Factors influencing the stormwater quality model of sewer networks and a case study of Louis Fargue urban catchment in Bordeaux, France

Congcong Sun [1,*], Benjamí Parellada [1], Jing Feng [2], Vicenç Puig [1] and Gabriela Cembrano [1,3]

[1] Advanced Control Systems Group at the Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Llorens i Artigas, 4-6, 08028 Barcelona, Spain;
[2] Suez Research Center, Nanjing West Road 800, Shanghai 40001, China;
[3] CETaqua, Water Technology Centre, Barcelona, 08904, Spain;
* Correspondence: csun@iri.upc.edu.

Benjami Parellada <benjamiparellada@gmail.com>;
FENG Jing <fengjing385@foxmail.com>;
Vicenc Puig Cayuela <vicenc.puig@upc.edu>;
Gabriela Cembrano <gabriela.cembrano@upc.edu>.

Abstract: Pollution caused by combined sewer overflows has become a global threat to the environment. Under this challenge, quality-based real-time control (RTC) is considered as an effective approach to minimize pollution through generating optimal operation strategies for the sewer infrastructure. To suit the fast computation requirement of RTC implementation, simplified quality models are required. However, due to the hydrological complexity, it is not easy to develop simplified quality models which are amenable to be used in real-time computations. Under this context, this paper contributes 1) a preliminary analysis of influencing factors for the quality models of sewer networks in order to give supportive knowledge for both model developing and applying. Conceptual quality models which were proposed previously by the authors, with total suspended solid (TSS) as quality indicator, are used in this study. 2) Clustering algorithm is used for exploratory analysis. Further analysis about the correlations between different factors and model performance is also carried out. 3) All the study and analysis are demonstrated on a real pilot based on Louis Fargue urban catchment in Bordeaux. Conclusive results about the influencing factors, flow rate, rain intensity and pipe length, as well as their correlations with the TSS models are elaborated.

Keywords: analysis, influencing factor, quality model, real-time control, total suspended solid

---

INTRODUCTION

In nowadays urban cities, combined sewer networks (SN) collect stormwater together with the wastewater and then send them to the wastewater treatment plants (WWTP) for purifying before releasing them into the receiving bodies (García *et al.* 2015). In storm weather, when the capacities of SN and the WWTP are overloaded, combined sewer overflow (CSO) occurs and generates harmful pollution to the environment (Butler & Schütze 2005; Becouze *et al.* 2009; Gasperi *et al.* 2008; Joseph-Duran *et al.* 2014). In order to provide proper service to the urban life, as well as to protect the water ecosystem, efficient management of SN through advanced control techniques is required.

In view of the recent advances in information and communications technology, real-time control (RTC) has been considered as an effective solution for the SN management due to the obvious advantages comparing with the traditional solutions of constructing infrastructure (Beeneken *et al.* 2013; Cembrano *et al.* 2004; Döring 1989; Dirckx *et al.* 2011; EPA 2006; Fanlin *et al.* 2017; Joseph-Duran *et al.* 2015; Puig *et al.* 2009; Schütze *et al.* 2004). Moreover, the integrated pollution-based RTC can generate system-wide effective strategies to reduce CSO volume, as well as the released pollution, through integrating the SN and WWTP considering both the quantity and quality measurements

(Fanlin *et al.* 2017; Sun *et al.* in press). This is also the main goal of the European project LIFE EFFIDRAIN (Efficient Integrated Real-time Control in Urban Drainage and Wastewater Treatment Plants for Environmental Protection) [Joseph-Duran *et al.* (2014, 2015); Ocampo-Martínez *et al.* 2013; Puig *et al.* 2009; Pleau *et al.* 2010; Schütze *et al.* 2004; Sun *et al.* (2017a, 2018a, b)].

There are plenty of physical models to describe quality dynamics in SN in detail, like the KUL Model (Combes 1982), Velikanov Model (van Rijn 1984) and Ackers White Model (Ackers & White 1973). Moreover, simulation platforms which elaborate the hydrological dynamics through embedding quality equations are also available, the popular ones are SWMM (Huber 1988; Rossman 2015), WaterCress (Clark 2002), Simba# (IFAK 2005), MIKE Urban (DHI 2007) and InfoWorks CS (MWH 2010). These high-fidelity models have sufficient capacities of representing quality dynamics accurately but need large computational effort and hard to be used into the RTC optimization. In order to take quality and its dynamics into RTC of urban drainage systems, simplified conceptual quality models must be used (Cembrano *et al.* 2004; Sun *et al.* in press). Taking into account the hydrological complexities of SN and the high data requirements for the model calibration, it is still a challenge to develop and apply conceptual quality models successfully.

To reduce the limitations of high-fidelity models and tools, new quality modelling approaches are proposed focused on total suspended solid (TSS) (see Vezzaro *et al.* 2014), Chemical Oxygen Demand (COD) and ammonia (see Fu *et al.* 2010; Lacour & Schütze 2011). Besides, Vezzaro *et al.* (2020) has also proposed just recently an online forecasting model and the performance evaluation of ammonium concentrations at the WWTP inlets. These modelling approaches were used to minimize the pollution loads from SN and WWTP inlets through including water quality into dynamic cost functions or prioritizing the CSO loads considering the recipient status (Fu *et al.* 2010; IFAK 2005; Lacour & Schütze 2011; Torres-Matallana *et al.* 2018; Vezzaro *et al.* 2014). However, these quality models cannot be mapped inside the SN and cannot be used into the optimization process directly, either. Afterwards, (Sun *et al.* 2017b) proposed conceptual quality models for the SN, with TSS as the quality indicator, which demonstrate how the TSS can be mapped inside the sewers and be involved into the control process directly. However, the development and application of the quality models are greatly affected by the topology of the SN, the climate and the physical characteristics of the pilot. What are the influencing factors and how they correlated with performance of the quality models are still unknown and are important to be investigated.

