the top layer (outlet) and wastewater is driven by pumping events in a plug-flow regime. As shown in Figure S1.1, eight sampling ports are distributed along the sewer pipe, i.e. at 0, 15, 45, 75, 105, 135, 195, and 240 m from the inlet for port #1 to #8, respectively.

Triplicate batch tests were carried out over 3 consecutive days (Day 1, 2 and 3) in the rising main. To achieve a typical hydraulic condition, the main pump was turned on for 1 min every hour with a flow rate of 236 L min^{-1}, producing a wastewater slug of 30 m and an intermittent shear stress of 0.6 N m^{-2} in the pipe. With the first pumping event of each test, a mixture solution of biomarkers and rhodamine was spiked into influent using an external peristaltic pump, resulting in the first and the only spiked wastewater slug at time 0 (t_0). This spiked wastewater slug was pushed 30 m downstream by the subsequent non-spiked wastewater slugs at every following pumping event and arrived at the final sampling port after 7 h of HRT in the pipe. Through matching the length of every wastewater slug with the distance between two sampling ports (i.e. 30 m), the central area of this spiked slug could be captured at #1 to #8 sampling ports in sequence after each hourly pumping event. Meanwhile, during every 1-h pump-off period, samples of the spiked wastewater were collected

at 15 min intervals through the sampling port where the spiked slug was located. Concurrently with the transportation of the spiked wastewater slug, a rhodamine sensor was connected to #1 to #8 sampling ports for the monitoring of the rhodamine concentration.

The ***Gravity sewer pipe*** comprises both water and air phases, between which gas transfer leads to 1~4 mg L^{-1} dissolved oxygen in the bulk liquid phase. A removable section of pipe showed the prevailing existence of sediments at the bottom. This gravity sewer pipe has an internal diameter of 225 mm and is constructed to spiral down from the top layer (inlet) to the ground layer (outlet) with a slope of 0.56%, where wastewater flow is driven by gravitation. The traveling time of wastewater from the pipe inlet to the outlet ranged around 8~10 min. Moreover, this pilot gravity sewer pipe has the unique capability of allowing recirculation. Under the recirculation mode, wastewater effluent at the pipe outlet is collected in a recirculation tank and re-directed to the feeding tank by a recirculation pump. The recirculation mode enables wastewater to flow in the gravity sewer pipe as long as required. The average flow velocity was 0.38 m s^{-1}; the in-pipe water depth was around 5 cm; and the average A/V ratio was approximately 27.5 m^{-1} (S1.2 and Figure S1.2). The average shear stress under this flow condition was 0.5 N m^{-2}. A sampling port is installed near the outlet, allowing access to wastewater in the pipe (Figure S1.1).

Triplicate tests were conducted over 3 consecutive days (Day 4, 5 and 6) in the gravity sewer pipe under recirculation mode. At the beginning of each test, a mixture solution of biomarkers and rhodamine was directly spiked into the pipe and continuously mixed with the flowing wastewater. Meanwhile, a rhodamine sensor was connected to the sampling port for online readings during experimental periods. As shown in preliminary tests, the spiked rhodamine presented 2~3 signal peaks over the first few cycles of recirculation and then a sufficiently mixed stage appeared at 1.5 h after the spiking event as indicated by the consistent signal intensity. Therefore, for experiments in the recirculating gravity sewer, t_0 of the non-spiked biomarker commenced at the beginning of each test. However, t_0 of the spiked biomarkers was defined when a homogeneously mixed status was reached, in order to minimize the uncertainty of mixing on the evaluation of the spiked biomarkers stability. The HRTs of the recirculating wastewater in the gravity sewer pipe were 3~4 h, during which wastewater samples were collected at 15 min intervals.

The collected samples were pretreated on site for the analyses of biomarkers and wastewater parameters, including sulfur species, dissolved methane, volatile fatty acids (VFAs), soluble chemical oxygen demand (SCOD), total and volatile suspended solids (TSS and VSS) (for sample pre-treatment and analytical methods see S2). Temperature and pH of samples were measured on site using a portable pH/temperature meter (TPS Aqua-pH pH/Temp meter).

Calibration of Drug Transformation Model

As widely applied in previous studies,[10, 14, 30-32] a first-order kinetics is adopted for illicit drug transformation model in this work (eq 1):

$$C_t = C_0 \cdot e^{-\left(k_{ww} + k'_{bio} \cdot \frac{A}{V}\right) \cdot t} \quad (1)$$

C_t is biomarker concentration (μg L^{-1}) at time t (h) and C_0 is the initial concentration (μg L^{-1}). k_{ww} (h^{-1}) represents the processes in the bulk liquid wastewater, mainly chemical hydrolysis assuming that the sorption to suspended solids or biofilm is limited for investigated biomarkers[5, 14, 16, 33] and the biological activity of suspended solids is negligible compared to the sewer biofilms or sediments.[34-36] It is further postulated that k_{ww} remains the same under aerobic and anaerobic conditions, according to the findings of McCall *et al.*[14] and Ramin *et al.*[16] (about the minor impact of redox condition on abiotic transformation rates). k'_{bio} (m h^{-1}) includes k'_{bioa} and k'_{bioan}, representing the biofilm effect under aerobic and anaerobic condition with the normalization of A/V ratio, respectively. The effects of mass transfer limitation on k'_{bio} is considered to be negligible for illicit drug compounds under the experimental conditions as discussed in S1.2. In conclusion, the overall in-sewer transformation rate (h^{-1}) depends on k_{ww} and the k'_{bio} coupled with a specific A/V ratio. Consequently, the percentage contributions from wastewater and biofilm processes to the overall transformation of a biomarker are quantified by the ratio of k_{ww} against $k'_{bio} \cdot \frac{A}{V}$. This study assumed that the mature sewer biofilms in the pilot

sewers were under steady state with negligible biomass growth over the course of experiments.

To calibrate the transformation model, experimental datasets obtained from the pilot rising main pipe (Day 1 and 2) and the pilot gravity sewer (Day 4 and 5) pipe were used to estimate k'_{bioan} and k'_{bioa}, respectively, through a Bayesian procedure described in Li *et al.*[10] The datasets of Day 3 (rising main) and Day 6 (gravity sewer) were reserved for the subsequent model validation. As shown in a previous field-scale study,[10] rhodamine signal was used to normalize the biomarker concentrations which effectively minimized the potential hydraulic uncertainty during wastewater transport. Therefore, for the spiked biomarkers in this pilot study, C_t is normalized by the ratio of the initial rhodamine signal against the signal at t. k_{ww} is estimated using the data on biomarkers transformation in the bulk wastewater (Figure S3.1), which is collected from the experiments using the control sewer reactor without biofilms and from previous *in-wastewater studies*.[10, 12] The estimation of k'_{bio} is carried out in R (Version 3.2) which executes the Bayesian method in OpenBUGS (http://www.openbugs.net).

Validation of the pilot-based model

In order to evaluate the validity of the transformation coefficients (*k* values) estimated by the abovementioned pilot-based model across diverse sewer environments, experiment datasets obtained from the pilot rising main at Day 3 and the pilot gravity sewer at Day 6 together with the literature data collected from three

222 independent sewer reactor studies[10, 14, 16] are used as the observation inputs
223 representing different testing sewer conditions (datasets in Table S3.2). For each
224 investigated biomarker, the pilot-based *k* values are used to generate predictions
225 (mean with 95% confidence bounds) with the specific A/V ratios of the four studies
226 (for descriptions of prediction scenarios see Table S3.2). Under each scenario,
227 rejection probability is determined by counting the percentage of experimental
228 observations (i.e. the data on biomarker transformation) located outside the
229 corresponding predictive region. The lower rejection probability reflects the higher
230 validity of the pilot-based *k* values.

