An integrated software architecture for the pollution-based real-time control of urban drainage systems
Luis Romero, Bernat Joseph-Duran, Congcong Sun, Jordi Meseguer, Gabriela Cembrano, Ramón Guasch, Montse Martínez, Eduard Muñoz and Vicenç Puig

ABSTRACT
This paper presents a complete methodology for the development of an integrated software architecture, which can achieve a closed-loop application between the integrated real-time control (RTC) and a virtual reality simulation for the urban drainage system (UDS). Quality measurements are considered during the simulation and optimization process. Model predictive control (MPC) and rule-based control (RBC) are the two main RTC methods embedded in this architecture. The proposed integration environment allows the different software components to efficiently and effectively communicate and work in a system-wide way, as well as to execute all the necessary steps regarding input parameter management, scenario configuration and results extraction. The proposed approaches are implemented into a pilot based on the Badalona UDS (Spain). Results from different scenarios with individual control approaches and rain episodes are evaluated and discussed.

Key words | key performance indicators, pollution-based control, software architecture, urban drainage system, water bodies protection

HIGHLIGHTS
- Closed-loop framework between the urban drainage system simulator and the real-time control module.
- Integration of the wastewater treatment plant state in the control strategy.
- Software architecture to handle operations and communication among specialized commercial software.
- Both hydraulic and quality measurements are considered during the control process.
- Implementation of the architecture to a case based on a real network.

INTRODUCTION
Combined urban drainage systems (UDS) have a fundamental role in collecting stormwater and wastewater together and conveying them to wastewater treatment plants (WWTPs) to minimize the associated pollution of the receiving environment (Butler et al. 2018). Their importance in developed areas lies in the influence over the modern society and the natural water cycle. The increasing urbanization, installation of complex infrastructure and frequent storm weather cause floodings and combined sewer overflows (CSO) due to the inability to manage high-intensity rainfalls, hence seriously polluting the receiving water...
bodies (Cembrano et al. 2004). New water detention and diversion infrastructure is typically installed in order to confront these challenges, but advanced control approaches are still required to proficiently operate the UDS.

Real-time control (RTC) is considered as an appropriate option among the different possible control strategies, iteratively computing the optimal control set-points through real-time measurements or even predicted values while benefiting from the development of Information and Communication Techniques (ICTs; Schüetze et al. 2004). According to the recent literature, there are numerous and different kinds of RTC approaches applied to the UDS, which can be generally classified into model-based strategies and knowledge-based algorithms. Model predictive control (MPC; Joseph-Duran et al. 2015) and linear-quadratic regulator (LQR; Lemos & Pinto 2012) are the representative of model-based approaches which require a mathematical model of the system behaviour, together with an objective function and boundary constraints, to produce the optimal strategy. Knowledge-based algorithms do not require a model, but a complete expertise about the network characteristics. One example of this technique is rule-based control (RBC), which considers different scenarios that may happen during the system operation by means of the assessment of if-else conditions and applies on-line predefined rules to generate control actions (Aulinas et al. 2011).

In order to validate these advanced control approaches for their application in a real system, they may be tested in conjunction with a high-fidelity simulator. This term is utilized in this work as presented in Lund et al. (2018). It behaves as virtual reality, playing the role of the sewer network (SN) and the WWTP. Typically, the implementation of a closed-loop software architecture (CLSA) for UDS is executed using a single software platform (Achleitner et al. 2007), limiting the utilization of powerful commercial tools. Most of the existing integrated frameworks only focus on the hydraulic model, so quality dynamics are not included into the optimization and simulation processes. Moreover, the integration problem is usually faced through the literature by means of the development of specialized software tools like SYNOPSIS (Butler & Schütze 2005), OpenMI (Gregersen et al. 2007), DynaMind (Urich et al. 2012) and CityDrain3 (Burger et al. 2016).

This paper proposes a methodology for the development of an integrated software architecture with the aim of achieving a closed-loop application between the RTC module and the wastewater system. The platform is prepared to allow the integration of the WWTP state in the control approaches, and therefore, different individual models (Bach et al. 2014) must be coordinated. Besides, quality measurements (Sun et al. 2020) can be considered into the closed-loop. In Rauch et al. (2005), the need for integration to cope with water-quality-based approaches is remarked due to the importance of identifying the most effective action in the UDS to solve any issue affecting the receiving water body properties.

Hence, the development of the integrated multi-software pollution-based control architecture presented in this paper entails a further step on the path to emulate the particular behaviour of a complex water system under the influence of a certain control approach, with sufficient reliability to be capable of extracting firm conclusions from the achieved results. In addition, the effects of including quality and integrating the WWTP in the control operation can be evaluated and this information can be employed as a decision-making tool.

**METHODS**

The general software architecture of an integrated closed-loop framework for UDS is presented in Figure 1.

The closed-loop coordination environment enables the communication between the integrated model simulator and the RTC controller. Therefore, the SN and WWTP simulation results are gathered, extracting the necessary data for the control algorithm operation. The RTC software computes the next required actuators set-points, which are, in turn, applied to the simulators for the next time instant.

![Figure 1 | General scheme of the CLSA framework.](Figure%201%20|%20General%20scheme%20of%20the%20CLSA%20framework.)
In this section, a detailed explanation of the different software constituents of the CLSA framework is provided, introducing the essential aspects defining each component and its operation. Those individual functionalities are linked with their role inside the closed-loop structure by explaining the coordination architecture. Therefore, a complete and hierarchic view of the proposed solution will be presented.

**Integrated model simulator**

The integrated model simulator consists of a set of individual high-fidelity models, representing different components of a complete UDS, whose operations are mutually influenced, as presented in Schütze et al. (1999). In the case of this paper, the integrated model is considered to be composed of two parts: a model to emulate the complex dynamics of an SN and another operating as a sufficiently convenient replacement of the real WWTP.

**SN simulator**

As mentioned earlier, an SN simulator consists of a software module that implements a high-fidelity physically based model of the network dynamics, in the form of a set of partial differential equations. The involved hydraulic variables in those expressions may include the sewer flow rates and velocities, pressure heads, surcharge level, as well as the water depths. Besides, in this work, quality measurements are also considered in order to obtain optimal pollution-based performance. These models emulate both hydraulic and quality behaviour of the SN, through routing rain scenarios with empirical surface runoff expressions and solving the related equations (Rubinato et al. 2013). A comparison among different water network simulators is presented in Bach et al. (2014).

