Identifying the best locations to install flow control devices in sewer networks to enable in-sewer storage

J. P. Leitão[1]*, J. P. Carbajal[1], J. Rieckermann[1], N. E. Simões[2], A. Sá Marques[2], L. M. de Sousa[1]

[1]Eawag: Swiss Federal Institute of Aquatic Science and Technology. Überlandstrasse 133, 8600 Dübendorf, Switzerland. Email: joaopaulo.leitao@eawag.ch

[2]MARE, Department of Civil Engineering, University of Coimbra, Portugal

Abstract

The activation of available in-sewer storage volume has been suggested as a low-cost flood and combined sewer overflow mitigation measure. However, it is currently unknown what the attributes for suitable objective functions to identify the best location for flow control devices are and the impact of those attributes on the results. In this study, we present a novel location model and efficient algorithm to identify the best location(s) to install flow limiters. The model is a screening tool that does not require hydraulic simulations but rather considers steady state instead of simplistic static flow conditions. It also maximises in-sewer storage according to different reward functions that also considers the potential impact of flow control device failure. We demonstrate its usefulness on two real sewer networks, for which an in-sewer storage potential of approximately 2,000 m^3 and 500 m^3 was estimated with five flow control devices installed.

Keywords: in-sewer storage, flow control device location, storm water management, best management practice, urban floods

1. Introduction

Urban pluvial floods have received a considerable amount of attention in recent years due to the negative impacts they can have on society. This type of flood event is caused by intense rainfall associated with the limited hydraulic capacity of drainage systems. Their consequences can range from small economic losses (e.g. in June 2016 storm flood events in

various regions of Switzerland caused delays in rail services[1]) and to the loss of human life (e.g. a flash flood event on the 23rd of August of 2015 in Montpellier (France) caused three deaths[2].,) Due to increasing urbanisation and changes in rainfall patterns, flood events are becoming more frequent (United Nations, Department of Economic and Social Affairs, Population Division, 2014) and may start occurring in areas that do not have a record of flooding (Houston *et al.*, 2011). It is thus necessary to find flexible and affordable solutions to address the problems caused by this type of natural hazard.

Various solutions, also referred to as Best Management Practices (BMPs), to prevent, or at least mitigate, urban pluvial flooding have been proposed based on the common concept of reducing flow peaks (Fletcher *et al.*, 2015). This can be achieved by increasing catchment infiltration rates, for example by constructing infiltration basins (e.g. Travis and Mays, 2008), or by storing storm water upstream of flood prone areas (e.g. Bennett and Mays, 1985; Cunha *et al.*, 2016). Perez-Pedini *et al.* (2005) proposed a methodology using these ideas to determine the optimal number and location of infiltration-based structures to reduce peak flood flows. Improving on a previous work by Mays and Bedient (1982), Bennett and Mays (1985) proposed a model to determine the location of detention basins. Following these early works, several other models have been proposed to find the optimal location and sizing of storm water storage units, e.g. Zhen *et al.* (2004), Lim *et al.* (2014), Oxley and Mays (2014) and Cunha *et al.* (2016). However, because these solutions usually require a significant amount of space to be built, they are difficult to implement in dense and consolidated urban areas (Baek *et al.*, 2015). Additionally, a number of these solutions come with large up front investments (construction costs) as well as imposing large operation and maintenance costs.

---

[1] http://www.swissinfo.ch/eng/society/pouring-down_switzerland-battered-by-heavy-rains-and-floods/42214938
[2] http://www.euronews.com/2015/08/23/deadly-storms-in-south-of-france-cause-travel-chaos

To address the pitfalls of the systems mentioned above, one can alternatively consider local flow control devices in sewer manholes to utilise in-sewer storage potential. This is the focus of the CENTAUR project[3]. Using the storage capacity in existing sewer networks is expected to have significant advantages over the more conventional flood mitigation solutions by 1) reducing upfront initial investment, 2) no additional space requirements necessary, which is especially important in densely urbanised areas, and 3) the potential for flexibility throughout the drainage system in dealing with uncertain future conditions, such as climate and catchment land use (and infiltration) changes. These advantages can be realized because flow control devices are comparably easy to install and remove as needs develop and change over time. The potential benefits of the installation of flow control devices to take advantage of in-sewer storage spare volume have been investigated in real-scale laboratory conditions by Shepherd et al. (accepted); the preliminary results of these investigations have shown that the water depth downstream the flow controlled section can be reduced by approximately 30% and the flood volume can be reduced by up to 25%. Further investigations are being carried out in real urban drainage networks as part of the CENTAUR project.

To maximise the impact of a local flow control device solutions in reducing pluvial flooding, it is necessary to identify the best location(s) to install such flow control devices. The various approaches to optimize BMPs suggest deterministic methods, such as linear models (Doyle *et al.*, 1976) and scatter-search meta-heuristics (Zhen *et al.*, 2004), dynamic programming algorithms (Bennett and Mays, 1985 and Travis and Mays, 2008), or stochastic, such as

genetic algorithms (Perez-Pedini *et al.*, 2005; Fu *et al.*, 2010; Lim *et al.*, 2014, Cimorelli *et al.*, 2016) and simulated annealing models (Oxley and Mays, 2014; Cunha *et al.*, 2016). The main objective in these different methods is to minimise costs subject to flood and capacity (size) constraints. These methods require an initial analysis step where possible locations for storage units need to be pre-defined. Subsequently, the number of units and their sizes are computed by relying on hydraulic simulations to determine the optimal solution.

In contrast, for the local flow control device solution the location model is, in this regard, more attractive because (i) the possible candidate locations are pre-defined and correspond to all existing sewer network manholes or one of their possible subsets obtained by excluding manholes with specific conditions, such as difficult physical access conditions, and (ii) the cost of the flow control device is likely to have less variability than the construction of new physical storage as it does not depend on the surrounding cityscape. The process thus becomes a ranking of possible configurations that depend on previous installation decisions and the available storage volume. However, identifying the best locations for installing a flow control devices is not trivial because knowledge on how many sewers are affected by the installation of one flow control device is not *a priori* known since, when closed, a flow control device can surcharge more than the immediately upstream sewer and can be hampered because hydraulic models are not always available. To make this solution valid for systems with these knowledge deficits, the location model should only be based on network topology data, such as manhole and drainage pipe data. In the event that the network topology is not available, subcatchment areas (or a Digital Elevation Model to automatically derive them), is necessary to generate a probable sewer network based on, for example, image analysis methods and empirical knowledge.

4

In this manuscript, we present a new location model which can be used as a screening tool to identify the best location(s) to install flow control devices to reduce urban pluvial flooding by maximizing the in-sewer storage potential without requiring hydraulic simulations. The main novelty of this work is a new flow control device location model, which not only proposes various reward (or objective) functions developed to encapsulate the available volume in the system and the potential impact of a flow control device failure, it also incorporates a proxy to the flow hydraulic energy (*steady state*) when calculating the storage volume available for each possible flow control device location. The suggested model uses manhole and drainage sewer register data, which can be obtained from geographic databases or hydraulic model data when they are available.

The manuscript is organised as follows: in section 2 the proposed location model is presented. The case study sewer networks are described in section 3. Results are presented and discussed in section 4. In section 5 the main conclusions are presented and future work is discussed.

2. The location model

2.1. General

The presented location model, called *CENTAUR.loc*, is based on a geospatial database that currently takes a text file as input. The file describes a sewer system network with nodes, edges and catchment elements in a tabular form. The data model defines a directed graph and is implemented in a relational database supported by the PostgreSQL database management system (Stonebraker and Kemnitz, 2006). This information is converted into formal geo-spatial objects relying on the PostGIS (Holl and Plum, 2009) extension. The model software

and database structures were developed using Java and PL/pgSQL programming languages. The code is available under an open source license, and is publically (and freely) available on a GitHub repository: https://github.com/ldesousa/centaur.loc. See the supporting information (Appendix A) for additional details on the software.