231 **Performance comparisons of different *k*-value sets**

232 To date, several lab-scale studies have assessed the in-sewer stability of illicit drug
233 biomarkers and estimated the transformation coefficients under their specific testing
234 conditions.[10, 14, 16] Although these *k*-value sets demonstrated adequate prediction
235 performance in the corresponding studies, their predictive abilities have neither been
236 validated beyond their original testing conditions nor jointly compared under the
237 realistic sewer conditions. Importantly, measurements of this pilot study are valuable
238 to model validation, which were obtained from an enlarged pilot-scale system with
239 typical operations and dynamic hydraulics of real sewer systems.

240 For this reason, multiple *k*-value sets ($k_{ww}^{M_i}$ and $k_{bio}^{\prime M_i}$) are obtained from four different
241 stability studies, i.e., this pilot study and three *sewer reactor studies* (defined as
242 M1~M4 in Table S3.2), and their prediction capabilities are jointly compared under

the conditions of the pilot-scale sewer system. Multiple prediction scenarios (mean with 95% confidence bounds) under the conditions of the pilot rising main and the gravity sewer are generated for each investigated biomarker (Table S3.2). Furthermore, in order to identify the *k*-value set with the highest agreement of prediction to the observations, performance statistics are computed for the predictive scenarios of M1~M4 under each condition, including a stochastic validation metric (the Bayes factor) and two accuracy measures (Pearson correlation coefficient *r* and variance explained by predictive models based on cross-validation (*VEcv*)) calculated using RStudio (Version 1.0.143).

1) The Bayes Factor

The Bayes Factor (BF) is a typical model validation metric, which quantitatively measures the agreement between predictions and measurements together with their internal uncertainties based on stochastic characteristics.[37] As explained by eq 2, the Bayes Factor quantifies the ratio of the probabilities of observations under null hypothesis H_0 and alternative hypothesis H_1, respectively, at each validation site (for visual illustrations see Figure S3.2):

$$BF = \frac{P(\text{data} \mid H_0)}{P(\text{data} \mid H_1)} \quad (2)$$

where H_0 is the null hypothesis representing the better match between the observations in pilot sewers and the predictions of a calibrated model, e.g. M1; H_1 is the alternative hypothesis representing the better match between observations and

the predictions of other competing models, e.g., M2~M4. For each biomarker under the pilot gravity sewer or the rising main condition, up to three log(BFs) can be computed depending on the available *k*-value sets in literature, i.e., log(BF12), log(BF13), and log(BF14). At a validation site, log(BF)>0 means H_0 is true, indicating that predictions by the *k*-value set of M1 is favoured by the observed data. A larger absolute value of log(BF) suggests the higher preference to the pilot-based M1 when log(BF)>0, or to the reactor-based M2, M3, or M4 when log(BF)< 0.

2) Accuracy of model predictions

The Pearson correlation coefficient *r* is widely used to assess model predictions against observed data via quantifying the strength and direction of a linear relationship between the two variables (eq 3). When observations x are perfectly linearly related to model predictions y (i.e., $y = \beta_0 + \beta_1 x$, β_0 and β_1 are coefficients), *r* is a suitable indicator of predictive accuracy:[38]

$$r = \frac{\sum_1^n (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_1^n (x_i - \overline{x})^2 (y_i - \overline{y})^2}} \quad (3)$$

where n is the number of observations; x_i is the observed value i; $\overline{x}$ is the mean of the observed values; y_i is the predicted value i; $\overline{y}$ is the mean of the predicted values.

However, when x and y are not well correlated, i.e., when noticeable noise ε appears in $y = \beta_0 + \beta_1 x + \varepsilon$, *r* becomes potentially biased and *VEcv* (eq 4) is recommended as the correct measure of predictive accuracy instead:[38]

$$VEcv = (1 - \frac{\sum_1^n (x_i - y_i)^2}{\sum_1^n (x_i - \bar{x})^2}) \times 100(\%) \quad (4)$$

RESULTS AND DISCUSSION

Wastewater Compositions and Biological Activities

The observations of wastewater parameters in this study were the same as reported in Gao *et al.*,[17] which conducted experiments in the same setup. The sewer characteristics in the pilot sewer system, in terms of wastewater parameters and biological activities, are also compared to the literature data. It is found that the variations of wastewater parameters in the pilot sewers (Figure S1.3) are similar to those in other sewer studies (Table S1.2).[10, 39, 40] For instance, wastewater pH kept relatively stable over 8 h in the pilot rising main (6.99±0.11), while pH in the pilot gravity sewer increased by ~0.25 units during the first 2 h, likely due to the CO_2 and H_2S stripping. Wastewater temperature remained consistently stable in the pilot sewers (22.9±0.6 ºC). TSS was higher in the pilot gravity sewer (500-800 mg L^{-1}) than the pilot rising main (200-410 mg L^{-1}) due to resuspension of sediments. The interday deviations of TSS were attributed to the daily variation of real wastewater, while the intraday changes of TSS were relatively insignificant over the experimental period.

Based on previous studies, sulfate reducing bacteria and methanogenic archaea are the primary microorganisms responsible for not only the carbon/sulfur transformation, but also the biomarker stability in biofilms.[14, 18, 41, 42] In this work,

sulfide and methane production rates are used as the major indicators of biological activities in the pilot sewers, which are comparable to the literature data (Table S1.2). The simultaneously increasing pattern of sulfide and methane in the pilot rising main (Figure S1.3) was also similar to that in real sewers.[43-46] In contrast, sulfide or methane production was detected to be negligible or even negative in the pilot gravity sewer, which could be explained by the faster sulfide oxidation than the concurrent sulfate reduction and/or the emission of sulfide and methane gas into the air phase.

VFAs and SCOD are the major substrates for heterotrophic bacteria and methanogens, which present varying consumption rates in different sewer environments (Table S1.2). Primarily being produced by fermentation process under anaerobic condition,[39, 47] the consumption rate of VFAs in the pilot rising main was lower than that in the gravity sewer (Figure S1.3). The consumption rate of SCOD was higher than VFAs in the pilot rising main, while in the pilot gravity sewer, the concurrent consumption of SCOD and VFAs was generally similar. Overall, this unique pilot sewer system has an environmental condition representative of real sewers.

Biomarkers Stability in the Pilot Sewer System

The online monitoring of rhodamine showed the dynamic flow patterns in the pilot sewer pipes during experiment periods (Figure 1). The rhodamine profiles shared the same results of Gao *et al.*[17] In the pilot rising main, the staged profiles of rhodamine

revealed the plug-flow regime as a response to the intermittent pumping events. In the pilot gravity sewer under recirculation mode, rhodamine profiles reflected the continuous mixing and indicated the commencing of a sufficiently mixed stage at 1.5 h after the spiking event.