The majority of software platforms utilized for this purpose gather a particular set of functionalities apart from the network dynamics simulation:

- **Network generation**: SN simulation packages normally include a graphical user interface (GUI) which allow the user to design a new network or represent an existing one by placing the necessary components (e.g. sewers, nodes, pumps, gates and catchments, etc.). Besides that, property details for these components can be configured due to the requirements of the SN structure.
- **Rainfall definition**: In the case of combined drainage networks, the water input to the system comes from two main sources: wastewater and stormwater. Regarding the fact that SN control approaches evaluation relies on the performance analysis in demanding scenarios, primarily characterized by intense and/or long-lasting rainfall events, the proper definition of those rain episodes becomes a capital task. Therefore, the precipitations configuration must be considered as a prior stage of any project dealing with UDS, generating a large database of real or fictitious scenarios with enough diversity to be capable of carrying out a proper evaluation of the control strategy operation.
- **Simulation configuration**: Several parameters must be set to achieve correct and effective simulations. Depending on the selected software, the amount and complexity of the settings may vary, though various of them need to be available at any simulator: simulation date and time, duration, rain event, build-up time (number of hours/days, depending on the simulation software units parameterization, from the previous rain event end to the starting point of the current simulation), numerical solver settings, wastewater profile (complementing the stormwater to supply the SN model with the total water input), control mode (apart from receiving external regulator set-points, the virtual drainage structure may employ local rule-based approaches to manage the actuators operation), hotstart file (record of a certain state of the virtual network, commonly used to feed the model as an initial condition), among others.
- **Results extraction**: It is a mandatory feature regarding the necessity of supplying several measurements of the virtual network state to the external controller (unless an internal local management approach is selected). Depending on the control strategy, the required input data may differ, but tank and catchment volumes and/or pollutants mass, certain sewers flow or water quality and water level at critical points are typically needed. Moreover, these results are also indispensable to perform an analysis of the virtual network representation resemblance to the modelled real system.
**WWTP simulator**

To consider an integrated control strategy, the presence of a WWTP model inside the simulation framework is essential. The physical, chemical and biological degradation processes that treat the pollutants contained in the water reaching the treatment plant are represented, and operation results are gathered to get information about the benefits of exploiting the integration in the considered control approach. Moreover, the collected data allows us to compare the performance regarding quality measurements between pollution-based and volume-based controllers.

The WWTP simulation software packages allow the inclusion of several operational blocks representing well-known processes like pre-treatment, clarifiers, gravity thickeners or bio-reactors (each one of them defined by a set of equations) (Jeppsson 1996); whose connection and configuration implies the implementation of the WWTP model.

As in the case of the SN simulation software, several general settings are mandatory for the WWTP operation, e.g. date, time, duration, hoststart information, among others. Moreover, depending on the designed WWTP type and its composing elements, various parameters regarding biological aspects of the plant functioning are required. If the SN simulator does not provide the necessary information as a result, an intermediate process must be designed to generate the parameters for the virtual treatment plant from the available pollutants information at the SN simulation results.

Regarding the integration implementation, a certain scope for action at the WWTP model performance is necessary. The degree of freedom in the treatment facility management from an external source is a design aspect for the closed-loop framework.

**Real-time controller (RTC)**

Through the UDS state produced by the integrated model simulator, the RTC module must be capable of computing the set-points for the different actuators.

As previously mentioned, there are several control strategies that can be implemented to accomplish this set-point computation. Thus, the approach selection ought to consider numerous aspects like its associated economic/technological cost and the network characteristics and complexity, highlighting the control degrees of freedom due to the number of accessible actuators at the system.

In addition, the RTC module must include a component whose function consists of allowing the integration of the WWTP with the control strategy, so that the treatment facility actual state is taken into account during the UDS management operation.

**Control algorithm**

Regarding the control strategy, model-based and knowledge-based approaches will be considered. In this work, about the former, an algorithm based on MPC is utilized as an example (it is extensively described in Sun et al. (2020)); whereas regarding the latter, a rule-based scheme referred to as RBC is introduced. A review of these and other approaches can be found in García et al. (2015). In both cases, hydraulic and quality dynamics are considered into the RTC controllers.

The distinctive features of these two control approaches will be presented, stressing issues like necessary input information or required software resources, as well as highlighting their advantages and drawbacks from the integrated UDS closed-loop architecture point of view.

**Model predictive control (MPC).** MPC is a model-based control strategy that employs the prediction of the system response to compute a set of suitable future control actions for a certain time horizon, using a simplified mathematical model of the system dynamics and an optimization process to minimize a certain cost function $J$. Its definition is essential and mainly includes an analysis of the different desired goals, like minimization of the floodings and/or CSO spilled to the water receiving bodies (in both quantity and quality), the WWTP efficient usage, safe and smooth operation of SN and WWTP, etc. All these objectives are represented by a term in the cost function, where each element is multiplied by its associated weight. The non-linear optimization problem
associated with the MPC can be expressed as follows:

\[
\begin{align*}
\text{minimize} \quad & J(x, u, w) \\
\text{s.t.} \quad & x(t) = x_0; \\
& x(k + 1) = f(x(k), u(k), w(k)), \quad k = t, \ldots, t + H - 1; \\
& h(x(k), u(k), w(k)) = 0, \quad k = t, \ldots, t + H; \\
& g(x(k), u(k), w(k)) \leq 0, \quad k = t, \ldots, t + H; \\
& x_{\text{min}} \leq x(k) \leq x_{\text{max}}, \quad k = t, \ldots, t + H; \\
& u_{\text{min}} \leq u(k) \leq u_{\text{max}}, \quad k = t, \ldots, t + H;
\end{align*}
\]

where \(x(t)\) corresponds to system states at time step \(t\), typically representing water volume and pollutants mass in tanks; \(u(t)\) is the vector of control actions at time step \(t\) (it depends on the studied network) and \(w(t)\) represents the disturbances at time step \(t\) (normally related to rain intensity and runoff). Function \(f()\) mainly includes mass and volume balance equations, while \(h()\) and \(g()\) represent the general constraints of the MPC problem, and \(u_{\text{min}}, u_{\text{max}}, x_{\text{min}}\) and \(x_{\text{max}}\) are the control actions and states physical limits. Constraints are given by the capacity of tanks/WWTP and flow limits at pipes/actuators. Finally, \(k\) is an index that represents time, which goes from the current time instant \(t\) to \(t + H\) (or \(t + H - 1\) considering \(f()\)), where \(H\) is the optimization horizon.

