Real-time foul sewer hydraulic modelling driven by water consumption data from water distribution systems

Zhang, Qingzhou; Zheng, Feifei; Jia, Yueyi; Savic, Dragan; Kapelan, Zoran

DOI
10.1016/j.watres.2020.116544

Publication date
2021

Document Version
Accepted author manuscript

Published in
Water Research

Real-time foul sewer hydraulic modelling driven by water consumption data from water distribution systems

Qingzhou Zhang[1], Feifei Zheng[2], Yueyi Jia[3], Dragan Savic[4] and Zoran Kapelan[5]

[1]**Qingzhou Zhang:** Postdoctoral Research Fellow, College of Civil Engineering and Architecture, Zhejiang University, China. wdswater@gmail.com.

[2]**Feifei Zheng:** Corresponding author, Professor, College of Civil Engineering and Architecture, Zhejiang University, China. feifeizheng@zju.edu.cn. Tel: +86-571-8820-6757. Postal address: A501, Anzhong Building, Zijingang Campus, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, China 310058.

[3]**Yueyi Jia:** PhD candidate, College of Civil Engineering and Architecture, Zhejiang University, China. yueyi@zju.edu.cn.

[4]**Dragan Savic:** Chief Executive Officer, KWR Water Research Institute, Dragan.Savic@kwrwater.nl, Professor, Centre for Water Systems, University of Exeter, North Park Road, Exeter, EX4 4QF, United Kingdom.

[5]**Zoran Kapelan:** Professor, Department of Water Management, Delft University of Technology, The Netherlands, z.kapelan@tudelft.nl

Abstract: One way to address many issues (e.g., illicit inflows) within foul sewer systems (FSSs) is via real-time hydraulic models. However, a bottleneck within real-time FSS modelling is the lack of spatio-temporal manhole inflow data. To address this problem, this paper proposes a new method to develop real-time FSS models driven by water consumption data from associated water distribution systems (WDSs) that often have a proportionally larger number of sensors. Within the proposed method, the FSS manholes are integrated with the WDS water consumption nodes based on their underlying physical connections. An optimization approach is subsequently proposed to identify the transfer factor *k* between nodal water consumption and FSS manhole inflows based on historical observations. These identified *k* values combined with the acquired real-time nodal water consumption data from the WDS equipped with a dense network of sensors drive the FSS real-time modelling. The proposed method is applied to two real FSSs and results show that it can produce sewer flow and manhole water depth simulations matching well with observations at the monitoring locations with averaged R^2, *NSE* and *KGE* (Kling-Gupta efficiency) around 0.99, 0.88 and 0.92, respectively. It is anticipated that real-time models developed by the proposed method can be useful for the efficient FSS management and operation.

Key words: foul sewer system; water consumption data; real-time models; water distribution system

1 Introduction

Sewer networks are traditionally designed to collect wastewater from residential, commerical and/or industrial clients or possible stormwater from urban surfaces due to rainfall events. Collected wastewater is transported then downstream to wastewater treatment plants (WWTPs) or released directly into rivers (Bailey et al., 2019). These sewer networks are often called combined sewer systems (CSSs), which have been widely used in large cities around the world (Li et al., 2014). In recent years, there is a growing trend in separating CSSs into independent storm drainage systems and foul sewer systems (FSSs, Schilperoort et al. 2013). The former are used to convey urban runoff solely to surface water bodies (e.g., rivers) and the latter deliver sewerage collected from houses and commercial buildings before being conveyed to treatment facilities. Such a separation is mainly driven by the purpose to improving urban water environments as combined sewer overflows (CSOs, Black and Endreny, 2006) would inevitably threaten the ecological health of the receiving water (Joseph-Duran et al., 2015).