Various patterns were found for the investigated biomarkers in the pilot sewer system (Figure 1). Over the consecutive tests in the pilot gravity sewer or the rising main, the interday divergence of the transformations for each biomarker was relatively limited as indicated by the low standard deviations. COC, MDMA, and MTD exhibited decreasing trends during the testing periods in both the pilot rising main and the gravity sewer, where MTD was observed to have the most rapid degradation with ~25% loss in 3 h, followed by ~20% loss of COC and MDMA. During the longer HRT periods in the pilot rising main, these biomarkers still presented similar degradation trends and overall 35~40% losses were observed after 8 h. By contrast, KET and METH showed relatively good stability with <20% losses in both pipes. In addition, formations of BE and MOR were observed over the experimental periods, which were the combined results of multiple concurrent in-sewer processes, including the transformation of biomarker itself, the back-transformation of parent compounds and/or the potential deconjugation of glucuronides in the raw wastewater.[11, 12] Data on the transformations of MOR and MTD in the pilot sewers was also reported previously.[17] Similarly, BE and MOR formations were also found in

other testing sewer conditions, e.g., in lab reactors[11, 12, 14, 16, 23, 31, 48] and real sewer pipe.[10]

Drug Transformation Models

Calibration and validation of the transformation model

Calibrations of the transformation model based on the measured data of this pilot study are presented in Figure S3.3 and the values of transformation coefficients are reported as mean with 95% credible intervals (CI) in Table S3.1. Among the estimated k'_{bio} values, the relatively high k'_{bioa} and/or k'_{bioan} for COC, MDMA, and MTD suggest the important effect of biofilm on their stability in sewers. Moreover, for most biomarkers, their k'_{bioa} values are found to be higher than the corresponding k'_{bioan} values, indicating the higher biofilm-specific impact in aerobic condition compared to anaerobic condition. Effect of mass transfer resistance on biodegradation process (k'_{bio}) is considered to be limited for most investigated biomarkers (see discussions in S1.2). Meanwhile, the relatively high k_{ww} for COC, KET, and MTD indicate their comparatively evident decreases in the bulk liquid wastewater due to the processes such as chemical hydrolysis.[10, 12] However, *k* value was not estimated for BE or MOR, which showed significant formations in the pilot sewers (Figure 1). This is because the knowledge on BE or MOR formation was limited since 1) it is difficult to accurately quantify the various contributing sources, such as the parent compounds and/or conjugated forms in the raw wastewater and 2) this study did not spike labelled BE or MOR to exclude the contribution from other precursors and hence to identify their

specific transformation in sewers. In consequence, the transformation model was not calibrated for the prediction of the in-sewer formations of BE or MOR.

Validity of the estimated pilot-based *k* values in diverse sewer environments are subsequently examined. For each investigated biomarker, the measurement datasets obtained from different stability studies and the predictive regions generated by the pilot-based *k* values under the corresponding scenario (e.g., A/V ratio and redox condition) are jointly plotted in Figure 2. Through showing whether the literature data points are encompassed or rejected by the prediction regions, Figure 2 visually reflects the transferability and applicability of the pilot-based *k* values to various sewer conditions. The rejection probability for each scenario is reported in Table S3.3.

The prediction capability of the pilot-based *k* values is successfully validated for KET and METH under all sewer conditions. As shown in Figure 2, the predictive regions almost completely encompass the observations from the four stability studies under aerobic and anaerobic conditions, except for a few data points over long HRTs. Validity of the pilot-based *k* values is successfully demonstrated for COC, MDMA, and MTD in the pilot sewers but is less successful against the testing conditions of the three *sewer reactor studies*. For these three biomarkers, relatively high rejection probabilities of the pilot-based *k* values are determined for different reasons: 1) the overestimation of COC decreases at the very high A/V ratio condition (175 m^{-1}) of Ramin *et al.*[21]; 2) the overestimation of MDMA transformation compared to its observed insignificant changes in lab reactors,[10, 14] suggesting that MDMA could be

partially degraded by certain microbes that existed in the realistic pilot sewers but not in lab-scale reactors. Similarly, McCall *et al.*[14] also found the biofilm specific transformation for MDMA and attributed it to the divergent biofilm growing conditions and the different transformation potentials; 3) the underestimation of MTD losses compared to the observed drastic decreases in the anaerobic reactor of Li *et al.*[10] The diffusion limitation could affect k'_{bio} in the pilot rising main (with intermittent pumping events) due to the high biodegradability of MTD (full discussion in S1.2).

Model performance evaluation I: the different k-value sets

Multiple *k*-value sets ($k_{ww}^{M_i}$ and $k_{bio}^{'M_i}$) are generated from four stability studies as explained in Table S3.2 and illustrated in Figure S3.3. Figure 3 shows the joint distributions of $k_{ww}^{M_i}$ and $k_{bio}^{'M_i}$, which visually reflect the consistency or discrepancy among the *k*-value sets estimated by different studies.

For most investigated biomarkers, the close distributions of $k_{ww}^{M_i}$ indicate the general congruence of $k_{ww}^{M_i}$ derived from different stability studies. Such consistency suggests that the processes in the bulk wastewater (mainly the abiotic physicochemical process) of different sewer environments could lead to similar transformation of biomarkers. However, certain deviations are found for COC, KET, and MTD under aerobic and/or anaerobic conditions, where the centroids of $k_{ww}^{M_1}$ and/or $k_{ww}^{M_2}$ show relative distances with those of $k_{ww}^{M_3}$ and/or $k_{ww}^{M_4}$. The deviations may be attributed to

the assumptions about the independent relationship of k_{ww} with wastewater compositions and properties such as the redox potential or suspended solids concentration. Specific experiments are required to evaluate the effects of wastewater composition, redox potential, and pH on k_{ww} and the associated k_{bio}.

By contrast, distributions of $k_{bio}^{'M_i}$ are more deviated for most investigated biomarkers. Since $k_{bio}^{'M_i}$ represents the contribution of biofilm to biomarker stability in sewers, such deviations reveal that the biofilm-specific effects in different studies are strongly related to the microbial communities, functions, and activities in biofilms. Generally speaking, $k_{bio}^{'M_1}$ and/or $k_{bio}^{'M_2}$ are usually found to be higher than $k_{bio}^{'M_3}$ and/or $k_{bio}^{'M_4}$, except for KET with closer distributions of $k_{bio}^{'M_i}$ at minor levels. This could be explained by the stronger biological activities of biofilms in the pilot sewer pipes and the sewer reactors of Li *et al.*[10] compared to biofilms in the reactors of McCall *et al.*[14] and/or Ramin *et al.*[16] Moreover, $k_{bio}^{'M_1}$ and/or $k_{bio}^{'M_2}$ appear to be generally equivalent for COC, KET, and METH, while present certain dissimilarity for MDMA and MTD under anaerobic and/or aerobic conditions. Such discrepancies could be due to the different microbial communities in biofilms grown in pilot-scale sewer pipes and lab-scale sewer reactors. Also, the continuous mixing condition in the lab reactors[10, 42] could accelerate the transformation of certain biomarkers in biofilms.

The embedded histograms in Figure 3 show the contributions from wastewater and biofilm processes to the overall transformation of biomarkers in the pilot sewer system. For most biomarkers, the *k*-value sets of M1 and M2 usually suggest the

dominant role of biofilms because of the relatively high $k_{bio}^{'M_1}$ and $k_{bio}^{'M_2}$ values, except for MDMA and METH under anaerobic condition and for KET under both aerobic and anaerobic conditions. The *k*-value sets of M3 and M4 also recognize the governing effect of biofilms on MTD under anaerobic condition and on KET, MDMA and METH under both aerobic and anaerobic conditions. The overall KET variation predicted by M3 is actually negligible and akin to the observed high stability of KET under multiple experimental conditions of McCall *et al.*[14] On the contrary, the *k*-value sets of M3 and M4 indicate higher impact of wastewater processes (k_{ww}) on MTD under aerobic condition and on COC under both aerobic and anaerobic conditions. These are consistent with the findings of McCall *et al.*[14] and Ramin *et al.*,[16] suggesting that processes in the bulk wastewater dominated the transformation of MTD (under aerobic condition) and COC, while sewer biofilms did not significantly enhance their transformation.