Therefore, the utilization of an optimization specialized software would be necessary if the MPC approach is considered, and a simplified model of the SN should be derived in order to implement the modelling equations at the optimizer. These conceptual models involve network components like nodes, pipes and tanks, as well as the evolution of their associated variables, like flow, total suspended solids (TSS; it is considered as a pollution representative in this paper; see Woodward & Curran (2006) for more information about this quality indicator), volume and mass. They mainly consider the flow and mass balances at the nodes and the tank volume and mass evolutions. Other elements like weirs may need custom models for the considered SN.

The set-points for the network actuators are derived from the optimization results. The achieved data for a certain set of sewers at the simplified model, at the first time step in the horizon, is evaluated by means of a flow-to-setpoint function. It allows the conversion of the corresponding flows into reference values for the position of the actuators. This function depends on the network characteristics and it must be meticulously designed.

Besides, as the optimization scheme employs the SN simplified model to emulate its dynamics by means of the implemented equations, information regarding the water inputs to the network is required. This knowledge may come from a conversion function that transforms the rain forecast into the flow reaching the network by each one of its basins.

**Rule-based control (RBC).** RBC consists of a decision-making approach that exploits gathered knowledge about system features to generate a certain set of rules that defines the cause–effect relation for a specific problem. The obtained set-point explicitly depends on the fulfillment of a settled group of conditions involving variables directly measured at the considered system.

In the case of UDS, variables (involving tanks level or volume, water level at critical points for flooding prevention, quality measurements at certain sewers or nodes, CSOs, etc., as well as WWTP inflow capacity in integrated strategies) must be evaluated to generate proper regulator values affecting the different system actuators.

As discussed in Aulinas et al. (2011), this system knowledge is usually converted into if-else structures like decision trees, although this implementation presents some disadvantages for complex networks, like less flexibility and lack of reasoning capabilities. Other approaches which are more suitable for complex water systems rely on fuzzy logic-based schemes as presented in García et al. (2015).

**WWTP inflow capacity calculator**

The information about the WWTP state is crucial for any considered regulator generation approach within an integrated architecture, since the selected control strategy must take this knowledge into consideration when deriving the controller set-points for the next time instant.
There exist different approaches to improve the WWTP operation under demanding conditions:

1. Increasing the WWTP inflow depending on its capacity instead of using a maximum inflow considered for the worst-case scenario (Müller & Krauth 1998).

2. Increasing the WWTP inflow, but bypassing the extra flow to be introduced into the secondary clarifier (Ahnert et al. 2008).

The first strategy is selected due to the larger leeway for action considering the integration of the WWTP state during the RTC operation, since the second approach considers WWTP internal procedures which are beyond the scope of this article. Thus, an inflow capacity computation module, henceforth referred to as a capacity calculator, must provide the plant state from measurements produced by the virtual reality UDS: input water flow and quality results from the SN simulator, as well as the current state of the WWTP, which is derived from the information achieved by the WWTP simulator. This maximum capacity applies to the plant entrance, but it may be computed considering the primary or secondary clarifier state, depending on the control objective, plant management goal and the facility layout. Therefore, the manner to design this WWTP inflow capacity calculator may vary among different projects.

Finally, this capacity value is provided to the control algorithm, so the way it handles this information about the WWTP state may differ among distinct RTC strategies:

- Considering MPC as the selected control approach, the WWTP inflow capacity knowledge must be utilized to compute an objective function term that increases the penalty as the difference between actual WWTP inflow and the mentioned capacity augments.

- Regarding RBC, the WWTP information may be used to evaluate a condition or for the explicit computation of a certain actuator set-point. The integration scheme would depend on the RBC algorithm definition.

**Closed-loop coordination environment**

To develop a complete closed-loop framework from the different presented ingredients, a binding element must be employed to host the intra-application operations and accomplish the necessary inter-application functionalities. The scheme of the complete architecture is presented in Figure 2.

The closed-loop approach implies the coordination of the different functionalities forming the framework. The SN simulation is configured, considering the event date, rain forecast and network initial state as internal configurations settings of the SN simulator, and the actuators set-points as external inputs to the simulation software. Then, this simulation is executed and the achieved results are extracted.

The information about the water quality, required for the WWTP simulation, is obtained from a quality transformation process (referred to as Quality Parameters Transformation in Figure 2). It is supplied with the gathered hydraulic and hydrologic results from the SN simulation and generates the required chemical and biological parameters for the WWTP simulator. The inflow capacity of the WWTP, used by its simulator, is computed by a capacity calculator (WWTP Capacity Calculator in Figure 2) from the previous treatment plant state and current SN information. After the WWTP detailed simulator launching, the resulting data can be gathered to compute the inflow capacity in the next iteration. Finally, the RTC algorithm must be fed with the required SN results such as flows, quality and initial volumes; the capacity for the WWTP and, depending on the control approach, a rain forecast and previous control module results.

The outcome of running the controller is a set of data that has to be converted from SN flows to actuator set-
points to configure the next step SN simulation. The selection of the coordination environment highly depends on the software elements performing the presented individual tasks:

- It must hold complete compatibility with all the applications, to carry out their launching and pull off the required communication demands.
- Files management involving numerous formats is indispensable to fulfil data inputting, software configuring and results extraction/presentation assignments.
- The importance of elaborating a closed application varies according to the purpose of the CLSA framework development, e.g. academic research and proprietary software; and the set of possible coordination software environments may be reduced if considering a professional application. However, a simpler environment may be enough for more relaxed development demands, so that the involved tasks are accomplished without including application layers to the solution.