Over the past decade, many FSSs around the world have experienced significant changes due to population growth and quick urbanization, which is especially the case in many developing countries such as China (Sweetapple et al., 2018). These changes are mainly represented by the expanded spatial scales of FSSs, the increased complexity in their topology structures and the aged systems (Rokstad and Ugarelli, 2015, Huang et al., 2018). This, consequently, results in significant challenges/difficulties for effective FSS management and operation, and hence many issues exist (Garda et al., 2016). A typical issue is the deposits in the FSSs, including sediments (Seco et al., 2018), fat, oil and grease (FOG, Liu et al., 2016) and toilet papers (Eren and Karadagli, 2012). All those deposits can directly affect flow capacity of the sewers, causing overflows from CSOs and manholes as well as potential water quality issues (e.g, odor issues, Liu et al., 2016; Talaiekhozani et al., 2016). Another common issue is the illicit discharges from local factories (Irvine et al. 2011; Banik et al. 2017), where these discharges often contain toxic substances (e.g., heavy metals) that are often beyond the processing capacity of the downstream WWTPs. This, therefore, can result in functional failures of WWTPs and consequently significant contamination of the receiving water body (McCall et al. 2016). In addition to issues of deposits and illicit discharges, leaks of the sewers, groundwater infiltration and illicit connections between FSSs and stormwater pipes are frequently reported, inducing serious contamination to the surrounding water environments (Lepot et al., 2017; Beheshti and Saegrov, 2019).

The issues mentioned above have appreciably affected the urban water environments due to the resultant overflows from CSOs/manholes and leaks of FSSs. One way to address these issues is the placement of sensors within the FSS to monitor the water depths and sewer flows, thereby triggering a warning when the observations are significantly higher or lower than the historical data (Ahm et al. 2016). However, due to the high purchase cost and intensive maintenance efforts associated with these sensors, the monitoring network is often sparse for the majority of the FSSs (Kleidorfer et al., 2012). Consequently, a warning associated with the potential issues (e.g., overflows or leaks) can be only available for the very limited number of FSS locations in the proximity of sensors. In addition, the abnormal observations at the monitoring locations may be caused by sudden discharges increases caused by the water users, rather than the illicit discharges, resulting in a potentially high likelihood of false warning (Koch et al. 2011). More importantly, relying solely on the observations from the sewer sensors cannot offer predictions on the hydraulic status of the entire FSS in future (Bruen and Yang, 2006). To this end, real-time sewer hydraulic modelling can be promising in addressing the issues associated with the FSSs as mentioned above, where the hydraulic variables such as water depths and sewer flows across the entire FSS are simulated in real-time. These simulations, combined with observations at the monitoring locations, can be used to deduce whether leaks, illicit discharges, deposits and illicit connections exist in the FSS, as well as facilitate the localization of such events.

Manhole inflow data (i.e., sewer discharges of the water users) at a high time resolution (say every 30 minutes) is the key to enable the development of a real-time FSS hydraulic model. However, discharge data of such a high temporal and spatial resolution are typically unavailable in engineering practice, resulting in a large challenge for real-time modeling of FSS (Breinholt et al., 2013). To deal with this problem, a widely used approach is to calibrate a model to estimate manhole inflows with the aid of limited in-sewer observations (Korving and Clemens, 2005). While some calibration methods are available in the literature (e.g., di Pierro et al., 2005, Khu et al., 2006, Broekhuizen et al., 2020), they mainly focus on calibrating the underlying rainfall-runoff relationship for the combined sewer systems in an off-line manner, thereby predicting the floods or sewer overflows caused by rainfall. These previously published methods, therefore, cannot be used or at least are difficult to estimate FSS manhole inflows in real-time.

The real-time management of the FSS has received great attention over the past few decades, with the main focus on system real-time control based on observations (Schütze et al., 2002; Sara et al., 2020). More specifically, real-time control is defined as a timely operation of an FSS based on continuously monitored process data. Those data are water levels and sewer flows in the system, with operations including the activation of pumps, sluice gates and weirs used to improve system performance (e.g., reduce the overflows, Schütze et al., 2003). However, these real-time control studies operated the hydraulic facilities (e.g., pumps) with the aid of system observations rather than FSS simulations, and hence they differ significantly from the real-time FSS hydraulic modelling, which is the aim of the present study.

The main difficulty associated with the calibration of FSS manhole inflows based on the limited number of monitoring sites is the "equifinality" (Khu et al., 2006). More specifically, a large number of manhole inflow combinations can produce similar agreements between simulated and observed water levels at monitoring locations. As a result, it is very difficult, if not impossible, to identify a particular parameter set (i.e., manhole inflow combination) that can represent the true underlying spatial distribution of the discharges from water users into the FSS.