Model performance evaluation II: the Bayes Factor, r, and VEcv values

Figure 4 shows the Bayes Factor, *r*, and *VEcv* across the predictive scenarios generated by the *k*-value sets of M1~M4 under the pilot sewer conditions (for detail numbers see Table S3.4 to S3.6). Predictions of M1~M4 versus observations in the pilot sewer pipes (Day 3 and Day 6) are jointly plotted for each biomarker in Figure S3.4.

As indicated by the results of performance statistics, the *k*-value set of M1 provides high prediction accuracy for COC under the pilot rising main and for MDMA and

MTD under both pilot gravity sewer and rising main conditions. In these cases, the likelihood of log(BFs) locating in the positive area is dominant, suggesting the closer agreements of pilot-sewer measurements to the outputs of M1 than those of M2~M4 (Table S3.4). Moreover, experimental observations and the outputs of M1 are more closely matched (i.e., close to $y = x$) compared to the correlations for M2~M4. Consistently, higher *r* and *VEcv* values are quantified for M1 (Table S3.5 to S3.6), corroborating the higher prediction accuracy of M1 for the discussed cases. It should be noted that the datasets of Day 3 and Day 6 from the pilot sewers may naturally favour M1, which is calibrated under similar conditions.

Meanwhile, all the *k*-value sets of M1~M4 provide similarly high prediction accuracy for COC under the pilot gravity sewer condition. The performances of M1~M2 are also comparable for KET under the pilot rising main condition, where a strong linear relationship between observations and the predictions of M1 or M2 is found. In these cases, log(BFs) mainly spread out within the area close to zero, implying the comparable prediction outputs of M1~M4 (Table S3.4). Moreover, the *r* and *VEcv* values of M1~M4 are similar and high for the discussed cases, which further verify the comparable and reasonable prediction capabilities of the *k*-value sets derived from different stability studies (Table S3.5 to S3.6).

For KET under the pilot gravity sewer and METH under both pilot gravity sewer and rising main conditions, the predicted overall transformations by all *k*-value sets are limited, i.e. $\leq 20\%$ as shown in Figure S3.4. The prediction accuracy of M1~M4 is

found to be deteriorated. In these cases, log(BFs) distribute overwhelmingly in the positive area or switch between positive and negative at different validation sites (Table S3.4). Performance statistics reveal that the *r* values of M1~M4 (Table S3.5) range from low (0.22 for METH in the pilot gravity sewer), moderate (0.53 for KET in the pilot gravity sewer) to high (0.75 for METH in the pilot rising main). By contrast, the *VEcv* values are low under all scenarios (Table S3.6), which could be interpreted as a consequence of the poor correlations between predictions and observations of KET and METH in the pilot sewers. This is because these two biomarkers had limited but random variations with time, which can hardly be well fitted by the transformation model. Hence, prediction capabilities of M1~M4 tend to be partially acceptable for the relatively stable biomarkers without clear temporal variation pattern in sewers.

On the contrary, validity of the *k*-value sets derived from different stability studies cannot be effectively assessed for BE or MOR because their actual stability in the pilot sewers might be masked by formations from other sources and/or deconjugation processes. Such in-sewer formations bring uncertainty to the back-estimation of the drugs that could be metabolized/transformed to BE (i.e., the consumption of COC) or MOR (e.g., the consumptions of codeine, heroin, and morphine itself in glucuronide form). Information on the sources contributing to BE and MOR formation, such as the concentrations and transformation pathways of

484 their parent compounds and/or conjugated forms, should be determined before the
485 formation processes are incorporated into the transformation model.

486 The comprehensive evaluations of different calibrated transformation models help to
487 choose the *k* values to give the best predictive performance in the application of
488 WBE. Collectively speaking, for the biomarkers presenting evident losses in sewers
489 (e.g., COC and MTD), both lab-scale and pilot-scale studies estimate relatively high *k*
490 values to predict such decreases over time. On the other hand, most stability studies
491 suggest the overall insignificant transformation of KET and METH in sewers. Their
492 predicted transformations are also limited over time based on all *k*-value sets,
493 although with some random variations. Importantly, due to the complex microbial
494 processes in real sewers, the pilot-based *k* values exhibit advantage in predicting the
495 fate of the biomarkers (e.g., MDMA) which are stable in lab-scale reactors but show
496 partial degradation in the pilot sewer pipes, especially over long HRT. Unfortunately,
497 none of the available *k*-value sets are capable of simulating the increasing patterns
498 of biomarkers (e.g., BE and MOR) in pilot sewers. In short, the transformation model
499 calibrated by the measurements in the pilot-scale sewer pipes can be applied as a
500 credible tool for the future WBE study.

501 ASSOCIATED CONTENT

502 **Supporting Information**

Additional information about the layout and parameters of the pilot sewer system, experimental observations of this pilot study, analytical methods for wastewater samples, and results of model validation and comparison are provided. Supporting information is available free of charge via the Internet.

AUTHOR INFORMATION

Corresponding Author

*Guangming Jiang: E-mail: gjiang@uow.edu.au

ACKNOWLEDGEMENT

This research was supported by the ARC Discovery project (DP150100645). Jiaying Li receives the support from China Scholarship Council. Jianfa Gao receives an ARC scholarship (DP150100645). Phong K. Thai was funded by the QUT VC Fellowship during part of this study. Jochen Mueller acknowledges the UQ Research Fellowship. Guangming Jiang is the recipient of an Australian Research Council DECRA Fellowship (DE170100694). We specially acknowledge the collaboration with Queensland Urban Utilities for the measurement campaign at Innovation Centre, Luggage Point Wastewater Treatment Plant. We thank for the assistance of Natasha Rossi and Tobias Hesse during the onsite experiments. We thank Chris Paxman for English proof reading.

REFERENCE

1. Daughton, C. G. Pharmaceuticals and Personal Care Products in the Environment: Overarching Issues and Overview. **2001,** *791*, 2-38.

2. Daughton, C. G., Illicit drugs in municipal sewage: Proposed new nonintrusive tool to heighten public awareness of societal use of illicit-abused drugs and their potential for ecological consequences. In *Pharmaceuticals and Personal Care Products in the Environment: Scientific and Regulatory Issues*, American Chemical Society: Washington DC, 2001; Vol. 791, pp 348-364.

3. Zuccato E; Chiabrando C; Castiglioni S; Calamari D; Bagnati R; Schiarea S; R., F. Cocaine in surface waters- a new evidence-based tool to monitor community drug abuse. *Environ. Health* **2005,** *4* (14), 7.

4. van Nuijs, A. L. N.; Castiglioni, S.; Tarcomnicu, I.; Postigo, C.; de Alda, M. L.; Neels, H.; Zuccato, E.; Barcelo, D.; Covaci, A. Illicit drug consumption estimations derived from wastewater analysis: A critical review. *Sci. Total Environ.* **2011,** *409* (19), 3564-3577.

5. European Monitoring Centre for Drugs and Drug *Assessing illicit drugs in wastewater Advances in wastewater-based drug epidemiology*; Luxembourg, 2016.

6. Ort, C.; Lawrence, M. G.; Rieckermann, J.; Joss, A. Sampling for Pharmaceuticals and Personal Care Products (PPCPs) and Illicit Drugs in Wastewater Systems: Are Your Conclusions Valid? A Critical Review. *Environ. Sci. Technol.* **2010,** *44* (16), 6024-6035.

7. Lai, F. Y.; Ort, C.; Gartner, C.; Carter, S.; Prichard, J.; Kirkbride, P.; Bruno, R.; Hall, W.; Eaglesham, G.; Mueller, J. F. Refining the estimation of illicit drug consumptions from wastewater analysis: co-analysis of prescription pharmaceuticals and uncertainty assessment. *Water Res* **2011,** *45* (15), 4437-48.