2.2. Flow energy line

In a recent work by Philippon et al. (2015), flow velocity is not considered to estimate the in-sewer storage potential, i.e. flow is assumed to have zero velocity (*static* assumption; left panel in Figure 1). However, because the flow control device may not be fully closed, originating a non-horizontal hydraulic energy line due to the flow velocity and corresponding energy (head) losses, the flow may be approximated to *steady state* (*steady state* assumption; right panel in Figure 1). Introducing the *steady state* assumption is one of the novel developments of the proposed model.
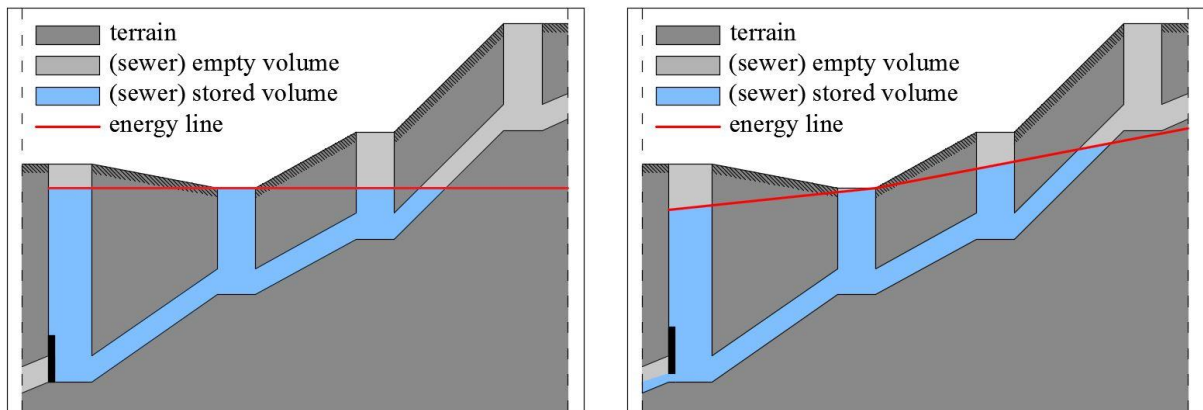

Figure 1. The flow control device location model improves the *Static* assumption (left) to a *Steady state* assumption (right).

The *steady state* assumption is a simplified approach based on the Manning-Strickler equation (Eq. 1) that avoids the computational cost of detailed hydraulic simulations.

$$Q_{max} = \frac{1}{n} \cdot A \cdot R_h^{\frac{2}{3}} \cdot J^{\frac{1}{2}} \tag{1}$$

where Q_{max} is the full cross-section flow, n is the Manning roughness coefficient, A is the cross-section area, R_h is the hydraulic radius and J is the head loss per length of sewer (or the sewer slope). The roughness coefficients used in Eq.1 are obtained directly from the input data (e.g. pipe and manhole data available at GIS inventory) and can vary from sewer to sewer. Since roughness coefficients influence the calculation of maximum flow (Q_{max} in Eq. 1), changing the roughness coefficient of a specific sewer can impact the in-sewer storage potential sewer ranking. Consequently, in case a sewer roughness coeficient is altered (updated), a different ranking maybe obtained. However, if roughness values of all sewers are affected by one same factor (i.e. $\pm$ x%), the final ranking is not changed because the relation between the roughness values of the different sewers is not altered.

The computation begins by calculating the maximum flow, i.e. full cross-section flow, for each network sewer based on the Manning-Strickler equation (Eq. 1). Then for each possible location of a flow control device (i.e. manhole), the maximum possible flow for the upstream sewers is calculated using the general formulation in Equation 2.

$$Q_{p,i} = min(Q_{max,i}, Q_{p,d_i}) \tag{2}$$

where $Q_{p,i}$ is the possible maximum flow for the sewer to which the flow is being calculated, $Q_{max,i}$ is the maximum flow of the i sewer calculated using the Manning-Strickler equation (Eq. 1), and $Q_{p,di}$ is the possible flow of the downstream sewer. This is done iteratively so that

the flow is always equal or smaller in the upstream sewer network. There are three special cases that use different formulations (Eqs. 3 to 5):

- Sewer immediately upstream of the gate: $Q_{p,i} = Q_{max,i}$　　　　　　　　　(3)

- Sewer join: $Q_{p,i} = min\left(Q_{max,i}; Q_{max,i} \times \frac{Q_{p,d_i}}{\sum Q_{max,u_i}}\right)$　　　　　　(4)

- Sewer bifurcation: $Q_{p,i} = min(Q_{max,i}; \sum Q_{p,d_i})$　　　　　　　　(5)

where $\sum Q_{max,ui}$ is the sum of the maximum flow (full section) of each of the immediately upstream sewers calculated with the Manning-Strickler equation. By taking these special cases into consideration, flow continuity is ensured.

As presented in Figure 2, the slope of the energy line (J_i) for each sewer is then calculated using Eq. 1, solved in order to J, with Q_{max} equal to the $Q_{p,i}$.
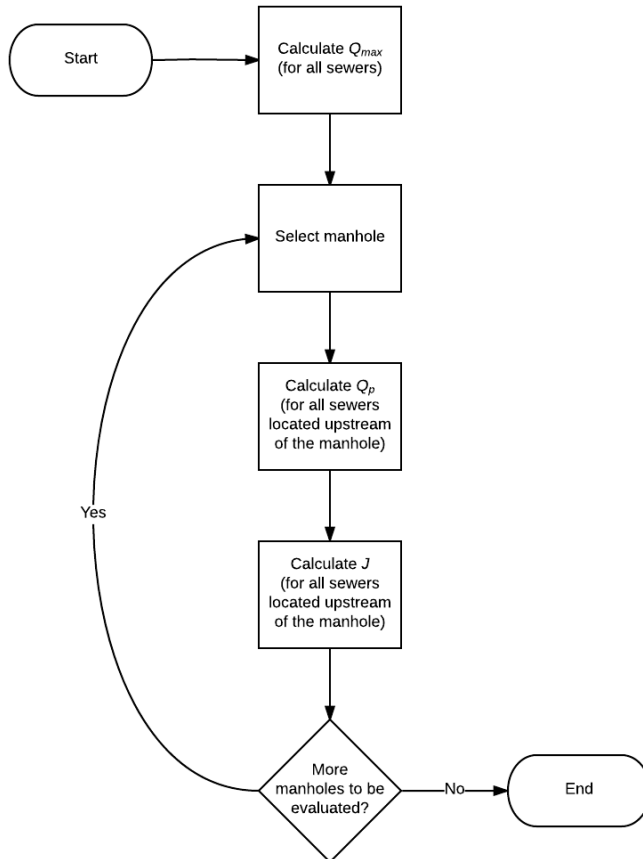
Figure 2. Flowchart of the calculation procedure of the slope of the flow energy line (J) of the steady-state assumption

This slope is then used to adjust the overflow height offset computed for the node and propagated upstream in the computation of flooded volumes. The height offset is calculated with the projected length of the sewer in order to take into account the accurate geometric relations between the slope and the length of the sewer.

2.3. Computation of the in-sewer storage potential

2.3.1. Pre-calculations

The database harbours a number of functions and views that are used to pre-calculate some elements that are required to subsequently correctly implement the flow control device location algorithm. A specific function is used to fetch all the nodes upstream of a given node and from this set of nodes the set of subcatchments that contribute with runoff to the node under assessment is identified. By using the subcatchment attributes it is thus possible to calculate the total and impervious areas that contribute runoff to the specific node in question. The database views also facilitate the display of network elements in a desktop GIS software, such as QGIS (Graser, 2013). In this way geo-spatial elements can be styled and used in cartograms as ordinary geo-referenced data.

2.3.2. Computation of in-sewer storage potential

The calculation of the in-sewer storage potential at each node is achieved with a recursive process developed in the Java language using the Java Hibernate technology (Holl and Plum, 2009) and is repeated for each node of the network. The process starts with the computation of the node surcharge level (ground level) by adding the node invert level to its depth. A

safety margin factor is subtracted from this value, which was in this case found to be 0.1 m. The search then proceeds recursively upstream, identifying the links arriving at the node, as well as their nodes of departure and so forth.

When arriving at an upstream node the algorithm compares the node ground level with the current flow energy line level. Three cases can present themselves at this stage:

1. The node invert level is higher than the flow energy line level: the search stops in this branch of the network.

2. The node invert level is lower than the flow energy line level and the ground level is higher than the flow energy line level: the search proceeds upstream in the branch.

3. The node ground level is lower than the flow energy line level: the flow energy line level is updated (with the former value) and the search proceeds upstream in the branch.

Additionally, the algorithm will no longer proceed when the following upstream situations are reached:

a. The node has no incoming links.

b. The upstream link is a pump (in which case the overflow head is updated to the invert level of the node).

c. The upstream link is a weir (in which case the overflow head is updated to the crest level).

d. More than one link departs from the node, i.e. the node is a bifurcation (in this case the overflow level is updated to the invert level of the node).

At the end of this search, a set of links and nodes surcharged by a hypothetical flow control device installed at the node under assessment is identified, plus an overflow level. This overflow level is then used to calculate the volume stored at each of the surcharged links, using their cross-section data.