Under this context, this paper contributes: 1) a preliminary analysis of the influencing factors for quality models in order to provide supportive guidelines for conceptual quality modelling and their further application at different pilots. In particular, the simplified TSS models proposed previously by the authors are used for analysis. Three possible factors rain intensity, sewer length, and flow rate are supposed very likely to affect the model performance. 2) To benefit from the new insight brought by the data processing techniques, clustering algorithm is used as the tool for grouping model performances and exploratory analyzing the influencing factors in this study. Afterwards, further analysis considering correlations between different influencing factors and the model performance is also carried out. The hydraulic and hydrology datasets for model calibration and validation, as well as the afterwards analysis, is produced through virtual reality simulations by EPA-SWMM5. 3) A real pilot based on the Louis Fargue urban catchment in Bordeaux (France) is used as case study. Conclusive results about the influencing factors (flow rate, rain intensity and pipe length) as well as their correlations with performance of the TSS models are provided afterwards.

MATERIALS AND METHODS

Conceptual TSS models

Physically, the TSS dynamics inside a sewer include solids transportation, sedimentation and erosions, which are easily affected by the hydraulic parameters. Three different conceptual models have been proposed in (Sun *et al.* 2017b). These simplified TSS models are analysed in this paper, whose model equations are briefly presented as:

Model 1

$$\boldsymbol{ss}_{out}(k+1) = (1-\alpha)\boldsymbol{ss}_{out}(k) + \alpha\boldsymbol{ss}_{in}(k) \tag{1}$$

Model 2

$$\boldsymbol{ss}_{out}(k+1) = \alpha_1\boldsymbol{ss}_{out}(k) + \alpha_2\boldsymbol{ss}_{in}(k) \tag{2}$$

Model 3

$$\boldsymbol{ss}_{out}(k+1) = \beta\boldsymbol{ss}_{in}(k-d) + e \tag{3}$$

where $\boldsymbol{ss}_{out} \in \mathbb{R}^{n_s}$ and $\boldsymbol{ss}_{in} \in \mathbb{R}^{n_s}$ represent the output and input vectors of TSS concentrations [mg/l] for the sewers, respectively; $k \in \mathbb{N}^+$ is the current time step; $d \in \mathbb{N}^+$ represents the TSS transportation delay; $\alpha, a_1, \alpha_2, \beta, d, e$ are the model coefficients which need to be calibrated for each sewer under different rain scenarios while n_s is the maximal number of sewers.

In order to analyse these simplified TSS models, the flow rate, rain intensity and sewer length are considered as the three most relevant hydraulic factors which are likely to affect the model performance.

Analysis tools

Before performance analysing, the conceptual TSS models presented in the previous section are firstly calibrated and validated using datasets generated through virtual reality simulations under different scenarios. The hydraulic datasets are produced by the simulator EPA-SWMM5. The quality data is generated through the library SWMM-TSS (Maruéjouls *et al.* 2012) from LyRE (R+D centre of Suez) based on extended Barre de Saint Venant equation set in SWMM5, which can reproduce the TSS transport, sediment accumulation and erosion in sewers and retention tanks (Wiuff 1985).

Figure 1 illustrates modifications made by SWMM-TSS in the SWMM5 library model, where the boxes correspond to the existing modules in SWMM5 while grey boxes are the ones added for representing the quality module. The abbreviations WW and DW represent the wet and dry weather respectively.

Besides the SWMM5 simulator, MATLAB and GAMS optimization software (Richard 2016) are also used in the calibration process. RStudio, which is an open source software for R (Krotov 2017), is used for grouping and analysing the model performances through clustering algorithm.

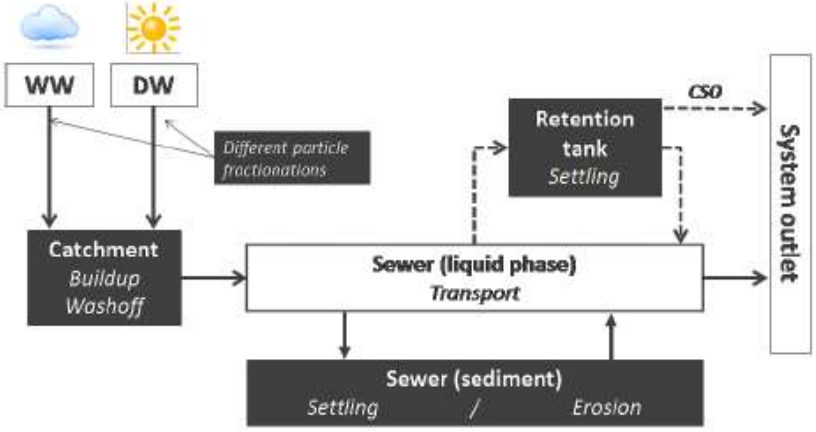

Figure 1. Structure of Quality Module SWMM-TSS

The Nash Sutcliffe model efficiency coefficient (NSE) (Nash & Sutcliffe 1970) is used to describe calibration performance of the TSS models:

$$fit_{i,j}^{m} = 1 - \sum_{k=1}^{K}\left(ss_{out}^{i,j,m}(k) - \widehat{ss}_{out}^{i,j,m}(k)\right)^2 / \sum_{k=1}^{K}\left(\widehat{ss}_{out}^{i,j,m}(k) - \overline{\widehat{ss}_{out}^{i,j,m}}\right)^2 \tag{4}$$

where $ss_{out}^{i,j,m}$ is the predicted TSS value using the model m ($m \in [1,2,3]$) for the sewer i under the rain scenario j, $\widehat{ss}_{out}^{i,j,m}$ is the corresponding TSS value read from the simulator and $\overline{\widehat{ss}_{out}^{i,j,m}}$ is average of it, $K \in \mathbb{N}^+$ represents maximal number of time steps in one rain scenario. The fitting performance $fit \in \mathbb{R}^{n_s * n_r * 3}$ can range from $-\infty$ to 1 (n_r represents maximal number of rain scenarios used for analysis). The higher it is; the better performance the model has.