To address the "equifinality" issue, this paper proposes a new method to enable the development of real-time FSS hydraulic model. Within the proposed method, the FSS model is integrated with its corresponding water distribution system (WDS) hydraulic model for the same area being considered. Such a model integration approach is possible as the WDS models have already been widely used (Walski et al., 2003). In addition, the number of sensors (e.g., smart demand meters, pressure sensors and flow meters) deployed in the WDSs can be large, which is, at least partly, driven by the quick developments of the Internet of Things in recent years (Zheng et al., 2018). Such a dense sensor network can greatly facilitate the estimation of real-time nodal water consumption for the WDS models as demonstrated in previous studies (Creaco et al., 2019). This is especially the case in recent years as smart demand meters have been increasingly used in many WDSs, providing water consumption data for many users (not only large users but also residential users) in a real-time manner (typically every 15 or 30 minutes, Creaco et al., 2018). Such near real-time and high-density spatial water consumption data can be assimilated with the limited in-sewer observations to develop a real-time FSS hydraulic model. This is the key feature and novelty of the method presented in this paper.

The concept of incorporating water consumption data into FSS modeling can be dated back to Bruke et al. (1986), where an FSS model was calibrated using monthly water use records. More recently, Bailey et al. (2019) presented a new FSS model, where the stochastically simulated water demands were imported into the sewer network model. While these limited previous studies have made great contributions in assimilating water use records into FSS modelling (mainly used for FSS design purpose), the water consumption data used are either collected manually at a very low time resolution (e.g., monthly, Bruke et al., 1986) or provided by a stochastic simulator (Bailey et al., 2019). Consequently, these data cannot represent the true underlying temporal and spatial variations of the manhole inflows. Therefore, they cannot be used to develop real-time FSS models, which is the focus of this study.

The key feature of the proposed method is that the real-time FSS model is developed using a large number of existing sensors within the WDSs. This implies that it is not necessary to deploy a large number of sewer sensors (which is often very expensive in terms of both sensor purchases and maintenance) to enable real-time sewer modelling, making the proposed method attractive for practical applications. This paper is organized as follows. The proposed methodology is described in Section 2, followed by the descriptions of the case studies considered in Section 3. Results and discussions are given in Section 4. Finally, the conclusion section (Section 5) shows the main observations and implications of this paper.

2. Methodology

2.1 The overall modelling concept

Fig.1 illustrates the overall concept of the proposed method, where a foul sewer system (FSS) and a water distribution system (WDS) for a small area are presented. Typically, raw water from reservoirs or rives is pumped into the water treatment plants in order to improve water quality to a required standard (Wu et al., 2011). Subsequently, the treated water is conveyed to the WDS, satisfying demands for various users including residents, schools, hospitals, industrial and commercial buildings, as shown in Figure 1. To ensure water supply safety, sensors are often deployed in the WDS (Figure 1), including pressure sensors, flow meters and smart demand meters. The latter have been increasingly being deployed in recent years to monitor water consumptions for the users in a near real-time manner (Creaco et al., 2018).

Consequently, such a dense sensor network enables the development of real-time WDS modelling, which has been an important trend within the water supply domain (both research and industry) due to its great merits in facilitating effective system management as highlighted in Creaco et al. (2019).