With the information on upstream impervious area, number of upstream subcatchments, set of surcharged links and their storage volume pre-calculated, the location algorithm for multiple flow control devices is straightforward. Each of these variables can be combined in different ways to guide the algorithm, following the reward functions described in Section 2.4. A database function is used to rank nodes according to a particular reward function and the highest scoring node is selected. This node, plus the set of surcharged nodes associated with it, are then flagged as occupied and thus discarded from further searches. The process is then repeated to locate additional gate(s). It is also possible to only search the network upstream of a specified node. This is useful if the goal is to protect a particular city district or block from flooding.

2.4. Reward functions

In the proposed location model, five different reward functions were implemented to either maximise the storage potential and/or maximise the storage potential while also minimising the potential negative impact of a flow control device failure. Flow control device failures can be caused by problems in flow control device operation due to mechanical or energy supply defects.

Reward (or objective) functions (*R(x)*) have the conceptual mathematical structure of Eq. 6, which can also be described in its logarithmic form (Eq. 7).

$$R(x) = \frac{f(x)}{g(x)}, \tag{6}$$

$$logR(x) = logf(x) - logg(x), \tag{7}$$

where $f(x)$, for the specific case of this study, represents the terms that increase storage potential, and $g(x)$ represents the terms that may contribute to increase the impact of a flow control device failure.

To explain the rationale behind the five reward functions proposed in this study to identify the best locations to install a flow control device, a simple example focused only on storage contribution in the first term and the impact of a flow control device failure in the second term is used (Eq. 8).

$$logR = log(vars.\ that\ increase\ storage) - log(vars.\ that\ increase\ failure\ impact) \tag{8}$$

For example, given the ith manhole in the network, the in-sewer potential is the sum of the volumes of the upstream sewers V_j that can be surcharged without causing overflow. This reward function can be described as Eq. 9.

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right), \tag{9}$$

where j runs over the m upstream pipes of manhole i.

Now, assuming that larger contributing runoff areas may increase the impact of a flow control device failure, this should also be considered in the reward function; it can be described as in Eq. 10.

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right) - log(A_i), \quad\quad (10)$$

where A_i is a measure of total operational impact in the ith manhole.

With this simple example, it can be seen that the particular form of the reward function depends on the details of the two terms. On the one hand the objective is to maximise the storage volume ensuring that there is enough water to fill the mobilised storage. On the other hand, an increase of the contributing runoff area, as well as of the number of contributing subcatchments, may contribute to increase the impact of an operational failure (due to the increase of the potential damage). To assign the same weight to the three variables and avoid variable unit dependencies, the reward functions are calculated using re-scaled values that range from 0 to 1. A simple re-scaling function was used: division of each variable value by its corresponding maximum value for the network under analysis.

The reward functions described below were implemented in the presented tool. In the future, these reward functions may be modified by adjusting the weights of their variables, or even new reward functions can be considered to represent different objectives.

RF1: [Volume]. Reward function RF1 is described by Eq. 11 and only considers storage volume, disregarding the impact of flow control device failure (or assume they have cancelled out).

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right) \tag{11}$$

RF2: [Volume] and [Contributing runoff Area]. Reward function RF2 (Eq. 12) also does not consider the possible impact of a flow control failure, or they have been cancelled out to give a numerator equal to $\sum_{k=1}^{n} A_k$, where *n* is the number of contributing catchments. If it is assumed that potential flow control failure impacts are ignored, the term $\sum_{k=1}^{n} A_k$ aims at representing the role of storable storm water in each sewer network location to reduce flooding or combined sewer overflow (CSO) discharge events.

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right) + log\left(\sum_{k=1}^{n} A_k\right) \tag{12}$$

RF3: [Volume] and [Contributing runoff Area] over [Number of Contributing Catchments]. This reward function (RF3) assumes that the number of upstream subcatchments can influence the assessment of the impact due to an operational failure of the flow control. It is a combination of the reward function RF2 and the flow control (impact) given by the number of upstream subcatchments (Eq. 13).

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right) + log\left(\sum_{k=1}^{n} A_k\right) - log\left(\sum_{k=1}^{n} n_i\right) \tag{13}$$

RF4: [Volume] over [Contributing runoff Area]. This reward function considers that the flow control device failure impact model is given by $(\sum_{k=1}^{n} A_k)^2$. When reward function RF2 is considered and modified with the flow control device failure impact model, the reward function can be presented as Eq. 14. Also, this reward function may also represent the objective to maximise the storage duration.

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right) - log\left(\sum_{k=1}^{n} A_k\right) \tag{14}$$

RF5: [Volume] over [Number of Contributing Catchments]. This reward function results from the combination of the reward function RF2 with the flow control device impact model given by $\sum_{i=1}^{p} n_i \sum_{k=1}^{n} A_k$; this results in Eq. 15.

$$log(R_i) = log\left(\sum_{j=1}^{m} V_j\right) - log\left(\sum_{i=1}^{p} n_i\right) \tag{15}$$

The storage potential (V) is a key variable for all of the reward functions and is therefore maximised in all reward functions. For the contributing runoff area (A) variable one can think of two different roles it can have in the proposed reward functions. First, by assuming that to achieve the objective of utilising the full storage potential, sufficient storm water is required - this corresponds to maximising the contributing runoff area variable. Second, by assuming that in case of a flow control device failure, an increase of the contributing runoff area may lead to larger water volumes, large pipe surcharges and consequently higher damage potential. In the latter case, the contributing runoff area should be minimised. Similarly, assuming that an increase of the number of contributing catchments (n) is a proxy for the total

size of the contributing catchment which in turn may increase the potential for flow control device failure, this variable should also be minimised in the reward function.

When considering the reward functions that aim only at maximising the storage volume (Table 1), the algorithm tends to select locations that mobilise large volumes, which are associated to large-diameter sewers. A failure of the flow control device may have a considerable impact, such as increased flooding in the vicinity of the flow control device location. In contrast, if small flow control devices are installed, the impact in the case of a failure is minimised. In this case the benefits of smaller installations outweigh the maximum storage volume per location.

Table 1. Reward functions and flow control device main objectives

| Reward functions	Maximise storage potential*	Maximise storage potential taking the potential impact of a flow control device failure into account**
RF1 *[Volume]*	x	
RF2 *[Volume]* and *[Contrib. Area]*	x	
RF3 *[Volume]* and *[Contrib. Area]* over *[Number of Contrib. catchments]*	x	x
RF4 *[Volume]* over *[Contrib. Area]*		x
RF5 *[Volume]* over *[Number of Contrib. catchments]*		x

* Also privileging the installation of a small number of large flow control devices.
** Also privileging the installation of various small flow control devices.

3. Case studies and model validation

3.1. Case studies

Two highly urbanised catchments (Figure 3) have reported severe pluvial flooding in recent years (Leitão et al., 2016; Leitão, 2009). The *Zona Central* catchment with approximately 1.5 km^2 is located in Coimbra (Portugal). The *D16&D17* catchments are located in Alcântara (Lisbon, Portugal) and cover an area of 0.9 km^2.
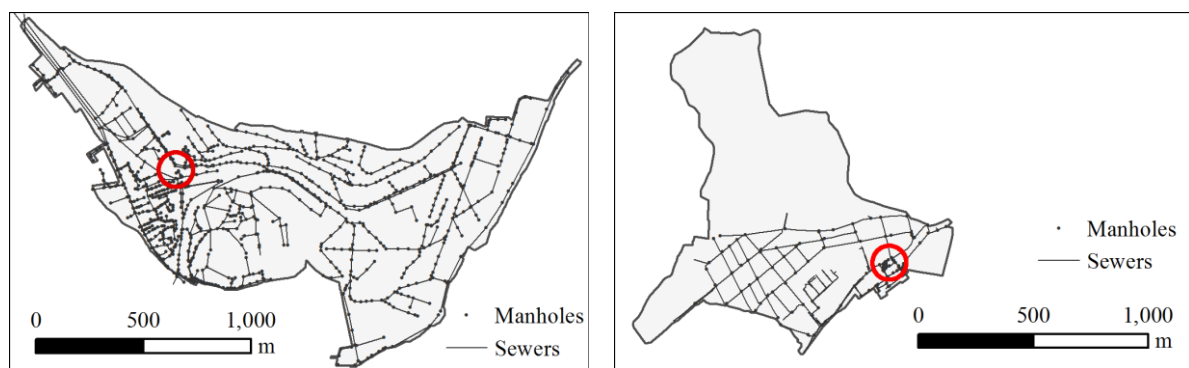
(a) *Zona Central* catchment, Coimbra, PT   (b) *D16&D17* catchments, Lisbon, PT

Figure 3. Catchments' elevation and respective sewer networks. The red circles represent the flood-prone areas.

The sewer networks of the two catchments are substantially different. The drainage system of the *Zona Central* catchment is based on a 34.8 km long combined sewer network, which comprises 911 subcatchments, 980 manholes and 1,001 pipes. The sewer network of the *D16&D17* catchments is also combined, approximately 7.5 km long and comprises 129 manholes and 136 sewers (Table 2).