Data structure

After the calibration process, enough data are generated to structure the matrix $D \in \mathbb{R}^{(n_s*n_r*3)}$ to be used for the performance analyzing. The matrix D includes fitting performance vector $\boldsymbol{fit} \in \mathbb{R}^{n_s*n_r*3}$, the corresponding flow rate vector $\boldsymbol{flow} \in \mathbb{R}^{n_s}$ [l/s], sewer lengths $\boldsymbol{len} \in \mathbb{R}^{n_s}$ [m], as well as the rain intensities $\boldsymbol{rain} \in \mathbb{R}^{n_r}$ [mm]. The format of matrix D is presented as:

$$D = \begin{bmatrix} fit_{1,1}^{1} & flow_1 & rain_1 & len_1 \\ fit_{2,1}^{1} & flow_2 & rain_1 & len_2 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ fit_{n_s,n_r}^{3} & flow_{n_s} & rain_{n_r} & len_{n_s} \end{bmatrix}$$

Clustering algorithm

A clustering algorithm (Rodriguez & Laio 2014; Soldevila *et al.* 2016) is used for the exploratory analysis of model performances based on the matrix D, which includes all the influencing features to be analyzed. Considering the given matrix, clustering algorithm can return a list of c clusters $C = \{c_1, \dots, c_c\}$. To achieve this, firstly, each data point is assigned a measurement grade $G = g_{i,j} \in [0,1], i = 1, \dots, (n_s * n_r * 3), j = 1, \dots, c,$ to indicate its possibility belongs to each of the cluster. And after that, the clustering process is proceeded through minimizing the following objective function:

$$\arg\min_{C} \sum_{i=1}^{n_s*n_r*3} \sum_{j=1}^{c} g_{i,j}{}^{\propto} \left\| x_i - c_j \right\|^2 \qquad (5)$$

where $\boldsymbol{x} \in \mathbb{R}^{n_s*n_r*3}$ is the data point being analysed; $\propto$ is a fuzzy number to determine level of cluster fuzziness. To have a crisp partitioning in this case, $\propto$ is put 1, and at that time, the measurement grade is converge to 0 or 1.

In this study, the three proposed TSS models are analysed firstly to conclude which model works best through clustering their fitting performance $\boldsymbol{fit}$. Under this objective, the number of returned clusters c is set as 3; the data point vector being analysed $\boldsymbol{x}$ is set as $\boldsymbol{fit}$; m is 1; and the measurement grade $g_{i,j}^{m}$ for the *m-th* cluster under the *j-th* rain scenario is computed for all the sewers $\forall i \in 1, \dots, n_s$ using the following equation:

$$g_{i,j}^{m} = \begin{cases} 0, & fit_{i,j}^{m} < \max\limits_{m \in C} fit_{i,j}^{m} \\ 1, & fit_{i,j}^{m} = \max\limits_{m \in C} fit_{i,j}^{m} \end{cases} \qquad (6)$$

The cluster which has maximal data points is concluded as the best performance model.

CASE STUDY

Louis Fargue Catchment

The Louis Fargue catchment is the most populated catchment in the Bordeaux Metropole, which has an area of around 7,700 ha and accommodate about 300,000 inhabitants. Figure 2 is a layout of this network in SWMM. The total sewer network length of Louis Fargue catchment is 1,340km with 80% combined while 20% separated. Most of these pipes have a circular shape with diameters ranging from 300 mm to 4500 mm. In the dry weather, the wastewater is collected and transferred all the way to the sub-catchment outlets (connections represented by red triangles in Figure 2) where it is then carried to the WWTP through a large interceptor. At each connection point (an outlet), there is a combined sewer overflow structure to discharge excessive combined water into the Garonne river during storm weather. All these descriptions are referred from (Ly 2019), where reader is referred to obtain more details about this catchment.

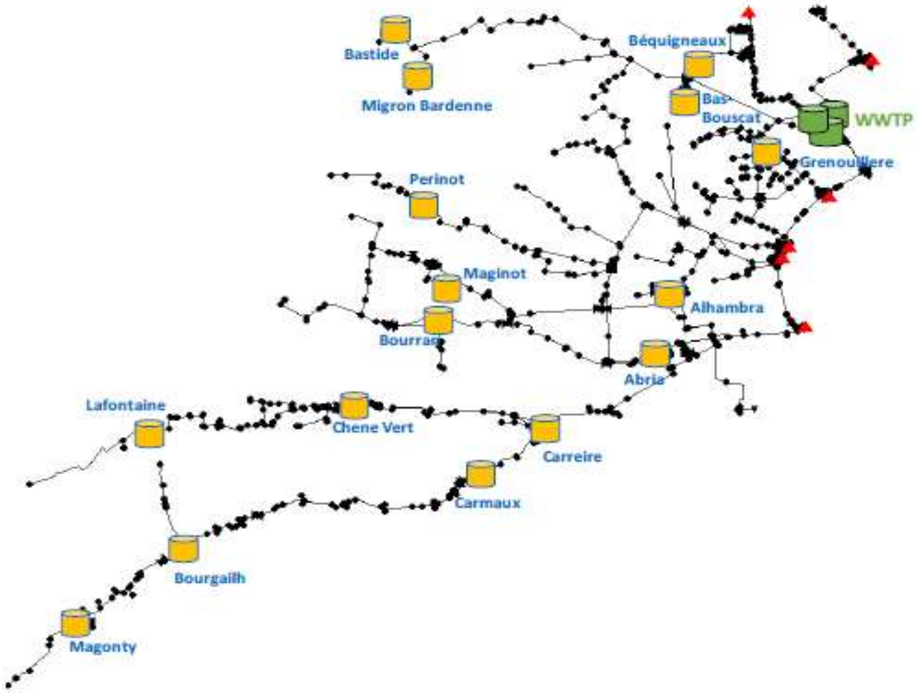

Figure 2. Louis Fargue Catchment in SWMM referred in (Ly 2019)

Rain scenarios for calibration and validation come from the real rainfall measured at Bordeaux in the year of 2016. The 129 hour raining data since 01/30/2016 is used for calibration, which corresponds to 14 different rain events according to their sub-catchment districts (Perinot, LaFontaine, Carreire, etc.).

Exploratory Analysis

In order to focus on analysing the sewer model, all the conduits in Louis Fargue catchment are recompiled. The conduits with one entrance and one exit are filtered for analysing, and there are in total 122 conduits. The corresponding rainfall measurements is assigned to each conduit by checking the sub-catchment it belongs to. Virtual reality simulation based on SWMM and SWMM-TSS is applied in a certain time period of 129 hours to generate hydraulic and hydrology datasets for the following calibration processes. Afterwards, all of these 122 sewers are calibrated for the three models through MATLAB and GAMS codes. Calibration and model performance results are obtained for the 122 sewers with the three models. All the sequential values from each of the simulations is then used to construct the Matrix D, where a total of 5124 feature vectors are analysed for the 122 different conduits under 14 different rain events with 3 TSS models.

The clustering algorithm is applied firstly to generate the measurement grade $g_{i,j}^{m}$ to check for each conduit which of the three models is the best. The results are presented in Table 1 showing the times each model presents best performance for the 122 conduits.