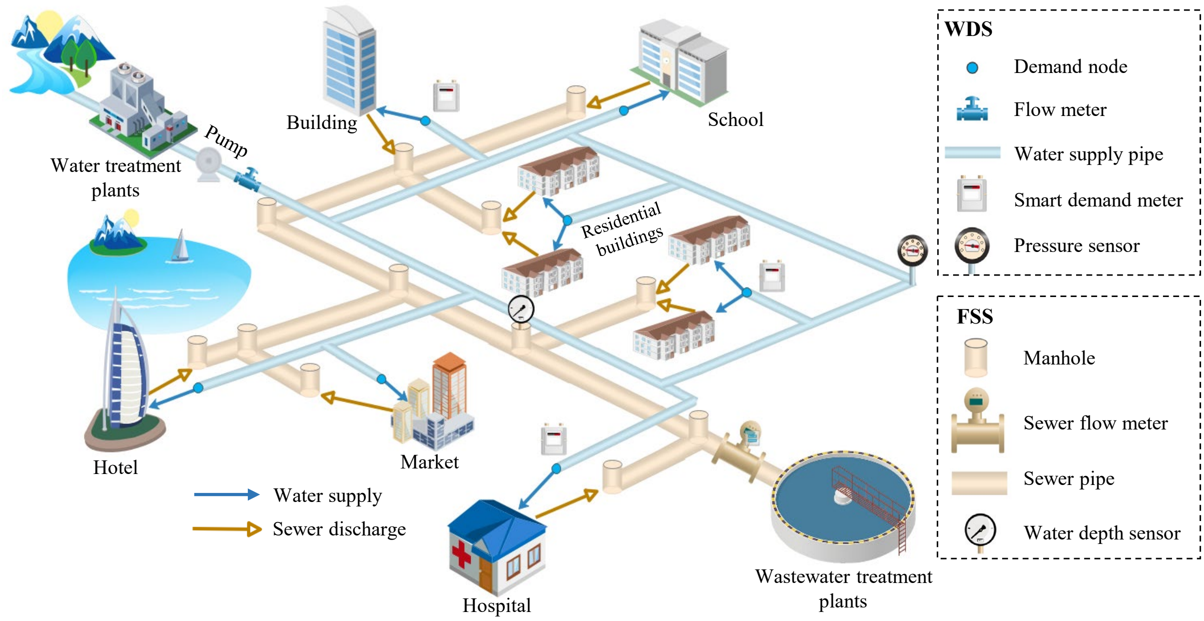

Fig. 1 An illustration of the concept for the proposed modelling method, where a water distribution system and a foul sewer system are presented

Inherently, local residents or commercial/industrial users discharge sewage after water consumption as illustrated in Figure 1. Sewer pipes collect and convey the sewage to downstream wastewater treatment plants, with a limited number of water depth or sewer flow sensors installed to monitor hydraulic state of the system. Consequently, the following equation can be used to represent the underlying relationship between water consumption and sewage discharge for user *i*:

$$d_i = F(q_i, k_i, t_i) \quad (1)$$

where d_i is the sewer discharge rate of user i (i.e., manhole inflow rate) resulting from its water consumption q_i taken from the WDS, t_i represents the time delay, i.e., the time between the clean water entering the user property and the time it reaches the local sewer network, k_i is the transfer factor for user i, representing the proportion of supplied water that ends up in the local sewer network; k_i typically has a value between 0.7 and 1.0 (Behzadian and Kapelan 2015). Equation (1) represents the fundamental rule/assumption in the proposed method used to build the connections between the WDS water consumption data and the FSS manhole inflows.

Fig.2 presents the overall methodology of the proposed method, with two modules involved. The first module consists of three phases, which are carried only once in an offline manner, and the second module only has the fourth phase (Phase 4) of the proposed method, which runs in real-time. The details are given below.

Phase 1: Integrate the WDS and FSS models (carry out once). Within this phase, the FSS and WDS models are developed with hydraulic facility information (e.g., water supply pipes, tanks, sewer pipes) taken from external sources such as the GIS or asset management system. This is followed by the building of the connections between each WDS demand node and the FSS manhole based on the spatial distance with details given in Section 2.2.

Phase 2: Calibrate nodal water consumptions of the WDS. It should be highlighted that the calibration of the nodal water consumptions in Phase 2 is conducted offline, which is used to provide data for Phase 3. More specifically, Based on a particular time period of historical data from pressures sensors, flow meters and smart demand meters deployed in the WDS, the nodal water consumption without smart demand meters are estimated for a given time resolution (often equals the time resolution of the flow or pressure sensors) with details given in Section 2.3

Phase 3: Estimate the transfer factor k for each manhole of the FSS. According to the identified relationship between WDS nodes and the FSS manholes in Phase 1, as well as the calibrated nodal water consumptions in Phase 2, the transfer factor k is determined. For this, an evolutionary algorithm (EA) is applied with the objective function defined in Eq. (9-13) and using sewer observations, with details given in Section 2.4.

Phase 4: Model the FSS in a real-time manner. Data from pressure sensors, flow meters and the available smart demand meters in the WDS are acquired at the current time t. These data are used as the inputs for the real-time WDS modelling to estimate water consumption for each node (q_i) within the WDS (Section 2.3). Eq. (1) is subsequently used to update the manhole inflows d_i based on the known q_i and identified k values (Phase 3). Finally, the FSS is modelled by updating manhole inflows d_i in real-time. This offers short-term hydraulic predictions (water depths at manholes and flow rates in sewer pipes) of the entire FSS with a particular time resolution (if say every 30 minutes used in this paper).