Table 2. Statistics of the case study drainage networks. The network node degree is the number of sewers that are connected to a manhole, which summarises the network topology.

	Zona Central catchment				D16&D17 catchments			
	Min.	Max.	Mean	St. dev.	Min.	Max.	Mean	St. dev.
Network node degree (-)	1	6	1.99	0.65	1	4	2.27	0.90
Sewer length (m)	1.98	345.65	34.25	33.14	1	175.21	56.69	40.52
Sewer cross-section (m)	0.16	2.63	0.46	0.40	0.2	2.55	1.01	0.58
Sewer slope (%)	-0.28	0.62	0.06	0.07	-0.07	0.3	0.04	0.05

*These values may be due to errors in the network registry.

The *Zona Central* catchment case study is used to: (Section 4.1) validate the model by comparing computed storage volumes with actual field measurements and hydrodynamic simulation results; (Section 4.2) show the potential of the proposed model to assess the in-sewer storage potential of a sewer system, and (Section 4.3) evaluate the influence of the five reward functions proposed on the flow control device location results. The *D16&D17* catchments case study is used to demonstrate the influence of these reward functions.

3.2. Location model validation

In order to validate the results of the proposed location model, the storage potential of two specific network locations (i.e. manholes) were analysed. The results obtained using the static assumption were compared with *in-situ* measurements, whereas the results of the *steady state* assumption were compared with hydraulic simulation results. The two locations considered were node 64 and node 75 in the Zona Central catchment sewer network.

3.2.1. Static assumption

A two steps calculation was used for this process; (i) identification of the lowest upstream manhole terrain level. This represents the maximum level that the water can reach upstream

of the flow control device. (ii) calculation of the total storage volume as the free volume inside manholes and conduits in the upstream branch(es) of the flow control device location, below the maximum level calculated in (i).

3.2.2. Steady-state assumption

In order to compare the *steady-state* assumption results of the proposed model, hydrodynamic simulations using the EPA SWMM model (Rossman, 2015) were conducted. The EPA SWMM model used has been calibrated based on rainfall events of different intensity, and has been used in other urban drainage/ urban flooding studies focusing on flood issues reported in the catchment (Sá Marques et al., 2016; Maluf et al., 2017). The simulations focused only on the sewer network upstream the two pre-defined locations. The inflows considered for the hydraulic simulations (EPA SWMM) were directly added to the network nodes (manholes) and correspond to those obtained during the calculation of the hydraulic grade line (the maximum possible flow, $Q_{p,i}$) with the steady-state assumption (Sub-section 2.2). Additional details of the hydrodynamic model and results are presented in Appendix B.

4. Results and discussion

4.1. Validation of model results and comparison between the static and steady state assumptions

In order to validate the proposed location model, the in-sewer storage potential estimated for two specific locations at *Zona Central* (manholes with IDs 64 and 75) using the *static* and the *steady-state* assumption were compared with the in-sewer storage volume calculated based on field measurements and hydraulic simulations, respectively (Table 3). The number of surcharged pipes and manholes differ significantly between the two test manholes (Figure 4).

Table 3. Comparison of the storage potential results using the location algorithm with the upstream volume calculated by *in situ* and hydrodynamic analysis (results obtained using reward function RF1)

Manhole ID	*in situ* analysis Volume (m³)	Hydrodynamic simulations Volume (m³)	Location model					
			Static assumption			*Steady state* assumption		
			Volume (m³)	Dif* (%)	Dif** (%)	Volume (m³)	Dif* (%)	Dif** (%)
64	87.2	114.5	87.1	**-0.1**	-23.8	111.3	27.6	**-2.8**
75	26.3	28.8	26.8	**1.9**	-6.9	27.8	5.7	**-3.5**

*Difference to the *in-situ* analysis values

**Difference to the hydrodynamic (EPA SWMM) simulation results

As seen in Table 3, the results obtained based on the *static* assumption are very close to the volumes calculated manually (*in-situ* analysis); the differences range from -0.1% to 1.9%. The results obtained using the *steady state* assumption were compared with the results obtained from the hydrodynamic simulations and in this case the differences are less than 3.5%. In both cases the proposed location model produces adequate results. It is also interesting to compare the results produced by the two assumptions: the values obtained can differ as much as 22%. The in-sewer storage process is not static because as long as there is storm water and/or wastewater flow velocity is not zero which creates changes in the hydraulic head (i.e. pressure). Therefore, the steady-state assumption is considered more realistic. Based on this, the further analyses throughout this study were conducted using only the *steady-state* assumption.
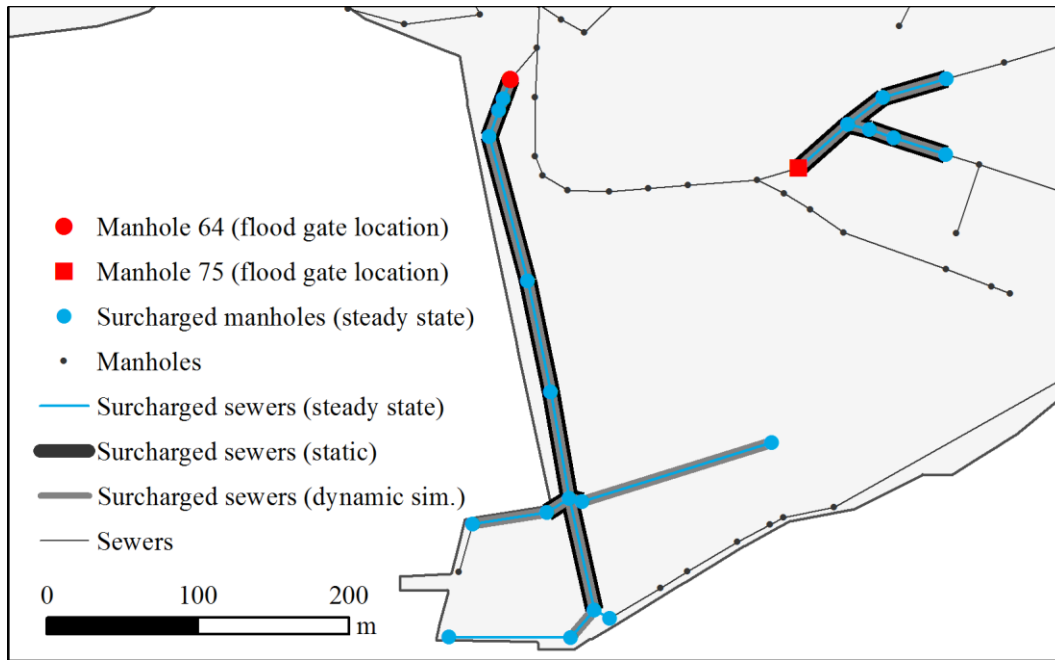
Figure 4. Surcharged manholes and sewers when installing (and closing) the flow control flow control device in manholes 64 and 75
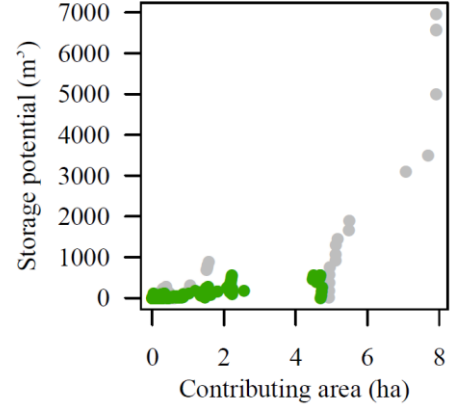
4.2. Evaluation of the in-sewer storage potential

The volumes calculated by the model represent the individual maximum in-sewer storage for each location (i.e. manhole), when considering only an individual and independent gate.

Figure 5 presents the individual in-sewer storage potential for each manhole location in the *Zona Central* catchment sewer network.
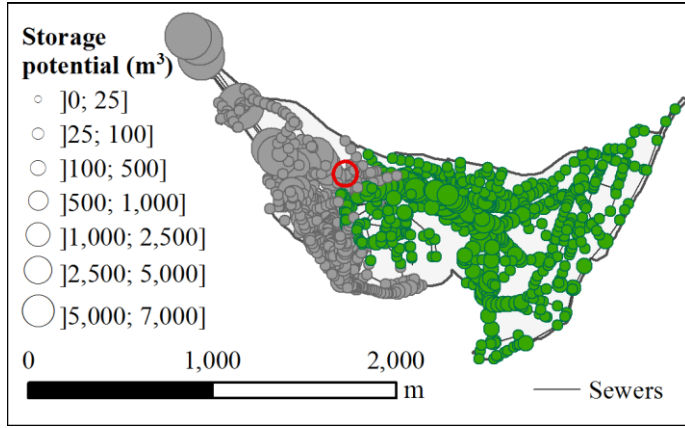
(a) Spatial distribution

(b) distribution of the storage potential in relation to the contributing runoff area

Figure 5. Potential in-sewer storage for each manhole, when considered independently. Grey dots represent manholes downstream of the flood-prone area and green dots represent manholes upstream the flood prone area marked with a red circle, with increasing dot size corresponding to increased storage potential.