8. Castiglioni, S.; Bijlsma, L.; Covaci, A.; Emke, E.; Hernandez, F.; Reid, M.; Ort, C.; Thomas, K. V.; van Nuijs, A. L. N.; de Voogt, P.; Zuccato, E. Evaluation of Uncertainties Associated with the Determination of Community Drug Use through the Measurement of Sewage Drug Biomarkers. *Environ. Sci. Technol.* **2013,** *47* (3), 1452-1460.

9. van Nuijs, A. L. N.; Lai, F. Y.; Been, F.; Andres-Costa, M. J.; Barron, L.; Baz-Lomba, J. A.; Berset, J. D.; Benaglia, L.; Bijlsma, L.; Burgard, D.; Castiglioni, S.; Christophoridis, C.; Covaci, A.; de Voogt, P.; Emke, E.; Fatta-Kassinos, D.; Fick, J.; Hernandez, F.; Gerber, C.; Gonzalez-Marino, I.; Grabic, R.; Gunnar, T.; Kannan, K.; Karolak, S.; Kasprzyk-Hordern, B.; Kokot, Z.; Krizman-Matasic, I.; Li, A.; Li, X. Q.; Love, A. S. C.; de Alda, M. L.; McCall, A. K.; Meyer, M. R.; Oberacher, H.; O'Brien, J.; Quintana, J. B.; Reid, M.; Schneider, S.; Simoes, S. S.; Thomaidis, N. S.; Thomas, K.; Yargeau, V.; Ort, C. Multi-year inter-laboratory exercises for the analysis of illicit drugs and metabolites in wastewater: Development of a quality control system. *Trac-Trend Anal Chem* **2018,** *103*, 34-43.

10. Li, J.; Gao, J.; Thai, P. K.; Sun, X.; Mueller, J. F.; Yuan, Z.; Jiang, G. Stability of Illicit Drugs as Biomarkers in Sewers: From Lab to Reality. *Environ. Sci. Technol.* **2018,** *52* (3), 1561-1570.

11. Gao, J.; Banks, A.; Li, J.; Jiang, G.; Lai, F. Y.; Mueller, J. F.; Thai, P. K. Evaluation of in-sewer transformation of selected illicit drugs and pharmaceutical biomarkers. *Sci. Total Environ.* **2017,** *609*, 1172-1181.

566 12. Thai, P. K.; Jiang, G.; Gernjak, W.; Yuan, Z.; Lai, F. Y.; Mueller, J. F. Effects of
567 sewer conditions on the degradation of selected illicit drug residues in wastewater.
568 *Water Res.* **2014,** *48*, 538-47.
569 13. McCall, A. K.; Bade, R.; Kinyua, J.; Lai, F. Y.; Thai, P. K.; Covaci, A.; Bijlsma, L.; van
570 Nuijs, A. L.; Ort, C. Critical review on the stability of illicit drugs in sewers and
571 wastewater samples. *Water Res.* **2016,** *88*, 933-47.
572 14. McCall, A. K.; Scheidegger, A.; Madry, M. M.; Steuer, A. E.; Weissbrodt, D. G.;
573 Vanrolleghem, P. A.; Kraemer, T.; Morgenroth, E.; Ort, C. Influence of Different Sewer
574 Biofilms on Transformation Rates of Drugs. *Environ. Sci. Technol.* **2016,** *50* (24),
575 13351-13360.
576 15. McCall, A. K.; Palmitessa, R.; Blumensaat, F.; Morgenroth, E.; Ort, C. Modeling
577 in-sewer transformations at catchment scale – Implications on drug consumption
578 estimates in wastewater-based epidemiology. *Water Res.* **2017,** *122*, 655-668.
579 16. Ramin, P.; Brock, A. L.; Causanilles, A.; Valverde Perez, B.; Emke, E.; de Voogt, P.;
580 Polesel, F.; Plosz, B. G. Transformation and sorption of illicit drug biomarkers in sewer
581 biofilms. *Environ. Sci. Technol.* **2017,** *51* (18), 10572–10584.
582 17. Gao, J.; Li, J.; Jiang, G.; Shypanski, A. H.; Nieradzik, L. M.; Yuan, Z.; Mueller, J. F.;
583 Ort, C.; Thai, P. K. Systematic evaluation of biomarker stability in pilot scale sewer
584 pipes. *Water Res* **2018,** *151*, 447-455.
585 18. Hvitved-Jacobsen, T.; Vollertsen, J.; Nielsen, A. H. *Sewer Processes: Microbial*
586 *and Chemical Process Engineering of Sewer Networks*. Second ed.; CRC Press: Boca
587 Raton, 2013.

588 19. Plosz, B. G.; Reid, M. J.; Borup, M.; Langford, K. H.; Thomas, K. V.
589 Biotransformation kinetics and sorption of cocaine and its metabolites and the
590 factors influencing their estimation in wastewater. *Water Res.* **2013,** *47* (7), 2129-
591 2140.
592 20. Ostman, M.; Fick, J.; Nasstrom, E.; Lindberg, R. H. A snapshot of illicit drug use
593 in Sweden acquired through sewage water analysis. *Sci. Total Environ.* **2014,** *472*,
594 862-871.
595 21. Ramin, P.; Libonati Brock, A.; Polesel, F.; Causanilles, A.; Emke, E.; de Voogt, P.;
596 Plosz, B. G. Transformation and sorption of illicit drug biomarkers in sewer systems:
597 understanding the role of suspended solids in raw wastewater. *Environ. Sci. Technol.*
598 **2016,** *50* (24), 13397-13408.
599 22. Castiglioni, S.; Zuccato, E.; Crisci, E.; Chiabrando, C.; Fanelli, R.; Bagnati, R.
600 Identification and measurement of illicit drugs and their metabolites in urban
601 wastewater by liquid chromatography-tandem mass spectrometry. *Anal Chem* **2006,**
602 *78* (24), 8421-8429.
603 23. Baker, D. R.; Kasprzyk-Hordern, B. Critical evaluation of methodology
604 commonly used in sample collection, storage and preparation for the analysis of
605 pharmaceuticals and illicit drugs in surface water and wastewater by solid phase
606 extraction and liquid chromatography-mass spectrometry. *J Chromatogr A.* **2011,**
607 *1218* (44), 8036-59.