Table 1 Exploratory Analysis Results

	Model 1	**Model 2**	**Model 3**
grade	13	1534	161

Table 1 shows that Model 1 is not working as well as the other two, whose measurement grade is 13, the minimal one among the three models. Only few conduits are found with performance evaluation higher than both of the Model 2 and Model 3. However, even though for the 13 scenarios when the Model 1 works better than the others two, it does not mean it is a better model. Table 2 shows in detail for the scenarios when Model 1 works better, which indicates that the fitting performances of Model 1 do not behave so well, either. Furthermore, it can also be seen that it is always the same conduits which provoke this deterioration in the other two models (all the names of the conduits are created due to confidential requirements of the pilot).

Table 2 Scenarios work best for Model 1

Scenarios	fit^1	fit^2	fit^3	From conduit	To conduit
S_1^1	0.000	-0.038	-0.084	Link_P_95_1	Link_BV_1
S_2^1	0.024	-0.114	-0.260	Link_P_95_1	Link_BV_1
S_3^1	0.359	0.027	-0.180	Link_EN_1	Link_905_1
S_4^1	0.000	-0.211	-11.10	Link_P_95_1	Link_BV_1
S_5^1	0.000	-2.940	-1.781	Link_G14_1	Link_G19_1
S_6^1	0.000	-0.085	-4.102	Link_P_95_1	Link_BV_1
S_7^1	0.074	0.055	0.036	Link_EN_1	Link_905_1
S_8^1	0.000	-0.066	-21.937	Link_P_95_1	Link_BV_1
S_9^1	0.000	-9.651	-5.681	Link_255_1	Link_260_1
S_{10}^1	0.000	-0.220	-4.657	Link_P_95_1	Link_BV_1
S_{11}^1	0.002	-3.713	-0.642	Link_BO_1	Link_AC_1
S_{12}^1	0.082	-1.176	-0.640	Link_G14_1	Link_G19_1
S_{13}^1	0.000	-0.004	-7.059	Link_P_95_1	Link_BV_1

Both Model 3 and Model 2 are good in the average case with higher measurement grades. The mean fitting value for Model 1 is 0.7787036 and for model 2 is 0.8009895. When comparing Model 2 and Model 3, it is clear that Model 2 performs better. Moreover, for Model 3, some poor results (negative fitting performances) have been obtained as illustrated in Table 3:

Table 3 Example scenarios work worst for Model 3

Scenarios	fit^1	fit^2	fit^3	From conduit	To conduit
S_1^3	0.000	0.191	-1.833	Link_OR1_1	Link_OR2_2
S_2^3	0.000	0.441	-8.106	Link_OR1_1	Link_OR2_2
S_3^3	0.000	0.594	-6.521	Link_OR1_1	Link_OR2_2
S_4^3	0.000	0.534	-6.405	Link_OR1_1	Link_OR2_2
S_5^3	0.002	0.003	-2.526	Link_OR1_1	Link_OR2_2
S_6^3	0.000	0.016	-5.399	Link_OR1_1	Link_OR2_2
S_7^3	0.000	0.016	-2.215	Link_OR1_1	Link_OR2_2
S_8^3	0.000	-0.066	-2.194	Link_P_95_1	Link_BV_1
S_9^3	0.000	0.768	-2.098	Link_AA_3	Link_PG_1

From Table 3, it is clear that the conduit from Link_OR1_1 to Link_OR2_2 is not represented particularly well. Among these first results, it can be concluded that Model 2, in most scenarios, works better with an average fitting performance of 89.81%, which takes into account mostly the physical processes with sufficient calibration space.

After grouping the best performance model (Model 2) through clustering algorithm, a principle component analyses (PCA) is carried out to initially evaluate influence of the three possible factors (principle components) flow rate, length and rainfall to the performance of the quality models. As shown in the Figure 3, where the different axis represents a linear combination for the three features, flow rate, length and rainfall from the available dataset. All the points are the centroids of each of the different individual data (e.g. for one simulation the conduit that goes from $linkX$ to $linkY$ is represented by a point on the plot). The objective of this PCA is to see which feature influence more of the performance. And it is clear to conclude that sewer length is the most influential feature for these TSS models, since it captures almost all of the variability of the dataset (99.57%). While rainfall is not very well represented in this case study based on the current simulations, and it is still hard to arrive to a final conclusion. Studiyng the other principal components, it is difficult either to separate which model works best for the three given features. In order to arrive more clear conclusion, further exploratory analysis is applied to search more correlations between the factors and the model performances.

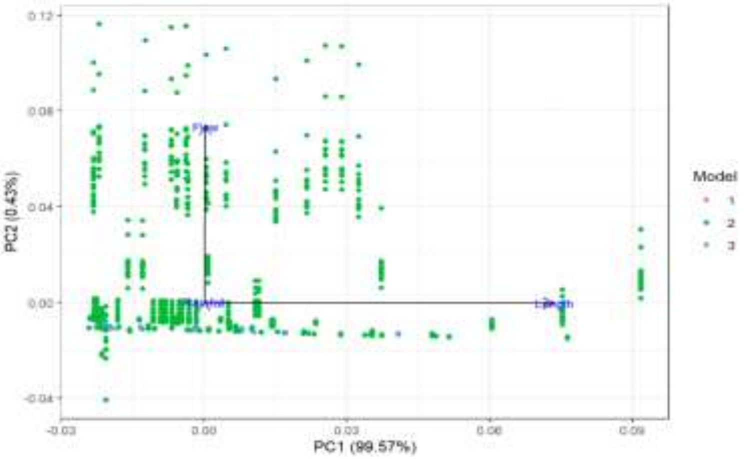

Figure 3. PCA for the influencing features

Correlation Exploratory Analysis

In order to explore further the correlation between each factor and the model performances, one of the sub-catchments of Luis Fargue catchment, the Perinot SN (Figure 4) with additional simulations under more representative rainfalls is applied.

The Perinot SN covers a total area of 260 ha with mainly residential uses. The sewer length for the Perinot SN is 3 km with an average slope of 0.007. A retention tank which is separated in three hydraulically connected bodies for a total storage volume of 35000 m^3 is also included. In order to simplify the tests and control afterward, sewers of similar dynamics in series are integrated as one, where five main sewers are presented (Figure 5).