However, the main tasks involving the closed-loop execution are conceptually identical regardless of the selected binding platform: virtual reality simulation preparation and control approach data feeding, as is represented in Figure 1.

Different programming languages like MATLAB® (Joseph-Duran et al. 2015), Python (Urich et al. 2012), as well as combinations of both of them (Riaño-Briceño et al. 2016), have been considered through the literature for this moderation task.

CASE STUDY

The case study consists of a realistic pilot based on the Badalona UDS. It considers a combined SN (represented in Figure 3 by the area that is highlighted in blue) and a WWTP that treats the collected water and releases it to the Mediterranean Sea. There is a single detention tank in this drainage network (DLES in Figure 3), which stands...
out due to its importance considering the control problem. It is conceived to be voluminous enough to be capable of fully storing the vast majority of the water coming from rainfalls. Additional information about the case study pilot characteristics, structure and location can be found in Sun et al. (2020).

When high-intensity rainfall occurs, the stormwater together with the sewage water may exceed the WWTP inflow capacity, so polluted water would reach the receiving body by means of the CSO, economically, environmentally and socially harming the area.

The local control strategy is based on two operative modes, both implementing a rule-based system that checks several SN hydraulic measurements. The anti-flooding approach is utilized as the main operation mode, basing its functioning on the water level at the most critical point of the network. Besides, an anti-CSO approach is also implemented, but due to its simplicity, the scope for action is large.

In order to improve the currently implemented local control strategy in the Badalona UDS, three different control approaches using the proposed CLSA are developed:

1. Volume-Based Model Predictive Control (VBMPC): Model-based strategy characterized by the main role of the CSO volume spilled to the environment during the optimization process. Despite the WWTP integration is possible due to the utilized architecture, the feature is not exploited by this approach for comparison with the next strategy.

2. Pollution-Based Model Predictive Control (PBMPC): The only differences with respect to the previous approach are the inclusion of the quality (TSS mass) during the optimization process and the integration with the WWTP.

3. Rule-Based Control (RBC): A set of simple rules is defined with the aid of various decision trees. They consider both hydraulic and quality features of the network state as well as the treatment plant inflow capacity, hence entailing a further step with respect to the local control approach. Its development is based on the gathered knowledge about the UDS characteristics during the development of the previous control approaches.

Therefore, the CLSA framework presented in this work must be implemented to be capable of applying the desired control strategies and assess their potential benefits. The coordination application that hosts the activities of the different software components has been programmed in Ruby, due to the facility of connection of this programming language with certain exploited commercial tools.

Following the structure of Section ‘Methods’, the individual software elements composed of the closed-loop framework for the case study will be presented; closing the exposition by defining the binding platform used to synchronize the previous elements.

Integrated model simulator

The elaboration, configuration and calibration of the integrated virtual reality simulator based on the Badalona UDS were developed as part of the LIFE EFFIDRAIN (LIFE14 ENV/ES/000860) project. The description of the integrated model simulator must be divided into its two different composing ingredients.

SN simulator

The demonstration SN, used as virtual reality for the simulation of the hydraulic and hydrology dynamics of the UDS, was developed using InfoWorks Integrated Catchment Modelling (or simply InfoWorks ICM).

The SN hydraulic and sediment transport dynamics are based on the 1D Saint Venant equations and the Velikanov model (Zug et al. 1998), respectively. The virtual network includes 14,280 nodes and 15,055 sewers approximately, together with 4 gate elements to control the amount of water entering the tank (there is a pair of InfoWorks ICM gates for each gate at the real network, as they convey water by two possible paths: to the tank or down to the network), and a pump for its emptying. The detention tank is included among the sewers, due to the impossibility of representing it as a node due to problems with the sedimentation phenomena.

Both the gates and the pump, which compose the SN complete actuators set, are managed by means of a regulator whose origin may be internal to InfoWorks ICM, via its local control scheme, or external, coming from the approach selected for the CLSA framework.
The simulation software presents the mentioned functionalities in Section ‘Methods’, Subsection ‘SN simulator’:

- Network generation at InfoWorks ICM: As previously commented, the definition of the virtual SN is part of a larger project that started with simpler network versions, even using a different SN modelling software, and ended with the latest improvements considering new data for calibration, gathered from Supervisory Control And Data Acquisition (SCADA) systems and measuring devices installed at the real network in the frame of external projects. Therefore, a deep explanation of this task goes beyond the scope of this paper.

- Rainfall events definition at InfoWorks ICM: To generate the stormwater information, historical data regarding rain scenarios that occurred during the last years at Badalona were employed. This knowledge can be converted into rainfall databases which are accessible from the InfoWorks ICM user interface. Then, during the simulation configuration, the selection of a certain precipitation dataset allows the user to assign the correct rainfall pattern to the particular simulated event. In the Badalona-based pilot, the alert criteria to classify the distinct rain scenarios, established by the city hall, was utilized to select a set of episodes ranging from small to large rainfalls.

- Simulation configuration at InfoWorks ICM: Events are prepared to be simulated once the previously commented aspects of the SN definition have been generated (network, rain database, wastewater profiles, etc.), as well as others like the time step duration, corresponding to 5 min for every simulation of the case study. The configuration process can be carried out manually, by means of the InfoWorks ICM GUI, as well as via external source, which is the main interest for this paper.

- Results obtained from InfoWorks ICM: It provides numerous sources of information about effectuated simulations regarding hydraulic and quality results, as well as computation performance. The simulation results files, generated in .csv format, present node and sewer-related variables like water depth (or height), flow rate and velocity, TSS, volume, pressure head and sedimentation depth.

**WWTP simulator**

The WWTP detailed model is implemented using the commercial software GPS-X, which allows the mathematical modelling, control and management of WWTPs. GPS-X provides a large suite of tools that allow generation of complex treatment plant models, run simulations and analyze the results.