For this specific case, it can be seen in Figure 5 that a large portion of the in-sewer storage potential is concentrated downstream of the flood-prone area. This is extremely relevant because only the storage potential located upstream the flood-prone area is effective for flood relief in this area. To further investigate the in-sewer storage potential one can also plot the relation between storage volume and contributing catchment area and assess if the storage is concentrated in a pipe or branch of the sewer system or if the storage potential is distributed over the entire network (Figure 5b).

4.3. Sensitivity of the best location to different reward functions

The values in Tables 4 and 5 show the in-sewer storage potential and the IDs of the manholes representing the best five locations obtained using the five reward functions (Section 2.4)

Table 4. *Zona Central* catchment - In-sewer storage (m^3) at the five best locations (manholes) according to the different reward functions, considering the location of the flood-prone area

Gate rank	Reward functions									
	RF1		RF2		RF3		RF4		RF5	
	Volume (m³)	Node ID	Volume (m³)	Node ID	Volume (m³)	Node ID	Volume (m³)	Node ID	Volume (m³)	Node ID
#1	606	159	568	515	606	159	14	206	66	116
#2	568	515	606	159	568	515	27	1_561	105	133
#3	350	157	271	1_668	350	157	29	113	27	363
#4	276	339	350	157	271	1_668	18	1040	18	1040
#5	271	1_668	188	313	276	339	35	144	17	369
TOTAL	2,070	-	1,983	-	2,070	-	123	-	233	-

Table 5. *D16&D17* catchments - Storage volume (m^3) at manhole (Node) ID of the five best flow control device locations according to the different reward functions, considering the flood-prone area

Gate rank	Reward functions									
	RF1		RF2		RF3		RF4		RF5	
	Volume (m³)	Node ID	Volume (m³)	Node ID	Volume (m³)	Node ID	Volume (m³)	Node ID	Volume (m³)	Node ID
#1	257	6	257	6	257	6	257	6	87	3
#2	88	7.2	59	7	88	7.2	36	7.4	257	6
#3	87	3	29	6.2	59	7	41	7.3	61	4
#4	75	5	41	7.3	29	6.2	4	8.1	75	5
#5	59	7	4	8.1	4	8.1	88	7.2	88	7.2
TOTAL	566	-	390	-	437	-	426	-	568	-

As can be seen in Table 4 (for the *Zona Central* catchment case study), only reward functions RF4 and RF5 are significantly different and, as expected, the variability in terms of the best locations slightly increases in lower rank positions. Based on this, the results obtained using the five reward functions follow the classification in two groups: (I) RF1, RF2 and RF3, and (II) RF4 and RF5. The ranking of group I is very similar (the same locations in the three top positions), leading to the same (or very similar) total in-sewer storage potential for the installation of five flow control devices. Since the reward functions use re-scaled values of the storage potential, contributing runoff area and number of contributing catchments, the available volume becomes the most relevant factor. For the reward function RF2, the

available volume is multiplied by a scalar (the re-scaled contributing runoff area) that is in this case positively correlated with the storage potential (see Figure 6a); this results from the drainage network design conventional practice. The results obtained with reward function RF3 are also very close to those obtained with reward function RF1. For this case study, the re-scaled contributing runoff area and the re-scaled number of contributing subcatchments are strongly linearly and positively correlated (Figure 6a), meaning that the area of the subcatchments does not change significantly and they cancel each other out.

(a) *Zona Central* catchment   (b) *D16&D17* catchments
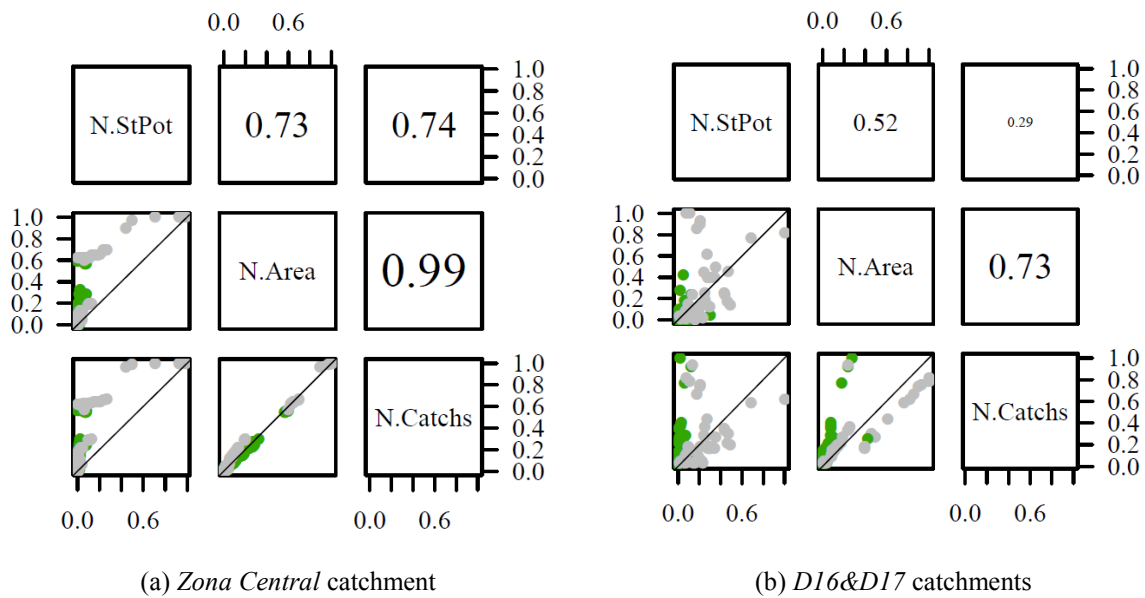
Figure 6. Correlation between the variables (re-scaled) used in the reward functions. Upper panel figures represent the degree of correlation between the variables. Grey and green dots represent candidate locations downstream and upstream of the flooding area, respectively. Correlation among the variables is much stronger for *Zona Central* catchment than for the *D16&D17* catchments case.

The results obtained using the reward functions of group II (RF4 and RF5) are markedly distinct. The total storage potential is approximately one order of magnitude smaller. Moreover, the results are significantly different between these two reward functions; although

the total potential storage is similar (Table 4), three out of the five locations are different. Using these two reward functions, the best locations are in the most upstream sectors of the network in which the contributing runoff area and the number of contributing catchments are small and therefore the potential impact of a flow control device failure is much smaller. This is clearly seen in Figure 7.