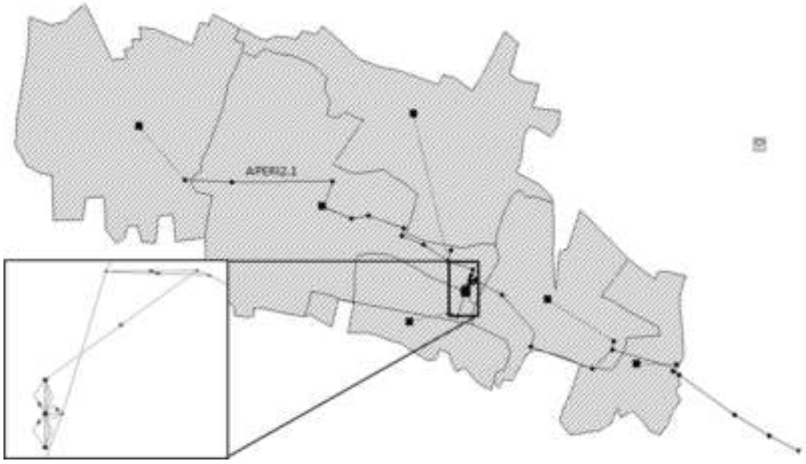

Figure 4. Detail topology of the Perinot sub-catchment

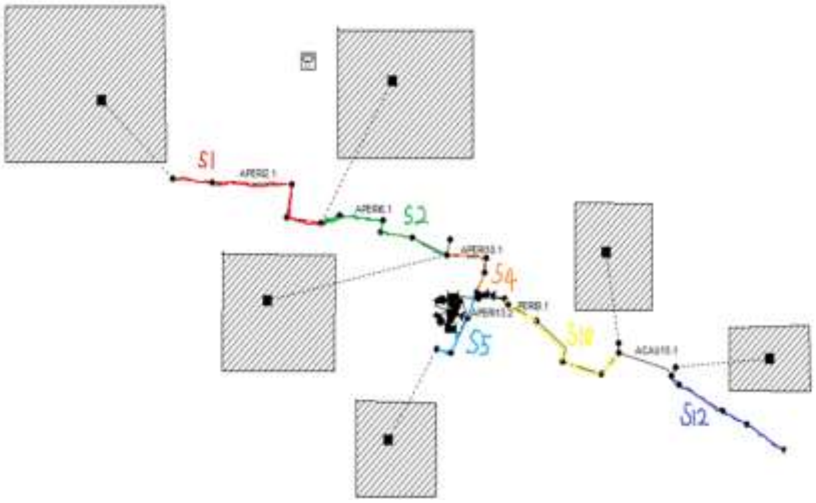

Figure 5. Perinot SN after pre-treatment

In order to consider a set of rain scenarios more representative of the rain intensities, rain scenarios for calibration and validation of the Perinot SN corresponding to real rainfall measured at France in the years 2003, 2007, 2011 and 2013. Besides, four different scenarios with various characteristics (Table 4) have been selected from historic data of 2007 (Figure 6), which represent different types of rainfalls with the 5-minute time step and 24 hours' duration.

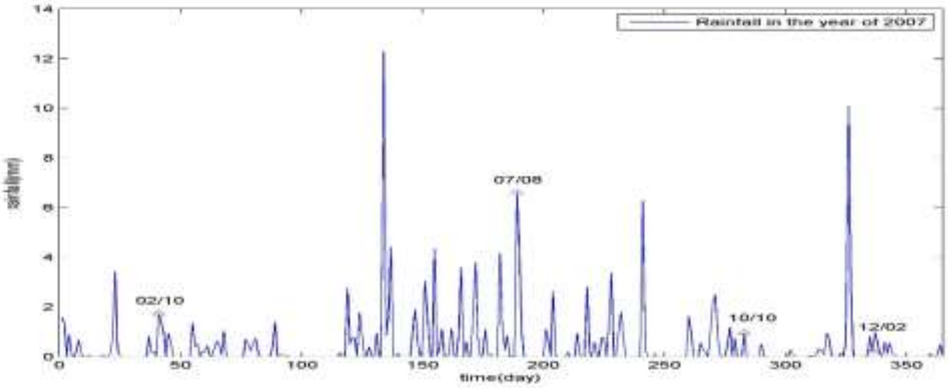

Figure 6. Rain Scenario of Perinot in the year of 2007

Table 4 Different rainfall scenarios

	1	2	3	4
Start time	Oct/10/2007	Dec/02/2007	Feb/10/2007	Jul/08/2007
	5	**6**	**7**	**8**
Start time	Aug/19/2003	Aug/02/2013	Jan/03/2011	Jan/03/2011
Duration	24h		**Time step**	5 min

The calibration process is also carried out using SWMM5 embedded with SWMM-TSS, Matlab and the GAMS optimization library. Table 5 provides more details about the arrangement about the rain scenarios, where the rain scenarios 1, 2, 5 and 7 will be used for calibration. After calibrating all the models, rain scenarios 3, 4, 6 and 8 will be used for validating the calibrated models.

Table 5 Test Arrangement

Rain	Calibration	Case1	Case2	Case3	Case4
	Validation	3 4	3 4	6 8	6 8
Sewer	*Si*	*Si* -1	*Si* -2	*Si* -5	*Si* -7
		Si -1-3	*Si* -2-3	*Si* -5-6	*Si* -7-6
		Si -1-4	*Si* -2-4	*Si* -5-8	*Si* -7-8

Si includes S1, S2, S4, S5, S10, S12; xx-xx-xx means sewer-calibration-validation

Correlation between sewer length and model performance

After the model calibration and validation, Table 6 and Figure 7 present the correlation relations between sewer length and the model performance. It seems that all three models present a similar tendency changing the sewer lengths. However, the performance of Model 3 changes more dramatically than Models 1 and 2. Table 7 shows how sewer length affects the coefficient parameters of Model 3. Sewer 5 and 10 perform worse with lower value of β but much higher e. Model 3 is generalized from the physical characteristics in a sewer, where the TSS dynamic is affected by the flow rate and time delays, which are affected directly by sewer length. Therefore, it can be concluded that up to some extent, the length of sewer has more impact on the performance of Model 3.

In conclusion, the length of sewer is more likely to influence the performance of Model 3. But in general, Model 1 and 2 seem to be good choices for sewers which length ranges from 400m to 900m.