The design, creation and implementation of the model of the plant consist of a prior task with respect to the purpose of this paper. In this case study, the Besòs WWTP was employed as a reference for the detailed virtual reality plant, including the same processes but scaling them considering the input flow to the designed model to be 14% of the real one (this resizing of the plant is required because the real WWTP treats water from different locations, whereas the case study only considers the Badalona area, which represents a 14% of the total). The modelled plant counts on a simplified sludge line without biosolid treatment. The water line is formed by a gravity clarifier as primary treatment, an aerated reactor and a conventional decantation as secondary treatment; and a biological treatment based on the Activated Sludge Model ASM1 (Henze et al. 2000).

Apart from various simulation settings which depend on the selected episode, different input parameters are required depending on the designed GPS-X model characterization, being related to the biological processes executed in the treatment operation. Regarding the Badalona-based pilot, the selection of the ASM1 model implies the utilization of several equations whose involved variables (see Jeppsson (1996) for a detailed explanation) value must be provided to GPS-X as prior information for each simulation. As InfoWorks does not directly generate all the required quality parameters for the GPS-X procedure, the utilization of a conversion process becomes essential. This task is carried out by the fractionation module (Martin & Vanrolleghem 2014), by means of information about the entrance flow and TSS concentration to the treatment plant and the event date. This process produces almost all the necessary data required by the WWTP simulator.

**Real-time controller (RTC)**

The RTC module for this pilot is composed of a control algorithm dependent on the selected approach and an operational unit in charge of the WWTP integration. Both of
them use input data from the simulators, as well as their previous results, to generate set-points for the involved variables, which are the gates position and pump activation mode for the InfoWorks SN, and the WWTP inflow capacity for the GPS-X model.

**Control algorithm**

The two implemented control approaches for this work, as previously mentioned, are MPC and RBC.

*Model predictive control (MPC).* The MPC strategy entails the necessity of a simplified model of the considered SN. The validation of the behaviour of this model in comparison with the InfoWorks ICM high-fidelity simulations is presented in Romero et al. (2019). Figure 4 shows its graphical representation.

To facilitate its comprehension, it is important to clarify some terms:

- The catchments (Cx and Baux) emulate the network basins from which the water enters the SN.
- Each Cubeta element corresponds to a small waterway located at the bottom part of the gates, so that there is always a minimum amount of water escaping to the downstream part. The detention tank corresponds to T1.
- The overflow-by-height elements divide the incoming flow into a part continuing the downstream and the

![Schematic representation of the simplified model of the Badalona-based SN (Romero et al. 2019)](image-url)
remaining flow (only if the water surpasses a certain level) is conveyed to the receiving body via CSO.

- The final part of the network is depicted as a tank (TWWTP) whose output utilizes a pump (QTWWT) to send water to the treatment plant. However, this structure does not exist in the real UDN. It indeed approximates the behaviour of a pumping station used to convey water to the WWTP, and where water can be stored until it reaches the pumping level.

The mathematical expressions relating all these elements, as well as the input information received from the integrated model simulator, allow the optimizer to compute the required reference values for the different actuators which give room to the objectives minimization. For the case study, the elements composing the optimization cost function are the following:

1. CSO minimization ($J_{CSO}$): It entails the maximum possible reduction of the water flow conveyed to the receiving body by means of the CSOs. To remark that, for the case study, and referring to Figure 4 scheme labeling, only CSO4 and the overflow of TWWTP, known as CSOWWTP, are considered by this objective, because the rest of CSO points of the SN are passive and they cannot be controlled.

2. CSO pollutants mass minimization ($J_{mass}$): It pursues a similar aim to the previous term, but seeking the pollution mass (it is computed as the product between flow and TSS) spilling minimization. The same CSO points are considered.

3. WWTP usage ($J_{WWTP}$): Its purpose consists of the comparison between the value supplied by the WWTP inflow capacity calculator and the inflow to the treatment facility. Then, if this difference is bigger, the value of this objective will grow.

4. SN elements safety ($J_{safety}$): It mainly involves the detention tank safe functioning, regarding the filling operation. Hence, it watches the tank volume and rapidly grows once a certain threshold is surpassed.

5. Actuators operation smoothness ($J_{smooth}$): The reduction of this objective depends on the velocity and frequency of change in the set-point value of the different actuators. Therefore, the smoother their operation, the minimal this objective is.

The presented terms are unified in a cost function, where each objective is accompanied by an associated weight:

$$ f = w_{CSO}J_{CSO} + w_{mass}J_{mass} + w_{WWTP}J_{WWTP} + w_{safety}J_{safety} + w_{smooth}J_{smooth} $$

(2)

The weights values must be established in accordance with the requirements of the UDN operators and the water utility. For this case study, after performing an analysis to study the effect of the weight values on the individual objectives, these weights are settled as: $w_{CSO} = 240$, $w_{mass} = 8,000$, $w_{WWTP} = 80,000$, $w_{safety} = 10,000$ and $w_{smooth} = 100$ (some of these terms are composed by several sub-objectives, and the given value corresponds to the weight of the highest of these sub-items. For example, the smoothness term is composed of three elements: two for the tank gates and one for the pump. While the pump term is accompanied by a weight of 100, the remaining sub-items are multiplied by 1.

Once the UDS simplified model and the objective function are described, the operation of the MPC-based approach must be detailed. In the case study, the MPC-RTC problem is solved through the General Algebraic Modelling System or GAMS optimization engine and the CONOPT3 algorithm.

The process starts extracting the required input information, associating it with the corresponding GAMS variables:

- The optimizer requires data about the water and mass inputs to the model at the whole horizon $H$ (30 min) with a time step of 5 min. For the case study, this knowledge about the basins that provide water to the network (Catchment in the legend of the simplified model in Figure 4) is supplied by a historical dataset of hydraulic and quality information, adapted to the simulated scenario and obtained from InfoWorks ICM. Therefore, the optimizer is considered to be receiving a ‘perfect’ dataset about the water input to the network. For further implementations, this knowledge may be furnished as the output of a module that employs the rain prediction.

- Information about the volume and mass of the tanks at the initial time step of the optimization horizon must be
provided: the first is obtained from InfoWorks ICM, whereas the latter is provided by the previous optimization results.

- Besides, a forecast about the flow values at every pipe in the network is utilized as input information for the quality computations during the optimization process, so that the product between flow and TSS becomes simpler. This information is obtained from the results achieved at the previous simulation (and default values for the first iteration).