For the *D16&D17* catchments case study, the total in-sewer storage obtained using the five reward functions are all similar (see Table 5). This is a completely different result when compared to the results obtained for the *Zona Central* catchment. This confirms the intuition that the network characteristics play an important role in identifying the best locations for the flow control devices. Regarding the *D16&D17* catchments, the correlation between the re-scaled contributing runoff area and number of catchments is scattered in two distinct groups (Figure 6b). This is most likely because the *D16&D17* catchments case study comprises two almost independent sewer networks. By comparing Figures 6a and 6b, one can see that the correlation between the re-scaled contributing runoff area and re-scaled number of catchments is less strong for the *D16&D17* catchments case than in the *Zonal Central* catchment one. This, for example, explains the storage potential differences obtained between reward functions RF1 and RF3, and between RF4 and RF5, in the *D16&D17* catchment case (Table 5 and Figure 8). Some of the top five best locations are the same for the five reward functions results. This was not the case for the *Zona Central* case study, in which more options could be considered due to the larger size of the sewer network.
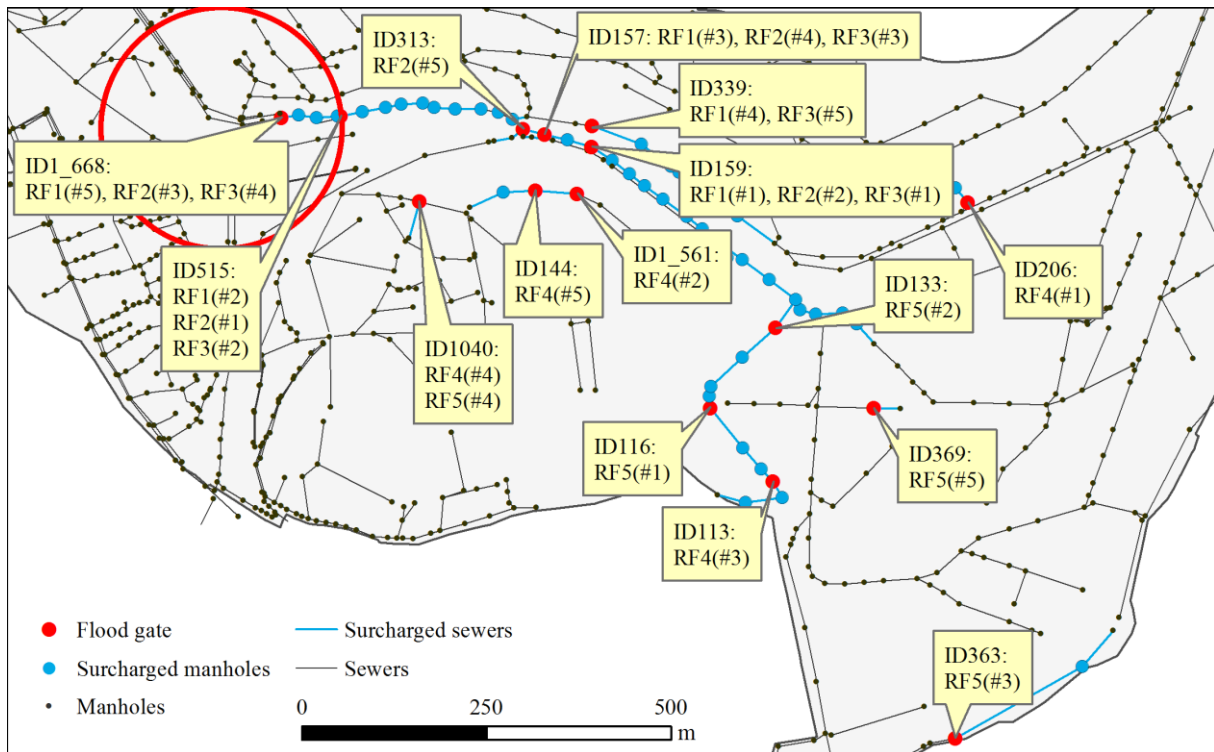
Figure 7. *Zona Central* catchment – best locations for in-sewer storage upstream the flood-prone area represented by the red circle
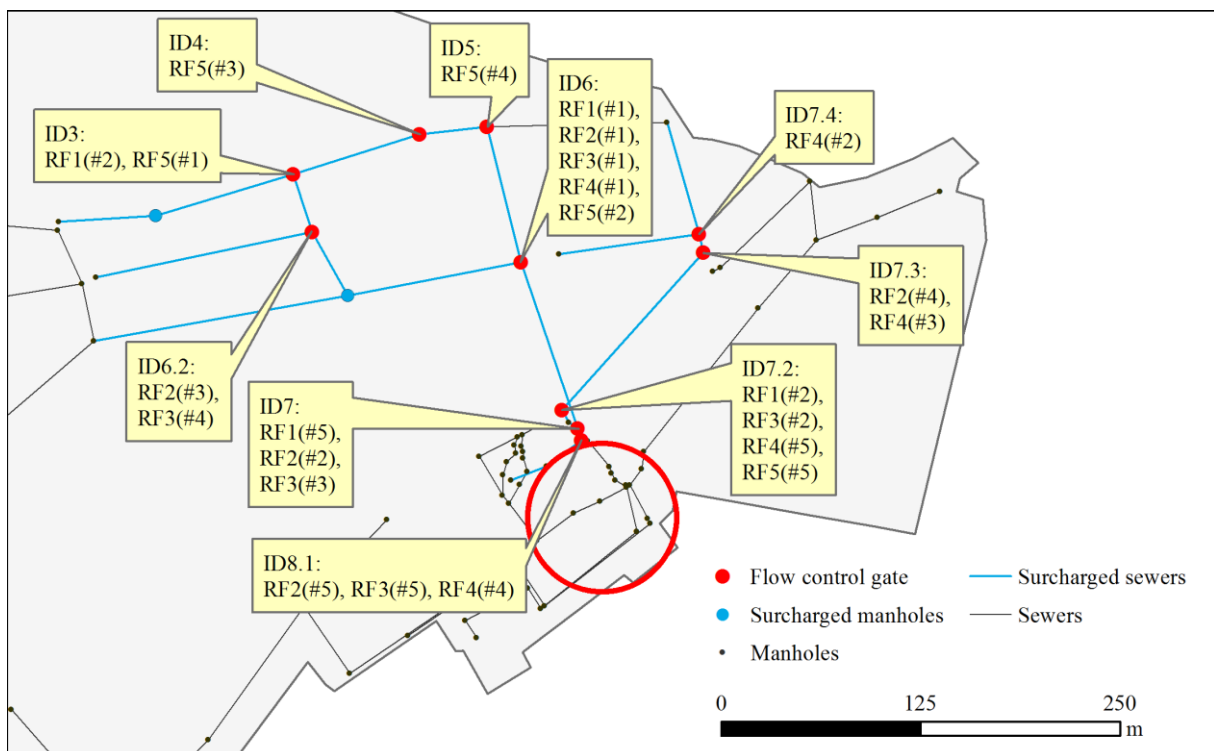
Figure 8. *D16&D17* catchments – best five locations for in-sewer storage upstream the flood-prone area represented by the red circle.

5. Conclusions

This study showed that in-sewer storage capacity in real sewer networks can be considerable. If local flow control devices are properly installed, in-sewer volume can be mobilised and become an interesting solution to mitigate floods and minimise CSO discharges. By using a screening tool such as the one presented in this study for rapid assessment of local in-sewer storage potential, the need for detailed hydraulic simulations can be avoided. The novelty of the proposed model includes the consideration of (i) different reward functions to identify the best locations to install flow control devices and (ii) a proxy to *steady state* flow conditions that best describe the system flow conditions when the flow control device(s) are not fully closed to estimate in-sewer storage potential. The reward functions implemented in the proposed methodology include not only terms that maximise the local storage potential, such as volume, but also terms that aim at reducing the impact of a flow control device failure and problems associated with large hydraulic forces. Upon detailed manual calculations and hydraulic simulations, the model performed well with a high degree of accuracy. This clearly shows that the consideration of the flow energy can have an impact on the estimation of the in-sewer storage potential, especially because the flow control device can operate partially closed.

The tests performed using two real networks showed that the choice of the reward function can have a large impact on flow control device location results. The differences in terms of storage volume for the best results using different reward functions can be as large as 93%.

Also, the analysis performed in the two case studies highlighted the fact that network topology and characteristics (e.g. sewer diameter and slope) have a strong influence in terms of the best flow control device location. The proposed location model can be used to support the decision of *where* to install flow control devices and also to determine *how many* flow control devices can or should be used to tackle a specific flood or CSO discharge problem. However, some limitations are worth noting and may be considered for future developments:

- The detail of the sewer network registry may have a strong influence on the results. As an example, the sewer network registries usually do not contemplate household connections which may affect in-sewer storage potential.

- Specific land uses (such as main roads) or areas defined as of difficult intervention (e.g. at the entrance of a hospital or in a main road) can hinder the installation of the flow control device where the in-sewer storage potential is maximised. Future work may consider the inclusion of this information in the methodology in order to make it more readily applicable and reliable.

It should be remembered that sewer network responses will change with time as a result of climate change and urbanisation, hence the benefit of any storage based solution is likely to decrease over time and so will need to be used in conjunction with other flood prevention or CSO discharge reduction measures. The local in-sewer storage solution is more adaptable to future conditions as additional local storage can be mobilised to maintain a specified flood risk level by installing additional flow control devices slowly over time as they are needed, rather than having to invest in a large scheme with a long design lifetime. These attributes make local in-sewer storage solution, and adaptable flood risk management measure, more desirable than traditional engineering solutions such as storage tanks which have less flexibility over the long term.

Acknowledgements

Author contributions. João P. Leitão designed the study, contributed to the development of the model, analysed the results and led the writing of the manuscript. Luís M. de Sousa contributed to the design of the model, developed and implemented the data model and software, analysed the results and contributed to the writing of the manuscript. Juan Pablo Carbajal contributed to the design of the model and to the manuscript writing. Nuno E. Simões and Alfeu Sá Marques contributed to the analysis, discussion and interpretation of the results. Jörg Rieckermann contributed to the design of the algorithm and the analysis of the results. All authors copy-edited the manuscript.

References

Baek, H., Ryu, J., Oh, J., Kim, T-H. (2015). Optimal design of multi-storage network for combined sewer overflow management using a diversity-guided, cyclic-networking particle

swarm optimizer – A case study in the Gunja subcatchment area, Korea. *Expert Systems with Applications*, 42(20), 6966-6975. doi: http://dx.doi.org/10.1016/j.eswa.2015.04.049.