Table 6 Relationship between sewer length and model performance

Sewer	Length (m)	Model 1 (%)	Model 2 (%)	Model 3 (%)
S4	156.20	90.31	90.45	80.37
S5	160.70	75.09	69.43	56.82
S2	482.60	93.10	93.12	86.61
S10	773.40	87.73	87.70	81.30
S12	879.20	94.02	94.09	92.46
S1	1181.90	79.53	79.91	62.63

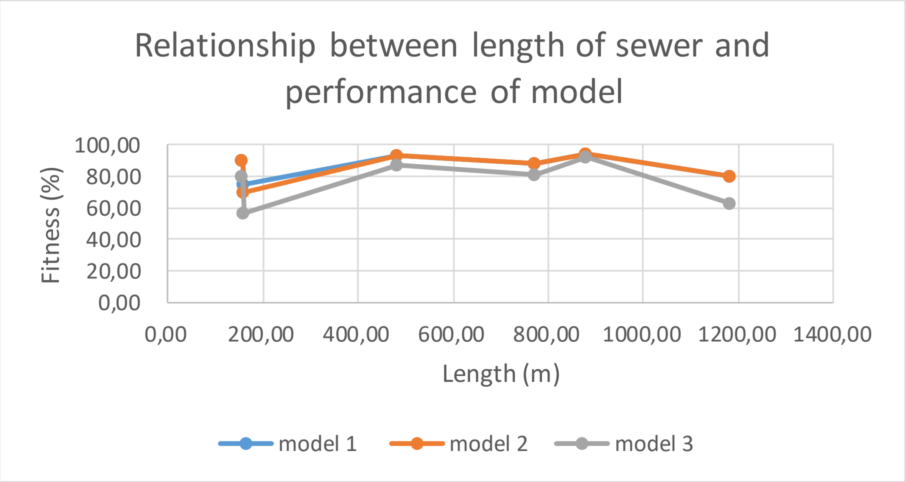

Figure 7. Relationship between sewer length and model performance

Table 7 Relationship between sewer length and Model 3 parameters

Sewer	Length (m)	β	e	Model 3 (%)
S4	156.20	0.96	10.35	80.37
S5	160.70	0.71	47.53	56.82

S2	482.60	0.97	9.65	86.61
S10	773.40	0.95	19.02	81.30
S12	879.20	0.97	11.74	92.46
S1	1181.90	0.73	86.04	62.63

Correlation between rainfalls and model performance

Table 9 illustrates correlation between rain intensity and model performances. As in Table 8, rainfall scenarios 1 and 2 can be regarded as light rain, while scenario 5 and 7 correspond to heavy rain. It seems that, with the increase of rainfall intensity, a decreases for Model 1; a_1 increases while a_2 decreases for Model 2; β decreases while e increases for Model 3. Overall, there is a tendency that the heavier rainfall is, the worse model performance will be obtained.

Table 8 Information of rainfall scenarios in calibration

Rainfall for Calibration	Total Depth of 24h(mm)	Intensity(mm/h)	Maximum Depth (mm)
1	5.53	0.23	0.04
2	0.25	0.01	0.06
5	1754.61	73.11	19.05
7	1667.78	69.49	45.71

Table 9 Relationship between rain intensity and model performance in calibration

Scenario	a	a1	a2	$\boldsymbol{\beta}$	e	Model 1(%)	Model 2 (%)	Model 3 (%)
case 1	0.47	0.48	0.47	0.89	15.66	92.79	92.87	80.98
case 2	0.46	0.54	0.46	0.92	21.07	92.30	92.36	78.17
case 3	0.35	0.65	0.35	0.88	36.30	88.78	91.16	71.71
case 4	0.42	0.58	0.42	0.84	49.86	80.35	81.62	67.25

Correlation between flow rate and model performance

Figure 8 shows the relationship between flow rate and TSS out of a sewer. Although there is no distinct evidence of the relationship between these two variables, we can still find that the trend of discharge is likely to be opposite against the trend of concentration of TSS in a sewer. This can be understood that the flow with large velocity will take away more TSS, thereby the concentration of TSS decreasing.

In SWMM, the manning equation is used to express the relationship between flow rate, slope, cross-sectional area and hydraulic radius in conduits. It shows that, the flow rate in a sewer is affected by many different parameters and physical characteristics of the conduit. However, how they affect the flow rate, and how they cross correlated among different parameters require further study in future. On another hand, flow rate and TSS are normally considered as the main variables to integrate SN and WWTP. In order to have a better integrated management of SN and WWTP, what is the underlying relationship between TSS and flow rate, as well as other potential parameters (i.e. slope, diameter) will be investigated as the next step.

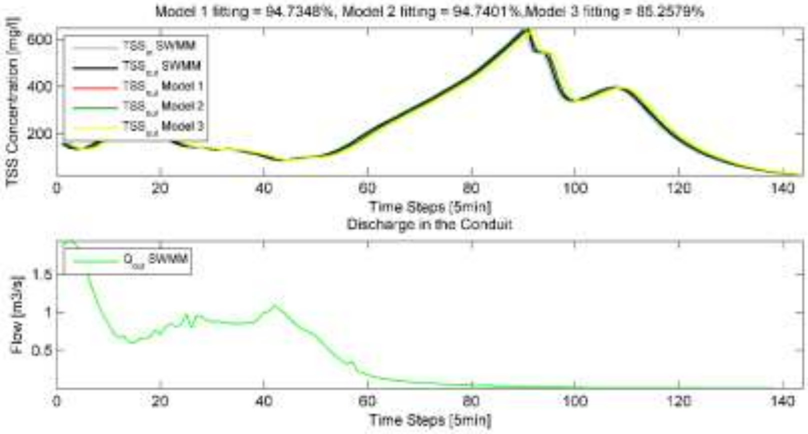

Figure 8. Flow rate and TSS behavior out of Sewer 1

CONCLUSIONS

This paper contributes a preliminary analysis of the influencing factors for the performance of SN quality models. In particular, a set of conceptual TSS models proposed previously are used for this study. Three possible factors rain intensity, sewer length, and flow rate are analysed through a series of tests. A clustering algorithm is used to explore preliminary the best performance model. Further exploratory analysis is carried out afterwards to obtain correlations between different influencing factors and the model performance. This study also contributes potentially to a better quality-based RTC management of SN and WWTP, which can lead to improvements in CSO and pollution reduction to the water environment.