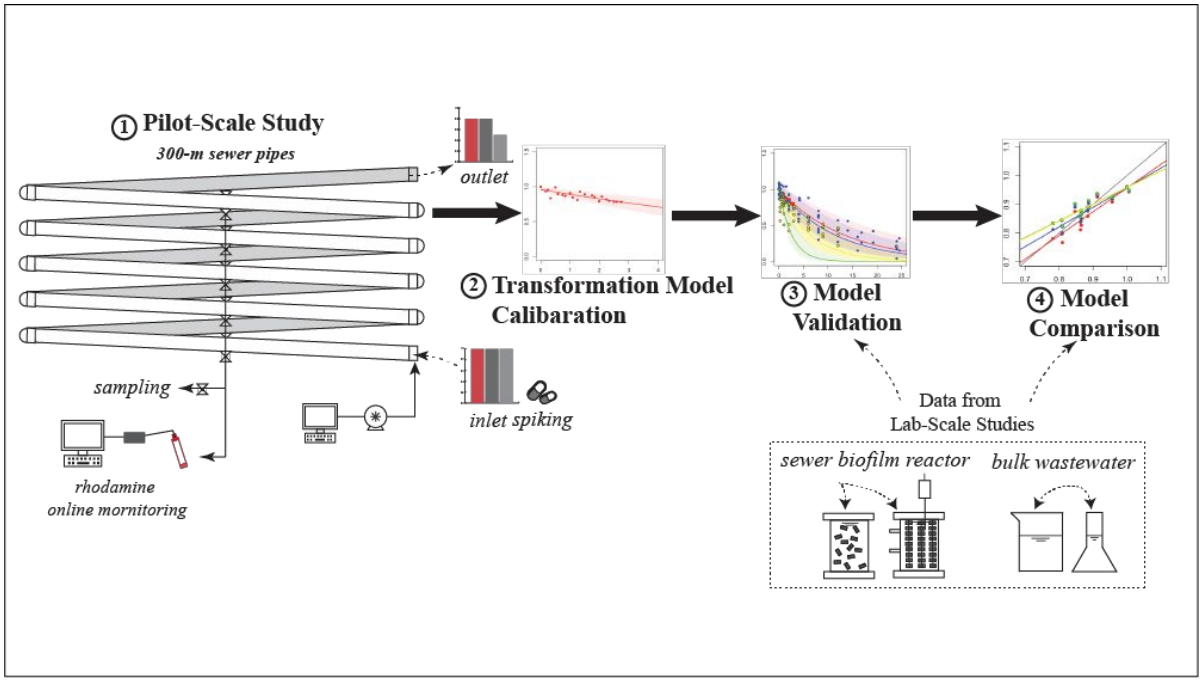
608 24. Chen, C.; Kostakis, C.; Irvine, R. J.; Felgate, P. D.; White, J. M. Evaluation of pre-
609 analysis loss of dependent drugs in wastewater: stability and binding assessments.
610 *Drug Test Anal* **2013,** *5* (8), 716-21.
611 25. Senta, I.; Krizman, I.; Ahel, M.; Terzic, S. Assessment of stability of drug
612 biomarkers in municipal wastewater as a factor influencing the estimation of drug
613 consumption using sewage epidemiology. *Sci Total Environ* **2014,** *487*, 659-65.
614 26. Jelic, A.; Rodriguez-Mozaz, S.; Barcelo, D.; Gutierrez, O. Impact of in-sewer
615 transformation on 43 pharmaceuticals in a pressurized sewer under anaerobic
616 conditions. *Water Res.* **2015,** *68*, 98-108.
617 27. Gao, J.; Li, J.; Jiang, G.; Yuan, Z.; Eaglesham, G.; Covaci, A.; Mueller, J. F.; Thai, P.
618 K. Stability of alcohol and tobacco consumption biomarkers in a real rising main
619 sewer. *Water Res.* **2018,** *138*, 19-26.
620 28. Ian A. Watson; Sascha E.Oswald; Steven A. Banwart; Roger S. Crouch;
621 Thornton, S. F. Modeling the dynamics of fermentation and respiratory processes in a
622 groundwater plume of phenolic contam- inants interpreted from laboratory- to field-
623 scale. *Environ. Sci. Technol.* **2005,** *39* (22), 8829-8839.
624 29. Shypanski, A. H.; Yuan, Z.; Sharma, K. Influence of pressure main pumping
625 frequency on sulfide formation rates in sanitary sewers. *Environmental Science: Water*
626 *Research & Technology* **2018,** *4* (3), 403-410.
627 30. Ramin, P.; Polesel, F.; Brock, A. L.; Plósz, B. G. The impact of temperature on the
628 transformation of illicit drug biomarkers in wastewater. *Sci. Total Environ.* **2018,** *644*,
629 1612-1616.

630 31. Senta, I.; Krizman, I.; Ahel, M.; Terzic, S. Assessment of stability of drug
631 biomarkers in municipal wastewater, as a factor influencing the estimation of drug
632 consumption using sewage epidemiology. *Sci. Total Environ.* **2014,** *487*, 659-665.
633 32. Devault, D. A.; Levi, Y.; Karolak, S. Applying sewage epidemiology approach to
634 estimate illicit drug consumption in a tropical context: Bias related to sewage
635 temperature and pH. *Sci. Total Environ.* **2017,** *584*, 252-258.
636 33. Baker, D. R.; Ocenaskova, V.; Kvicalova, M.; Kasprzyk-Hordern, B. Drugs of
637 abuse in wastewater and suspended particulate matter--further developments in
638 sewage epidemiology. *Environ Int* **2012,** *48*, 28-38.
639 34. Gutierrez, O.; Park, D.; Sharma, K. R.; Yuan, Z. Effects of long-term pH elevation
640 on the sulfate-reducing and methanogenic activities of anaerobic sewer biofilms.
641 *Water Res.* **2009,** *43* (9), 2549-2557.
642 35. Guisasola, A.; Sharma, K. R.; Keller, J.; Yuan, Z. Development of a model for
643 assessing methane formation in rising main sewers. *Water Res.* **2009,** *43* (11), 2874-
644 2884.
645 36. Liu, Y. W.; Tugtas, A. E.; Sharma, K. R.; Ni, B. J.; Yuan, Z. G. Sulfide and methane
646 production in sewer sediments: Field survey and model evaluation. *Water Res.* **2016,**
647 *89*, 142-150.
648 37. Liu, Y.; Chen, W.; Arendt, P.; Huang, H.-Z. Toward a Better Understanding of
649 Model Validation Metrics. *J Mech Design* **2011,** *133* (7).
650 38. Li, J. Assessing the accuracy of predictive models for numerical data: Not r nor
651 r2, why not? Then what? *Plos One* **2017,** *12* (8), e0183250.

652 39. Æsøy, A.; Storfjell, M.; Mellgren, L.; Helness, H.; Thorvaldsen, G.; Ødegaard, H.;
653 Bentzen, G. A comparison of biofilm growth and water quality changes in sewers with
654 anoxic and anaerobic (septic) conditions. *Water Sci. Technol.* **1997,** *36* (1), 303-310.
655 40. Sharma, K.; Ganigue, R.; Yuan, Z. pH dynamics in sewers and its modeling.
656 *Water Res.* **2013,** *47* (16), 6086-6096.
657 41. Jiang, G.; Sharma, K. R.; Guisasola, A.; Keller, J.; Yuan, Z. Sulfur transformation
658 in rising main sewers receiving nitrate dosage. *Water Res.* **2009,** *43* (17), 4430-4440.
659 42. Thai, P. K.; Jiang, G. M.; Gernjak, W.; Yuan, Z. G.; Lai, F. Y.; Mueller, J. F. Effects of
660 sewer conditions on the degradation of selected illicit drug residues in wastewater.
661 *Water Res.* **2014,** *48*, 538-547.
662 43. Sharma, K. R.; Yuan, Z.; de Haas, D.; Hamilton, G.; Corrie, S.; Keller, J. Dynamics
663 and dynamic modelling of H(2)S production in sewer systems. *Water Res.* **2008,** *42*
664 (10), 2527-2538.
665 44. Foley, J.; Yuan, Z.; Keller, J.; Senante, E.; Chandran, K.; Willis, J.; Shah, A.; van
666 Loosdrecht, M.; van Voorthuizen, E. *N2O and CH4 emission from wastewater*
667 *collection and treatment systems: technical report*; Global Water Research Coalition:
668 London, United Kingdom, 2011.
669 45. Guisasola, A.; de Haas, D.; Keller, J.; Yuan, Z. Methane formation in sewer
670 systems. *Water Res.* **2008,** *42* (6), 1421-1430.
671 46. Liu, Y. W.; Sharma, K. R.; Fluggen, M.; O'Halloran, K.; Murthy, S.; Yuan, Z. G.
672 Online dissolved methane and total dissolved sulfide measurement in sewers. *Water*
673 *Res.* **2015,** *68*, 109-118.