Bauer, C. and King, G. (2006). *Java Persistance with Hibernate*. Dreamtech Press. ISBN: 817722719X, 9788177227192

Bennett, M. S., Mays, L. W. (1985). Optimal design of detention and drainage channel systems. *Journal of Water Resources Planning and Management*, 111(1), 99-112

Cunha, M. C., Zeferino, J. A., Simões, N. E., Saldarriaga, J. G. (2016). Optimal location and sizing of storage units in a drainage system. *Environmental Modelling & Software*, 83, 155-166. doi: http://dx.doi.org/10.1016/j.envsoft.2016.05.015

Doyle, J. R., Heany, J. P., Huber, W. C., Hasan, S. M. (1976). Efficient storage of urban storm water runoff. In *Proceedings of the Conference on Environmental Modeling and Simulation*, 139-143, 19-22 April, Cincinnati, OH, USA

Fletcher, T.D., Shuster, W., Hunt, W.F., Ashley, R., Butler, D., Arthur, S., Trowsdale, S., Barraud, S., Semadeni-Davies, A., Bertrand-Krajewski, J.-L., Mikkelsen, P.S., Rivard, G., Uhl, M., Dagenais, D., Viklander, M. (2015). SUDS, LID, BMPs, WSUD and more – The evolution and application of terminology surrounding urban drainage. *Urban Water Journal*, 12(7), 525-542. doi: 10.1080/1573062X.2014.916314

Fu, G., Khu, S-T., Butler, D. (2010). Optimal Distribution and Control of Storage Tank to Mitigate the Impact of New Developments on Receiving Water Quality. *Journal of Environmental Engineering*, 136(3), 335-342. doi: 10.1061/ASCEEE.1943-7870.0000161

Graser, A. (2013). *Learning QGIS 2.0*. Packt Publishing Ltd.

Holl, S., & Plum, H. (2009). PostGIS. *GeoInformatics*, 3(2009), 34-36

Houston, D., Werritty, A., Bassett, D., Geddes, A., Hoolachan, A., McMillan, M. (2011). *Pluvial (rain-related) flooding in urban areas: the invisible hazard*. Joseph Rowntree Foundation, York, UK. ISBN: 978-1-85935-863-4

IPCC (2013). *Annex I: Atlas of Global and Regional Climate Projections* [van Oldenborgh, G.J., M. Collins, J. Arblaster, J.H. Christensen, J. Marotzke, S.B. Power, M. Rummukainen and T. Zhou (eds.)]. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Kuczera, G., Kavetski, D., Franks, S., Thyer, M. (2006). Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *Journal of Hydrology*, 331(1-2), 161–177. doi: 10.1016/j.jhydrol.2006.05.010

Leitão, J. P. (2009). *Enhancement of Digital Elevation Models and Overland Flow Path Delineation Methods for Advanced Urban Flood Modelling*. PhD Thesis, Imperial College London, London, UK

Lim, S., Ho, V. H., Lee, S. Y., Yoo, D.G., Kim, J.H. (2014). Determination of optimal location and capacity of detention facilities. *Procedia Engineering*, 70, 1037-1045. doi: http://dx.doi.org/10.1016/j.proeng.2014.02.115

Maluf, L., Shepherd, W., Ostojin, S., Simões, N., Sá Marques, A., Mounce, S., Skipworth, P., Leitão, J.P. (accepted). Real time flow control to utilise existing in-sewer storage. In *14th International Conferences on Urban Drainage*, 10-15 September, Prague, Czech Republic

Oxley, R. L., Mays, L. W. (2014). Optimization – Simulation Model for Detention Basin System Design. *Water Resources Management*, 28, 1157–1171. doi: 10.1007/s11269-014-0552-z

Perez-Pedini, C., Limbrunner, J. F., Vogel, R. M. (2005). Optimal Location of Infiltration-Based Best Management Practices for Storm Water Management. *Journal of Water Resources Planning and Management*, 131(6), 441-448. doi: 10.1061/ASCE0733-94962005131:6441

Philippon, V., Riechel, M., Stapf, M., Sonnenberg, H., Schütze, M., Pawlowsky-Reusing, E., Rouault, P. (2015). How to find suitable locations for in-sewer storage? Test on a combined sewer catchment in Berlin. In *10th International Conference on Urban Drainage Modelling*, Quebec, Canada

Quentin B. Travis, Q. B., Mays, L. W. (2008). Optimizing Retention Basin Networks. *Journal of Water Resources Planning and Management*, 134(5), 432-439. doi: 10.1061/ASCE0733-94962008134:5432

Rossman, L. (2015). *Storm Water Management Model User's Manual Version 5.1*. U.S. Environmental Protection Agency, Cincinnati, OH, USA

Sá Marques; A., Simões; N., Maluf; Milad, L. Jajarmizadeh, M. (2016). *Initial strategy for CENTAUR system deployment in Coimbra*. CENTAUR project deliverable. Freely available at:

Shepherd, W., Mounce, S. R., Ostojin, S., Abdel-Aal, M., Schellart, A., Skipworth, P., Tait, S. (accepted). Optimising a Fuzzy Logic Real-Time Control System for Sewer Flooding Reduction using a Genetic Algorithm. In *15th International Computing & Control for the Water Industry Conference*, 5-7 September, Sheffield, UK

Stonebraker, M. and Kemnitz, G. (1991). The POSTGRES next generation database management system. *Communications of the ACM*, 34(10), 78-92

United Nations, Department of Economic and Social Affairs, Population Division (2014). *World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)*. United Nations (available at: http://esa.un.org/unpd/wup/Publications/Files/WUP2014-Highlights.pdf)
[Accessed: 19 May 2016]

Zhen, X-Y., Yu, S. L., Lin, J-Y. (2004). Optimal Location and Sizing of Stormwater Basins at Watershed Scale. Journal of Water Resources Planning and Management, 130(4), 339-347. doi: 10.1061/(ASCE)0733-9496(2004)130:4(339)

Appendix A. Data management of the flow control device location model

The developed flow control location model is based on a geospatial database that takes as input a text file describing a sewer system network in a tabular form, including various types of elements and their geo-spatial location. This tabular formulation is prone to redundancies and cannot guarantee referential integrity between the various information entities it describes. The first implementation step was thus the translation from this tabular data into a formal and normalised data model. The current version of the proposed model takes as input a text file with a structure similar to the SWMM model (Rossman, 2015) *.inp* file; nevertheless, other model and GIS formats can, in the future, be imported into the database.

Figure A.1 presents an overview of the data model used. At its core, this model defines a directed graph, composed by elements of types *Node* or *Link.* Each *Link* element refers to two nodes: one from where it departs and a second where it ends - in the hydraulic context they are referred as upstream and downstream nodes. Neither of these types exist in the *.inp* file, and introduce a higher level of abstracting that later on simplifies computation.

Three different entities of those described in the *.inp* file are identified with the *Node* type: *Outfall*, *Junction* and *Storage*. *Outfall* represents ending points of the network, where rainwater is returned to the natural environment. *Junction* is a point between two conduits, usually accessible through a manhole; starting points of the network are also of this type.
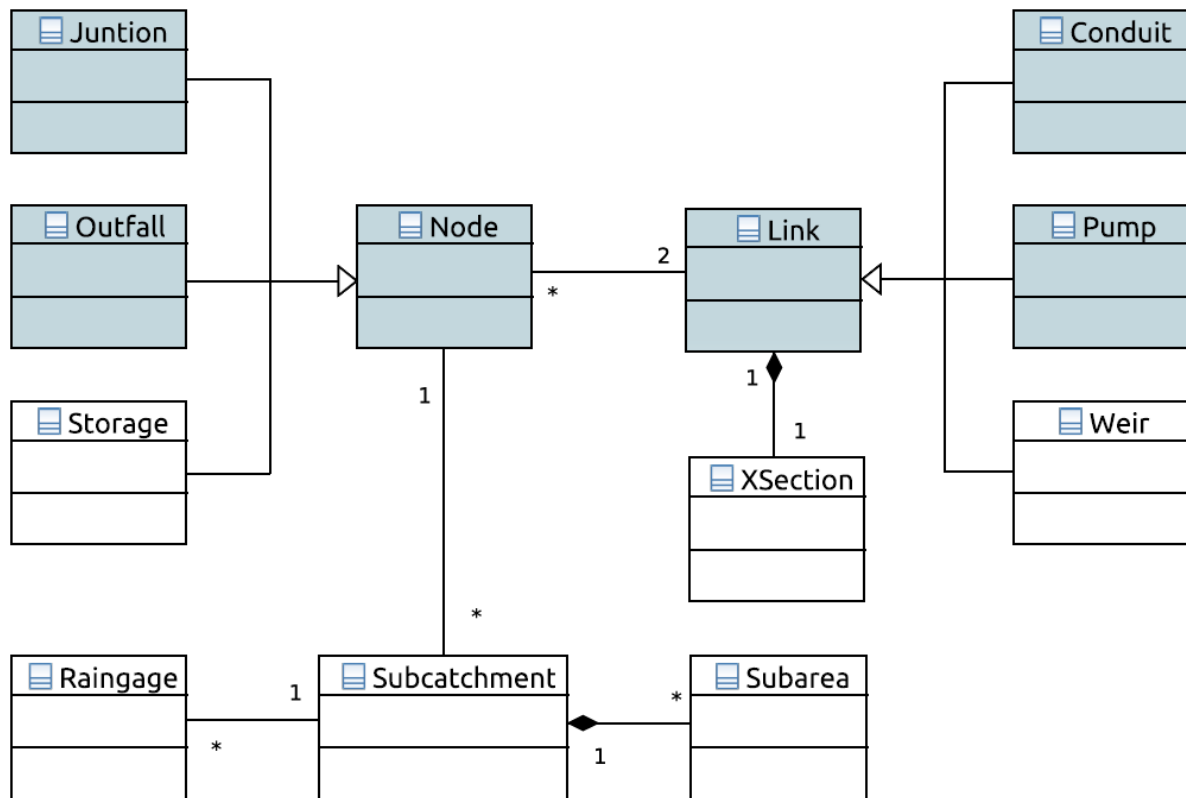
Figure A.1. UML diagram of the data model.