- The WWTP inflow capacity must be supplied to the optimizer in order to compute the $J_{\text{WWTP}}$ objective, importing this knowledge from the WWTP inflow capacity calculator. However, as the optimizer needs a forecast of this value over the horizon, this capacity calculator must be able to produce as many output values as inputs it receives.

Then, GAMS is prepared to be utilized, carrying out the optimization with the objective of minimizing the cost function described in Equation (2).

Finally, the actuators reference values are achieved from certain flows at the simplified model (after the optimization) through a flow-to-setpoint function (concretely, a conversion table). For the Badalona-based pilot, and considering the terminology applied in Figure 4, the measured flows would correspond to pipes S3 (G11/Gate 1 to tank), S4 (G21/Gate 2 to tank), S50 (G12/Gate 1 downstream), S70 (G22/Gate 2 downstream) and S8 (P1/Pump).

The output result is provided to the InfoWorks ICM regulator for the control of the actuators during the next iteration of the closed-loop. Therefore, the employed flow values must come from the first step of the optimization horizon, as it represents the next simulation time instant.

**Rule-based control (RBC).** The knowledge-based strategy proposed in this work entails the generation of a set of rules based on conditions regarding the SN and WWTP states. Their evaluation allows to achieve values for the actuators set-points. For this case study, they are the position of the four gates and the pump functioning mode. The designed decision-tree has been continuously refined by means of the knowledge gained from the pilot performance and the MPC-based approach results.

The first step consists of extracting the necessary information from the available data: the water level at the critical point of the network (the least prepared location to endure a flooding), the level of the detention tank, the flows and TSSs coming from the basins represented by $C_1$ and $C_3$ (they directly affect the tank gates), as well as the flow conveyed to the WWTP and the treatment plant inflow capacity (coming from the inflow capacity calculator).

Depending on the comparison between the previous variables and a set of thresholds, which were derived from the knowledge of the system, different actions are triggered with respect to the tank emptying and filling operations. The former checks the tank level to evaluate the necessity of activating (or deactivating) a certain number of pumps, which depends on the WWTP inflow capacity, the actual flow entering the WWTP and the current mode of the pumping operation. The latter assesses the difference in pollution concentration arriving at the tank from the basins $C_1$ and $C_3$, so that the most polluted water enters the tank with a higher priority. Consequently, the set-point values are derived depending on the desired tank operation mode, as well as the values of the previously mentioned variables.

The RBC approach has been directly programmed in Ruby, without needing extra software, designing and implementing the set of required functions at the control module of the coordination environment.

**WWTP inflow capacity calculator**

The importance of the WWTP inflow capacity calculator as part of an integrated closed-loop architecture has been remarked through the document in various sections. It provides the WWTP inflow capacity to the controller and GPS-X, allowing consideration of the treatment facility state during the decision-making process carried out by the control approach.

In the Badalona-based pilot, the capacity calculator has been derived from a criteria based on the State Point Analysis (SPA) or the Operating Point Analysis method (Wahlberg 2001). It is applied to the secondary clarifier with a double objective: to detect faults in its operation and to tune the return activated sludge or RAS, which is the part of the outcome obtained from the secondary clarifier that is conveyed...
back to the aerated bio-reactor to maximize the plant inflow capacity.

The calculator receives the entrance flow to the treatment facility and the mixed liquor settleable solids concentration or XMLSS (obtained from the previous GPS-X simulation), as well as it employs several parameters, highlighting WWTP design values or the Sludge Volume Index (SVI).

The implementation of this module has also been developed in Ruby, hence avoiding the necessity of incorporating new external components.

Closed-loop coordination environment

Once the integrated framework ingredients have been individually explained for the case study, the closed-loop coordination environment must be presented to completely define the proposed architecture.

As mentioned in the case study introduction, a Ruby module has been selected to play this role, as presented in Figure 5.

The similarities between this diagram and the scheme depicted in Figure 2 are evident, as the different components used in the case study have substituted their general counterparts. The main difference lies in the inclusion of the flow forecast input to the control algorithm, necessary for the MPC approach.

The implementation is based on the instantiation and sequential usage of several classes defined in Ruby, involving each one of the necessary elements:

- InfoWorks ICM Exchange is an Application Programming Interface (API) developed by Innovyze to manage InfoWorks ICM resources from a Ruby program or module. It has been exploited to develop a new class for the management of InfoWorks ICM. It is capable of, by means of the API methods, carrying out the configurations to the slightest detail, proceeding to simulate and saving the achieved results in the desired format.
- A class has been designed and programmed from scratch to manage the WWTP simulator operation. It implements GPS-X related tasks like input information gathering from InfoWorks ICM, simulation configuration and launching, and results storage. Additionally, the fractionation module and the WWTP inflow capacity calculator are included in this class. There are two different functioning modes for this calculator: one provides a single capacity value and another supplies a complete forecast (necessary for the MPC approach) depending on the size of the input values array.
- About the control approaches, on the one hand, another class has been generated to operate the MPC strategy at GAMS from the Ruby environment. The optimization can be configured using the necessary information obtained from InfoWorks ICM results files, previous optimizations and the WWTP inflow capacity supplied by the calculator. Besides, the optimization can be launched and its results gathered for further analysis or utilization during the next iterations.
- On the other hand, a different class has been defined for the RBC operation, implementing the explained functionalities, as well as the results files generation for posterior analysis.
- Finally, two more classes have been created for general purpose: one deals with the creation and management of Excel files from Ruby, and the other allows to generate a complete simulation results file from the ones that InfoWorks ICM creates each iteration, i.e. the separated information regarding each individual iteration is encompassed into a general results record.

To manage the usage of all these classes, a main Ruby program serves as a binding element. It controls the proper execution of the scheme depicted in Figure 5: receiving simulation settings from an external source, configuring the control approach (the employed classes and the input information that these same classes receive changes
among different strategies: Volume-Based MPC, Pollution-Based MPC or RBC), managing the coordination of the different objects created from the previously presented classes inside the loop, and saving the final results of the employed platforms; or even gathering data for subsequent analysis like the consumed time by each software.