674 47. Wang, K. J.; Zeeman, G.; Lettinga, G. Alteration in Sewage Characteristics Upon
675 Aging. *Water Sci. Technol.* **1995,** *31* (7), 191-200.
676 48. Ramin, P.; Valverde-Perez, B.; Polesel, F.; Locatelli, L.; Plosz, B. G. A systematic
677 model identification method for chemical transformation pathways - the case of
678 heroin biomarkers in wastewater. *Sci Rep* **2017,** *7* (1), 9390.

679



680

681 Table of Contents

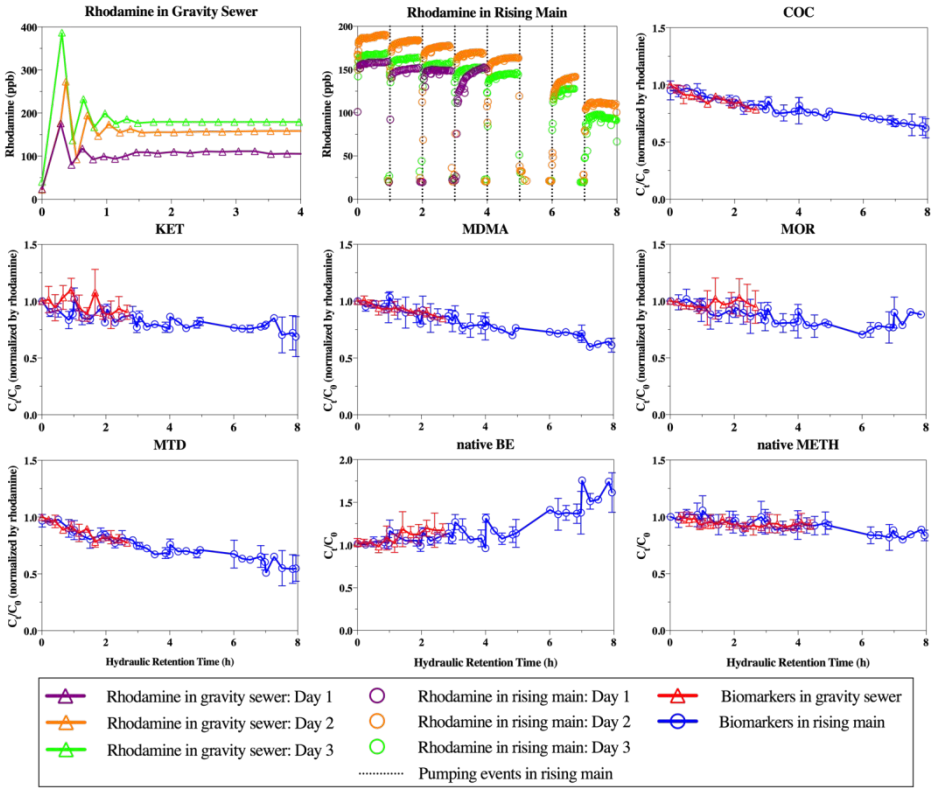
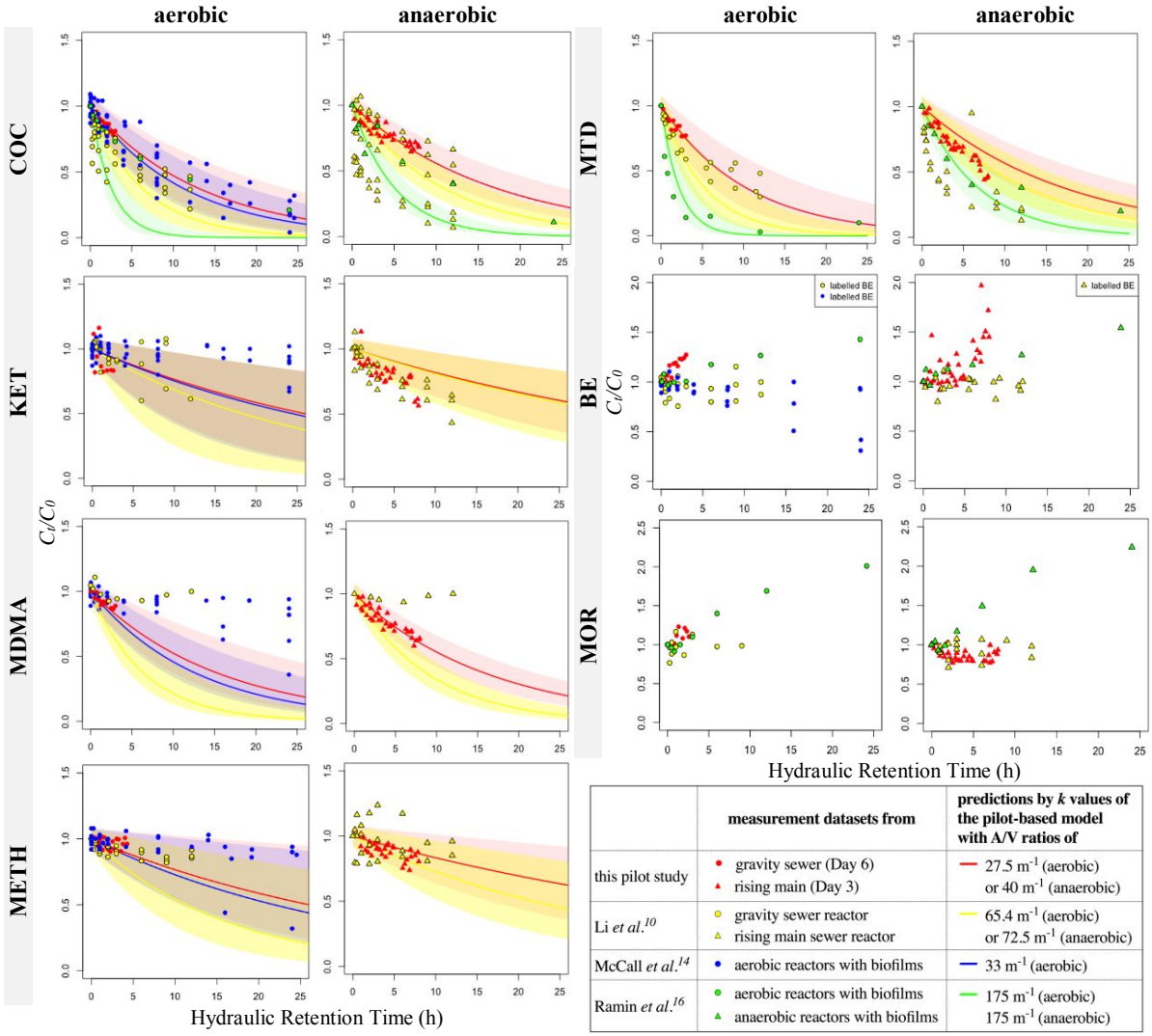


Figure 1. Transportation of rhodamine as a flow tracer and stability of biomarkers in the pilot rising main pipe and the pilot gravity sewer pipe. (Transformations of MOR and MTD in the pilot sewer pipes were also reported recently[17])



	measurement datasets from	predictions by *k* values of the pilot-based model with A/V ratios of
this pilot study	gravity sewer (Day 6) rising main (Day 3)	27.5 m^{-1} (aerobic) or 40 m^{-1} (anaerobic)
Li *et al.*[10]	gravity sewer reactor rising main sewer reactor	65.4 m^{-1} (aerobic) or 72.5 m^{-1} (anaerobic)
McCall *et al.*[14]	aerobic reactors with biofilms	33 m^{-1} (aerobic)
Ramin *et al.*[16]	aerobic reactors with biofilms anaerobic reactors with biofilms	175 m^{-1} (aerobic) 175 m^{-1} (anaerobic)

Figure 2. Validation of the pilot-based transformation coefficients through the predictive scenarios of the corresponding experimental observations in different stability studies.