Regarding the *Link* class, three entities from the *.inp* format text file are assigned to it: *Conduit*, *Pump* and *Weir*. it The *Conduit* is a hollow structure that transports flow from one node in the network to another; each one has a specific cross section geometry, described in the *XSection* type. A *Pump* also links two network nodes and transfers mechanical energy to the flow (hydraulic energy). The *Weir* is a flow control structure that also links two nodes of the network.

Another relevant entity present in the *.inp* file is the *Subcatchment*. It defines the area of the surface that drains to a particular node in the network. The *Subcatchment* may be divided in various *Subareas*. Each *Subcatchment* is connected to one, and only one, node in the network.

This data model was implemented in a relational database supported by the PostgreSQL database management system (Stonebraker and Kemnitz, 2006). A programme was developed in the Java language that reads the *.inp* file and imports its contents to the database, using the Java Hibernate technology (Holl and Plum, 2009). It normalises the data, converting it into the model described above, creating all the necessary relationships.

The *.inp* file contains specific tables with cartographic coordinate pairs for each of the geo-spatial elements, these being points for *Node* type elements and polygons for *Subcatchment* elements. This information is converted into formal geo-spatial objects and stored in the database as a geometry type field in the respective table. To this end is employed PostGIS (Holl and Plum, 2009), the geo-spatial enablement extension for PostgreSQL.

Appendix B. Hydraulic simulation results

In order to compare the results of the proposed location model to identify the best locations to install in-sewer flow control devices, a set of hydraulic simulations were conducted using the EPA SWMM (Rossman, 2015). The hydraulic profiles obtained are presented in Figures B.1 and B.2 for the case study in Coimbra (Zona Central catchment) and in Alcântara, Lisbon (D16&D17 catchments.
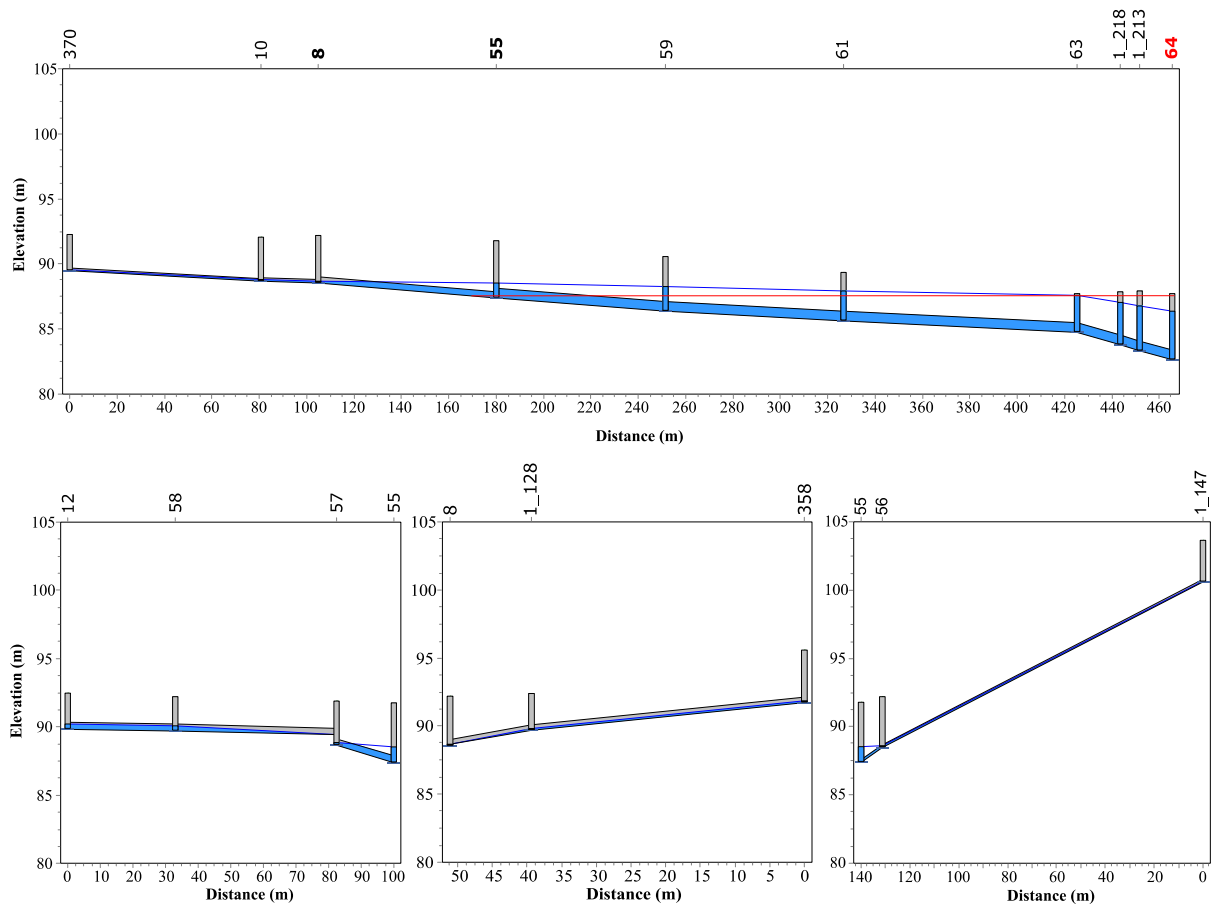

Figure B1. *Zona Central* catchment. Results obtained with the hydrodynamic simulations conducted using EPA SWMM model (blue line represents the hydraulic grade line obtained using the hydraulic simulation; red line represents the hydraulic grade line for the *static* assumption)
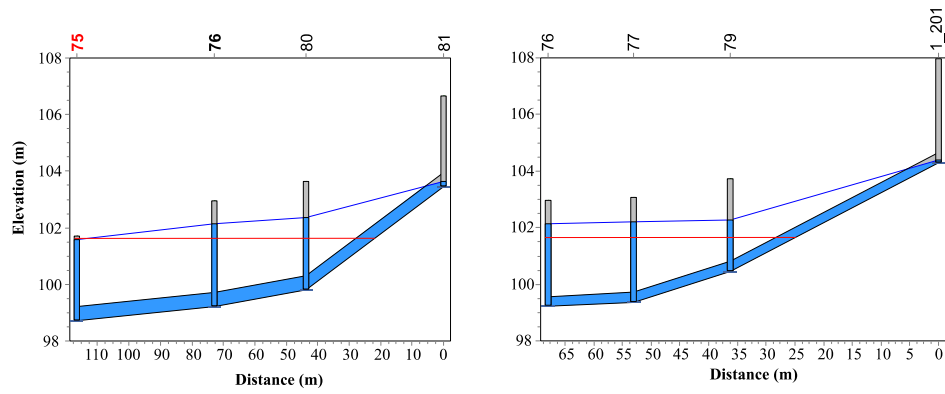
Figure B2. *D16&D17* catchments. Results obtained with the hydrodynamic simulations conducted using EPA SWMM model (blue line represents the hydraulic grade line obtained using the hydraulic simulation; red line represents the hydraulic grade line for the *static* assumption).