RESULTS AND DISCUSSION

Once the proposed integrated closed-loop approach and the case study implementation have been presented, a set of simulation scenarios are employed to show the correct operation of the CLSA and the enhanced performance of the pollution-based approach.

As introduced in Section ‘Case study’, Subsection ‘SN simulator’, there exists an alert criterion to classify the different rainfall episodes, based on the relation between the precipitations intensity considering a 20-min time step and a 60-min time step ($I_{20}$ and $I_{60}$, respectively). Badalona City Hall defines it as follows:

(a) Alert Level 0: there is no rain.
(b) Alert Level 1: the rain is considered as very small ($I_{20} < 0$ mm/h or $I_{60} > 0$ mm/h).
(c) Alert Level 2: the rain is small ($I_{20} < 10$ mm/h or $I_{60} > 5$ mm/h).
(d) Alert Level 3: it is a medium-sized rain ($I_{20} < 20$ mm/h or $I_{60} > 10$ mm/h).
(e) Alert Level 4: the rain is considered as big ($I_{20} < 40$ mm/h or $I_{60} > 20$ mm/h).
(f) Alert Level 5: the rain is very big ($I_{20} < 60$ mm/h or $I_{60} > 30$ mm/h).

Therefore, regarding the necessity of considering high-intensity episodes, which demand a suitable tank management from the control approach, the following rain events have been selected:

- **22 August 2014**: this episode presents an alert level of 3, due to an $I_{20} = 42.6$ mm/h and an $I_{60} = 17.8$ mm/h. About the quality information, the previous rainfall occurred 20 days before this one. The associated return period is 0.44 years.
- **18 June 2016**: it is classified as an event with alert level of 4, due to an $I_{20} = 60.3$ mm/h and an $I_{60} = 24.40$ mm/h.

Regarding the quality data, the preceding rain happened 20 days before this event, and the return period is 0.92.

- **24 March 2017**: it presents an alert level of 3 because of an $I_{20} = 39.6$ mm/h and an $I_{60} = 24.6$ mm/h. In terms of quality, the period between this event and the previous precipitation is 20 days, as well as the return period is 0.38 years.
- **T10**: this event was artificially designed as an episode with alert level of 5 (an $I_{20} = 110.5$ mm/h and an $I_{60} = 52.9$ mm/h), in order to challenge the control approach capabilities due to the occurrence of the tank complete filling. It is assumed a period of 10 days of dry weather flow between the preceding event and this one. It also presents a return period of 10 years.

The performance at these scenarios is evaluated by means of the following key performance indicators (KPIs):

1. $CSO_{volume}$: total volume spilled through all the considered CSO points (from both SN and WWTP).
2. $CSO_{mass}$: total mass emitted by the previously mentioned CSO points.
3. $Vol_{WWTP}$: water volume reaching the WWTP.
5. $Vol_{primtreat}$: water volume that passes through the primary treatment.

In order to compare the different explained strategies, i.e. Local Control, Volume-Based MPC, Pollution-Based MPC and RBC; the evaluation of the previous KPIs are presented in Table 1.

The presented results demonstrate the benefits and potential of the detailed software architecture and demonstrate the effectiveness of the proposed solution. The proposed methodology achieved explainable and expectable performance indicators for the different tested rain events and control strategies:

- The Local Control approach turns out to be the worst for all the considered indicators. Two upshots may be extracted from this fact: the implementation of the proposed control frameworks seems to be suitable and desirable, and the functioning of the communication between the RTC module and the virtual reality software environments works in the desired way, correctly
applying set-points coming from an external source to the SN virtual actuators.

- Regarding the pollutants emission to the receiving body, the PBMPC strategy presents the best performance, especially at intense and long rain events like T10. This not only demonstrates the effectiveness of the pollution-based strategy but also the proper operation of the several elements composing the software architecture: the quality terms of the simplified model, the proper configuration of the RTC module depending on the selected control strategy, etc.

- The usefulness of the WWTP integration is demonstrated at the PBMPC and RBC strategies by two facts: the lower CSO spilled volume at these integrated approaches (due to a better management of the treatment facility, so that its internal CSOs are smaller) with respect to the VBMPC, as well as a higher treated volume (this difference gets higher in the further treatment steps) despite the sent volume to the treatment plant is similar. Besides, the achieved results indicate the correct functioning of the integration scheme regarding its implementation, including the configuration of the WWTP simulator, the programming of the inflow capacity calculator, etc.

- Furthermore, the comparison among the different rain scenarios shows that the selection of the proper execution of different operations, like the SN simulation configuration and the results extraction.

Therefore, the presented methodology consists of an effective framework for the deployment of RTC strategies for UDS. The software architecture gathers the most important features of previous state-of-the-art approaches: the exploitation of an efficient inter-application communication between modules as in Riaño-Briceño et al. (2014), the possibility of implementing pollution-based control strategies, as in Joseph-Duran et al. (2015), which have been demonstrated to be more effective by the results in Table 1; and the integration of the WWTP state during the control operation, as explained in Butler & Schütze (2005).

For an extended presentation of control strategies and results of additional scenarios, see Sun et al. (2020).

### CONCLUSIONS

This work presents a complete methodology for the development of an integrated closed-loop framework to implement pollution-based control approaches for UDS.

The structure has been composed of an integrated model simulator, combining the SN and WWTP functionalities to operate as virtual reality; and the RTC module composed by the control algorithm and the WWTP inflow capacity calculator. The communication among the different software components, as well as the organization of the