Through this study, the following conclusive results are obtained:

1) Preliminary analysis shows that Model 2 and Model 3 perform better than Model 1. After comparing the Model 2 and Model 3, Model 2 shows better performance in more scenarios with an average fitting performance of 80.10%. On the other hand, PCA analysis concludes that the sewer length is the factor which can influence the performance most;
2) Further exploratory analysis illustrates that the sewer length is more likely to influence Model 3. Models 1 and 2 are good choices for sewers with length ranges from 400m to 900m. Also, there is a tendency that the heavier rainfall is, the worse model performance will be obtained.
3) From the further exploratory analysis, it may also be concluded that it is better to have larger parameters a, a_2, β and smaller a_2, e for the models to perform better when there is lack of rainfall data for calibration. Besides, the trend of discharge is likely to be opposite to the trend of TSS in a sewer, which can be explained because the flow rate with large velocity takes away more TSS.
4) More conclusive results were in expectations which can indicate clearer usability of the models in RTC, when specific parameter sets can be used directly; when re-calibrations are required, etc. However, the current analysis based on the available pilots did not provide enough evidences to support these conclusions. Therefore, further analysis is still required.
5) Cross correlation between model performance and the potential influencing factors sewer slope, diameters and velocity will also be investigated as the next step.

Acknowledgement

The authors wish to thank the support received by the European Commission research grant of project LIFE EFFIDRAIN (LIFE14 ENV/ES/000860), thank the support from Spanish national project DEOCS (DPI2016-76493-C3-3-R). The authors also want to thank Bordeaux Metropole and Aigües de Barcelona for the financial and technical support. This work is also supported by the Spanish State Research Agency through the María de Maeztu Seal of Excellence to IRI (MDM-2016-0656), by the Institute of Robotics Industry through the TWINs project.

References

Ackers, P. & White, W. 1973 Sediment transport: a new approach and analysis. *J. Hydra. Divi.*, 99(11), 2041-2060.

Butler, D., & Schütze, M. 2005 Integrating simulation models with a view to optimal control of urban wastewater systems. *J. Environ. Modell. and Softw.*, 20(4), 415-426, 2005. doi.org/10.1016/j.envsoft.2004.02.003

Becouze, C., Bertrand-Krajewski, J. -L., Dembélé, A., Cren-Olivé, & Coquery, M. 2009 Preliminary assessment of fluxes of priority pollutants in stormwater discharges in two urban catchments in Lyon, France. *In: 13th of IWA intern. confer. on Diffuse Pollu. and Integ. Waters. Manage (DIPCON), Seoul/South Korea*, 12-15 October 2009.

Beeneken, T., Erbe V., Messmer, A., Reder, C., Rohlfing, R., Scheer, M., Schuetze, M., Schumacher, B., Weilandt, M. & Weyand, M. 2013 Real time control of urban drainage systems–a discussion of the additional efforts compared to conventionally operated systems. *Urban Water Journal*, 10(5), 293–299. doi.org/10.1080/1573062X.2013.790980

Combes, V. 1982 *Etude de modèles mathématiques de transport des matériaux solides en réseau d'assainissement (Study of mathematical models of transport of solid materials in sewerage network).* DEA de mécanique, Institut National Polytechnique, Toulouse, France.

Clark, R., Pezzaniti, D. & Cresswell, D. 2002 Watercress—Community Resource Evaluation and Simulation System—A tool for innovative urban water system planning and design. *In: Proceedings of the Hydrology and Water Resources Symposium, Melbourne/Australia*, 20–23 May 2002.

Cembrano, G., Quevedo, J., Salamero, M., Puig, V., Figueras, J., & Martí, J. 2004 Optimal control of urban drainage systems. A case study. *J. Contr. Engin. Pract.*, 12(1), 1-9. doi.org/10.1016/S0967-0661(02)00280-0

Döring, R. 1989 *Real-time control of urban drainage systems. The state-of-art.* IAWPRC, International Association on Water Pollution and Control, University of Michigan.

DHI. 2007 MIKE URBAN. https://www.mikepoweredbydhi.com/products/mike-urban (accessed 24 October 2018)

Dirckx, G., Schütze, M., Kroll, S., Thoeye, C., De Gueldre, G. & Van De Steene, B. 2011 RTC versus static solutions to mitigate CSO's impact. *In: 12nd International Conference on Urban Drainage (ICUD)*, Porto Alegre/Brazil, 11-16 September 2011.

Ly, D.K. 2019 *Water quality-based real time control of combined sewer systems.* PhD thesis, Environmental Engineering. Université de Lyon - INSA Lyon, France.

EPA. 2006 *Real time control of urban drainage networks*, Technical Report EPA/600/R-06/120, Environmental Protection Agency, Washington D.C., United States.

Fu, G., Khu. S., & Butler, D. 2010 Optimal distribution and control of storage tank to mitigate the impact of new developments on receiving water quality. *J. Environ. Engine.*, 136(3), 335-342. DOI: 10.1061/(ASCE)EE.1943-7870.0000161

Fanlin M.; Guangtao F. & David B. 2017 Cost-effective river water quality management using integrated real-time control technology. *Environ. Sci. Technol.*, 51(17), 9876-9886. doi.org/10.1021/acs.est.7b01727

Gasperi, J., Garnaud, S., Rocher, V., & Moilleron, R. 2008 Priority pollutants in wastewater and combined sewer overflow. *J. Scien. of the Tot. Environ*, 407(1), 263-272. doi.org/10.1016/j.scitotenv.2008.08.015

García, L., Barreiro-Gomez, J., Escobar, E., Téllez, D., Quijano, N., & Ocampo-Martínez, C. 2015 Modelling and real-time control of urban drainage systems: A review. *Advances in Water Resources*, 85, 120-132. doi.org/10.1016/j.advwatres.2015.08.007

Huber,W.C. & Dickinson, R.E. 1988 *Storm Water Management Model*, User's Manual, U.S. Environmental Protection Agenc, Athens, GA, USA.

IFAK. 2005. SIMBA# 5.0 user's guide, https://nextcloud.ifak.eu/s/simba3?path=%2FFlyer#pdfviewer (accessed 1 October 2019).