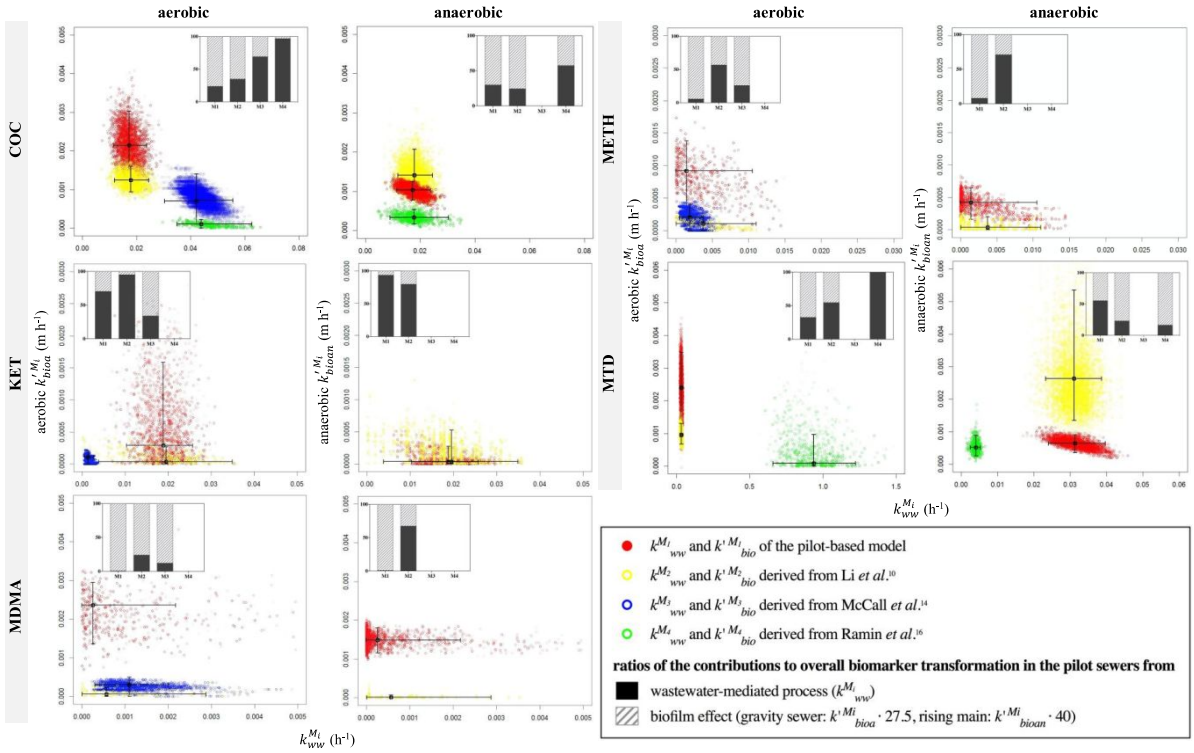


Figure 3. Joint distributions of the transformation coefficients ($k^{M_i}_{ww}$ and $k'^{M_i}_{bio}$) derived from the four stability studies (M1~M4) under aerobic and anaerobic conditions. The embedded histograms indicate the contributions from wastewater processes against biofilm effects to the overall biomarker transformation in the pilot sewer pipes.

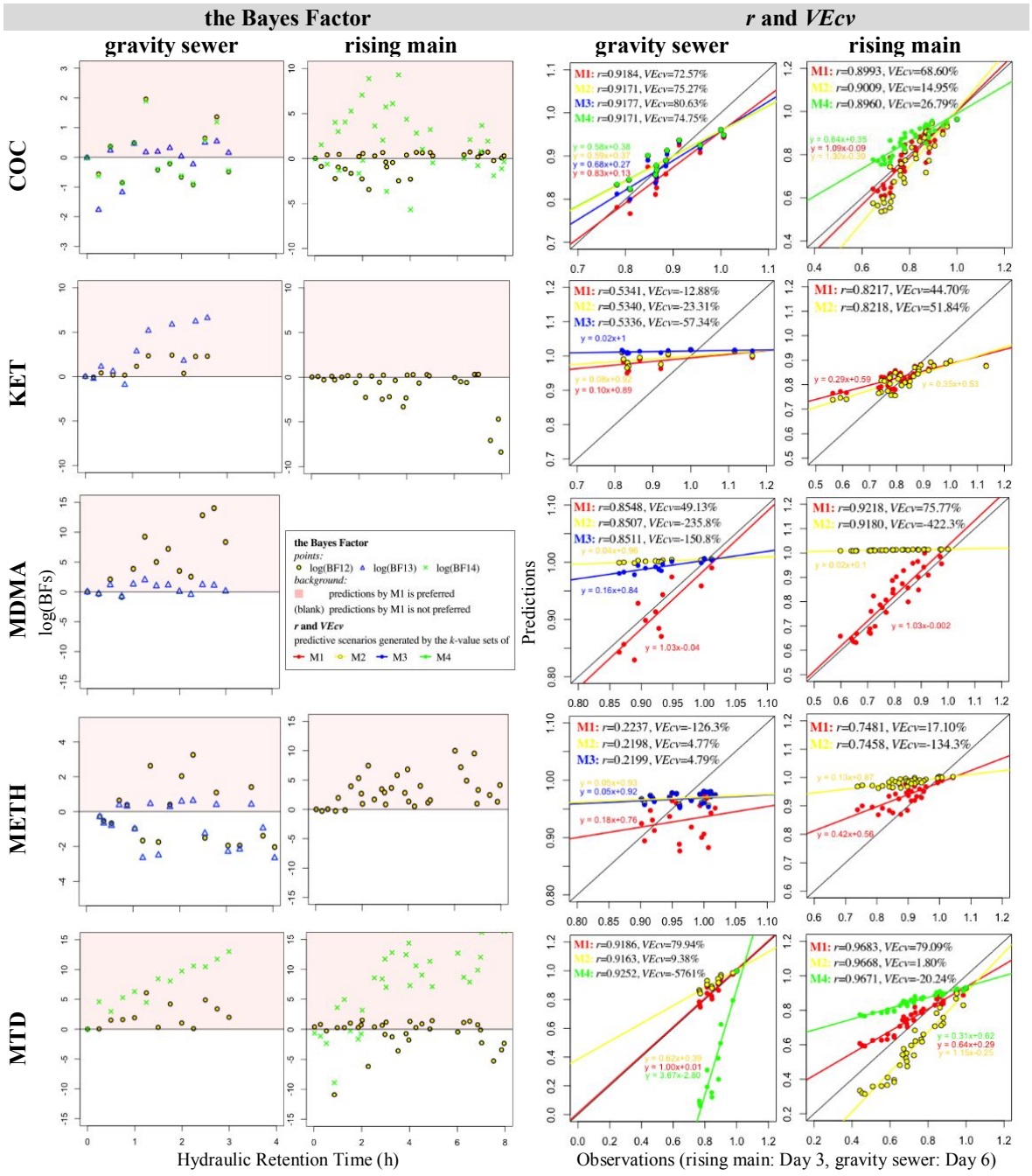


Figure 4. Model performance comparison: results of the Bayes factor, *r*, and *VEcv* for comparing the predictive scenarios generated by the *k*-value sets of four stability studies (M1 to M4) under the conditions of the pilot sewer pipes.