### Table 1 | Performance results for the presented rain scenarios

<table>
<thead>
<tr>
<th>Control strategy</th>
<th>Local control</th>
<th>Volume-based MPC</th>
<th>Pollution-based MPC</th>
<th>Rule-based control</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Aug 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSO\text{volume} (m³)</td>
<td>83,207</td>
<td>78,443</td>
<td>71,820</td>
<td>72,763</td>
</tr>
<tr>
<td>CSO\text{mass} (kg)</td>
<td>56,387</td>
<td>54,375</td>
<td>53,531</td>
<td>53,950</td>
</tr>
<tr>
<td>Vol\text{WWTP} (m³)</td>
<td>114,314</td>
<td>119,364</td>
<td>117,284</td>
<td>117,161</td>
</tr>
<tr>
<td>Vol\text{pretreat} (m³)</td>
<td>104,509</td>
<td>114,029</td>
<td>113,796</td>
<td>112,485</td>
</tr>
<tr>
<td>Vol\text{primtreat} (m³)</td>
<td>88,965</td>
<td>94,587</td>
<td>101,172</td>
<td>99,625</td>
</tr>
<tr>
<td>Vol\text{sectreat} (m³)</td>
<td>86,560</td>
<td>92,180</td>
<td>98,733</td>
<td>97,218</td>
</tr>
<tr>
<td>18 Jun 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSO\text{volume} (m³)</td>
<td>406,482</td>
<td>399,975</td>
<td>390,354</td>
<td>386,144</td>
</tr>
<tr>
<td>CSO\text{mass} (kg)</td>
<td>88,863</td>
<td>88,778</td>
<td>87,402</td>
<td>87,660</td>
</tr>
<tr>
<td>Vol\text{WWTP} (m³)</td>
<td>235,505</td>
<td>260,507</td>
<td>260,664</td>
<td>262,011</td>
</tr>
<tr>
<td>Vol\text{pretreat} (m³)</td>
<td>208,505</td>
<td>236,791</td>
<td>236,948</td>
<td>236,902</td>
</tr>
<tr>
<td>Vol\text{primtreat} (m³)</td>
<td>144,097</td>
<td>154,757</td>
<td>164,122</td>
<td>167,844</td>
</tr>
<tr>
<td>Vol\text{sectreat} (m³)</td>
<td>134,295</td>
<td>144,035</td>
<td>153,676</td>
<td>157,791</td>
</tr>
<tr>
<td>24 Mar 2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSO\text{volume} (m³)</td>
<td>250,399</td>
<td>247,216</td>
<td>246,968</td>
<td>233,037</td>
</tr>
<tr>
<td>CSO\text{mass} (kg)</td>
<td>56,677</td>
<td>52,965</td>
<td>49,596</td>
<td>55,196</td>
</tr>
<tr>
<td>Vol\text{WWTP} (m³)</td>
<td>112,893</td>
<td>117,360</td>
<td>118,913</td>
<td>117,331</td>
</tr>
<tr>
<td>Vol\text{pretreat} (m³)</td>
<td>107,563</td>
<td>113,992</td>
<td>114,227</td>
<td>113,649</td>
</tr>
<tr>
<td>Vol\text{primtreat} (m³)</td>
<td>98,729</td>
<td>105,782</td>
<td>105,425</td>
<td>105,692</td>
</tr>
<tr>
<td>Vol\text{sectreat} (m³)</td>
<td>78,523</td>
<td>81,583</td>
<td>82,242</td>
<td>95,758</td>
</tr>
<tr>
<td>T10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSO\text{volume} (m³)</td>
<td>22 Aug 2014</td>
<td>247,216</td>
<td>246,968</td>
<td>233,037</td>
</tr>
<tr>
<td>CSO\text{mass} (kg)</td>
<td>56,677</td>
<td>52,965</td>
<td>49,596</td>
<td>55,196</td>
</tr>
<tr>
<td>Vol\text{WWTP} (m³)</td>
<td>112,893</td>
<td>117,360</td>
<td>118,913</td>
<td>117,331</td>
</tr>
<tr>
<td>Vol\text{pretreat} (m³)</td>
<td>107,563</td>
<td>113,992</td>
<td>114,227</td>
<td>113,649</td>
</tr>
<tr>
<td>Vol\text{primtreat} (m³)</td>
<td>98,729</td>
<td>105,782</td>
<td>105,425</td>
<td>105,692</td>
</tr>
<tr>
<td>Vol\text{sectreat} (m³)</td>
<td>78,523</td>
<td>81,583</td>
<td>82,242</td>
<td>95,758</td>
</tr>
</tbody>
</table>
involved tasks, has been implemented inside a coordination environment.

A pilot based on the Badalona UDS has been presented to exemplify the usage of the proposed methodology in several scenarios, as well as demonstrating the capabilities of the advanced framework to generate effective control commands for the SN management. Results involving various rainfall events are exposed to highlight the appropriateness, efficiency and superiority of the presented methodology.

Therefore, new control strategies can be designed and initially tested by implementing a sufficiently representative pilot of a UDS by means of the proposed methodology, achieving different conclusions via the employment of a virtual reality application.

The generated closed-loop architecture deals with stand-alone and/or proprietary software components whose operation was not conceived to be coordinated by an external element. The election of the programs and applications composing the architecture of this case study arises from different reasons. The high-fidelity simulators correspond to proprietary software selected by the network and plant operators involved in the project. The optimization software was chosen due to the wide previous experience at its usage. Finally, the coordination platform was designed in Ruby because of the existence of an API to manage one of the simulators, facilitating its operation. These constraints in the software selection lay bare the difficulty of developing a proper coordination algorithm. However, the proposed solution entails a successful solution to the inclusion of these commercial tools, making the most of their outstanding capabilities while efficiently and effectively coordinating their tasks.

These achievements entail a further step in the UDS control field, due to the wide range of possible software solutions that can be incorporated to the presented methodology, as well as the several features that it allows to exploit, like pollution-based approaches, WWTP integration, etc.

Several future lines of work remain open. Regarding the software development, the design of a closed application would be of interest, including an application layer in order to improve the operator experience. Moreover, the efficiency of the methodology might be tested and compared with other approaches. About the water system operation, numerous control strategies can be implemented and evaluated, as well as new case studies may be considered. Finally, the implementation of the methodology into a real network would be the definitive step to assess the effectiveness of the solution.

ACKNOWLEDGEMENTS

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

DATA CONFIDENTIALITY STATEMENT

Data, tools and models employed in this work are confidential or commercial, and hence, we refer other researchers to contact SUEZ Spain Group (for access to detailed, simplified models; the Closed-loop Simulation Framework software; and the rain data) and Innovyze (for licenses to InfoWorks ICM).

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES


Aulinas, M., Nieves, J. C., Cortés, U. & Poch, M. 2011 Supporting decision making in urban wastewater systems using a knowledge-