Joseph-Duran, B., Ocampo-Martínez, C., & Cembrano, G. 2014 Hybrid modelling and receding horizon control of sewer network. *Water Resources Research*, 50(11), 8497-8514. doi.org/10.1002/2013WR015119

Joseph-Duran, B., Ocampo-Martínez, C., & Cembrano, G. 2015 Output-feedback control of combined sewer networks through receding horizon control with moving horizon estimation. *Water Resources Research*, 51(10), 8129-8145. doi.org/10.1002/2014WR016696

Krotov, V. 2017 *A Quick Introduction to R and RStudio*, Technical report. Murray State University, Murray, Kentucky. DOI: 10.13140/RG.2.2.10401.92009.

Lacour, C., & Schütze, M. 2011 Real-time control of sewer systems using turbidity measurements. *Water Science & Technology*, 63(11), 2628-2632. doi.org/10.2166/wst.2011.159

MWH 2010. InfoWorks CS. http://www.innovyze.com/products/infoworks_cs/ (accessed 24 October 2018)

Maruéjouls, T., Vanrolleghem, P., Pelletier, G., & Lessard, P. 2012 A phenomenological retention tank model using settling velocity distributions. *J. Wat. Res.*, 46, 6857-6867. doi.org/10.1016/j.watres.2011.11.067

Nash, J.E. & Sutcliffe, J.V. 1970 River flow forecasting through conceptual models part I-A discussion of principles. *Journal of Hydrology*, 10(3), 282-290. doi.org/10.1016/0022-1694(70)90255-6

Ocampo-Martínez, C., Puig, V., Cembrano, G. & Quevedo, J. 2013 Application of predictive control strategies to the management of complex networks in the urban water cycle [applications of control]. *IEEE Control Systems*, 33(1), 15–41. DOI: 10.1109/MCS.2012.2225919

Puig, V., Cembrano, G., Romera, J., Quevedo, J., Aznar, B. R., & Cabot, J. 2009 Predictive optimal control of sewer networks using CORAL tool: application to Riera Blanca Catchment in Barcelona. *Water science & Technology*, 60(4), 869-878. doi.org/10.2166/wst.2009.424

Pleau, M., Fradet, O., Colas, H. & Marcoux, C. 2010 Giving the rivers back to the public. Ten years of real time control in quebec city. *In: International Conference on Sustainable Techniques and Strategies for Urban Water Management (NOVETECH), Lyon/France*, 28-30 June 2010.

Rodriguez, A. & Laio, A. 2014 Clustering by fast search and find of density peaks. *Science*, 344 (6191), 1492-1496. DOI: 10.1126/science.1242072

Rossman, L. 2015 *Storm Water Management Model*, Users' Manual Version 5.1, *Envir. Prot. Agn*, Washington D.C., USA.

Richard, E. 2016 *GAMS - A Users' Guide.* Washington, DC, USA.

Schütze, M., Campisano, A.C., Colas, H., Schilling, W. & Vanrolleghem, P.A. 2004 Real time control of urban wastewater systems - where do we stand today? *J. Hydro.*, 299(3-4), 335-348. doi.org/10.1016/j.jhydrol.2004.08.010

Soldevila, A., Blesa, J., Tornil-Sin, S., Duviella, E., Fernandez-Cantí, R.M. & Puig, V. 2016 Leak localization in water distribution networks using a mixed model-based/data-driven approach. *Control Engineering Practice*, 55, 162-173. doi.org/10.1016/j.conengprac.2016.07.006

Sun, C.C., Joseph, B., Maruejouls, T., Cembrano, G., Muñoz, E., Meseguer, J., Montserrat, A., Sampe, S., Puig, V. & Litrico, X. 2017a Efficient integrated model predictive control of urban drainage systems using simplified conceptual quality models. *In: 14th IWA/IAHR International Conference on Urban Drainage (ICUD)*, Prague/Czech, 10-15 September 2017.

Sun, C.C., Joseph-Duran, B., Maruejouls, T., Cembrano, G., Meseguer, J., Puig, V., & Litrico, X. 2017b Real-time control-oriented quality modelling in combined urban drainage networks. *In: IFAC 2017 World Congress, Toulouse/France*, 9-14 July 2017.

Sun, C.C., Cembrano, G., Puig, V. & Meseguer, J. 2018a Cyber-physical systems for real-time management in the urban water cycle. *In: 4th International Workshop on Cyber-Physical Systems for Smart Water Networks*, *Porto/Portugal*, 10 April 2018.

Sun, C.C., Joseph, B., Cembrano, G., Puig, V. & Meseguer, J. 2018b Advanced integrated real-time control of combined urban drainage systems using MPC: Badalona case study. *In: 13th International Conference on Hydroinformatics (HIC)*, *Palermo/Italy*, 1-6 July 2018.

Sun, C.C., Romero, L.B., Joseph-Duran, B., Meseguer, J., Craviotto, E.M., Guasch, R.P., Martinez, M.P., Puig, V. & Cembrano, G. 2020 Integrated Pollution-based Real-time Control of Sanitation Systems. Journal of Environmental Management. To appear.

Torres-Matallana, J.A., Leopold, U., Klepiszewski, K. & Heuvelink, G.B.M. 2018 EmiStatR: A Simplified and Scalable Urban Water Quality Model for Simulation of Combined Sewer Overflows. *Water*, 10, 782. doi.org/10.3390/w10060782

van Rijn, L. 1984 Sediment transport, part II: suspended load transport. *J. Hydra. Engine.*, 110(11), 1613-1641. doi.org/10.1061/(ASCE)0733-9429(1984)110:11(1613)

Vezzaro, L., Christensen, M.L., Thirsing, C., Grum, M. & Mikkelsen, P.S. 2014 Water quality-based real time control of integrated urban drainage systems: A preliminary study from Copenhagen, Denmark. Procedia Engineering 70, 1707-1716. doi.org/10.1016/j.proeng.2014.02.188

Vezzaro, L., Pedersen, J.W., Larsen, L.H., Thirsing, C., Duus, L.B. & Mikkelsen, P.S. 2020 Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations at WWTP inlets. *Water Science & Technology*, 81(1), 109-120. doi.org/10.2166/wst.2020.085

Wiuff, R. 1985 Transport of suspended material in open and submerged streams. *J. Environ. Eng.* 111(5), 774-792. doi.org/10.1061/(ASCE)0733-9429(1985)111:5(774)