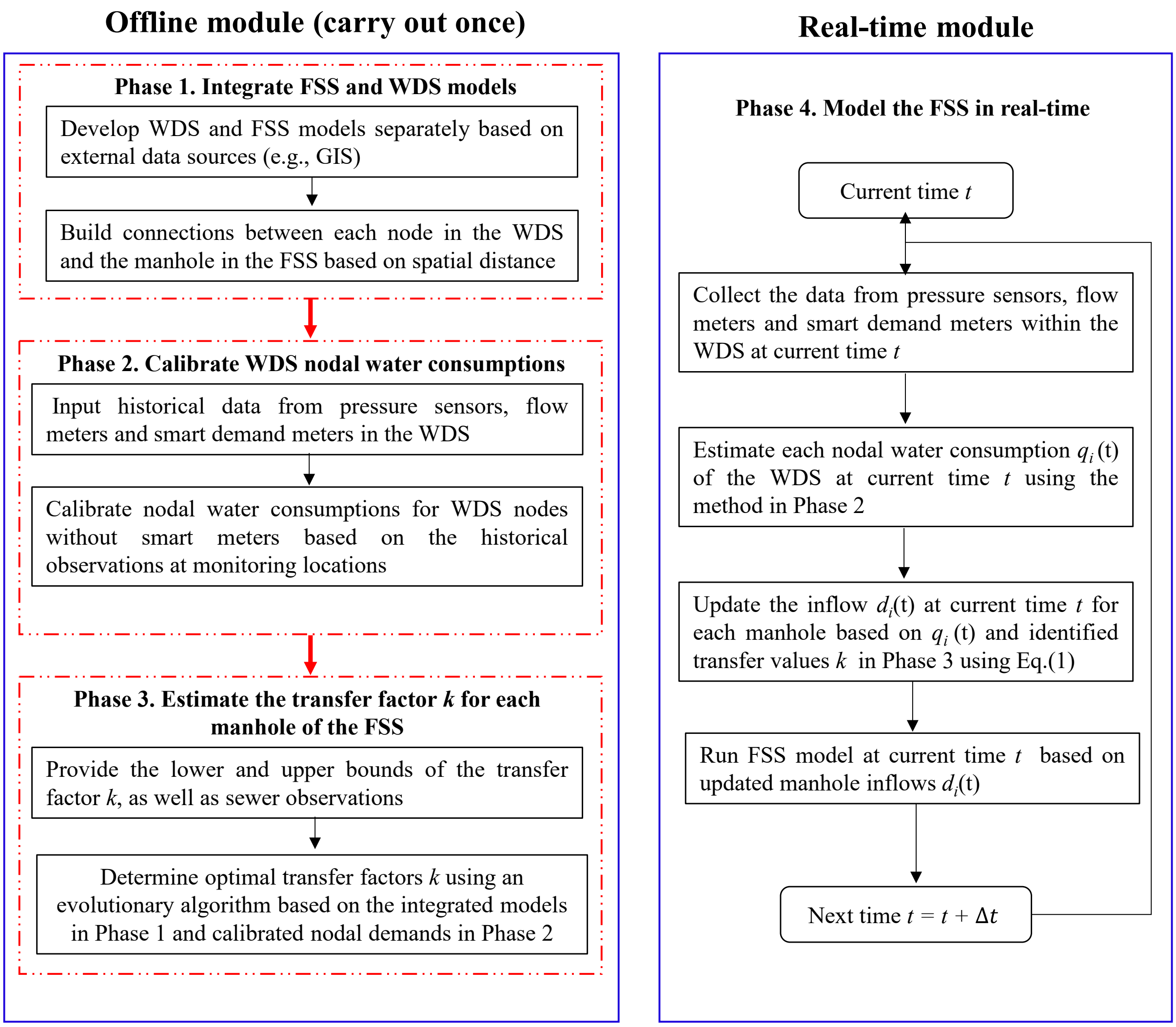

Fig. 2 The overall methodology of the proposed method

It is noted that a few assumptions are made in the proposed method, with the justification given below.

(i) Given that the proposed FSS real-time modelling method is driven by water consumption data from the WDS, the number of available smart demand meters in the WDS is important to ensure the high accuracy of the FSS simulations. For the two case studies considered in this paper, the number of smart demand meters is reasonably high, making them perfectly suited for the demonstration of the proposed method. However, some WDSs may have relatively low coverage of the smart demand meters (e.g., only installed for large demand users). While such a case would not affect the application of the proposed method, the accuracy of the WDS nodal water consumption values and the FSS real-time simulations can be affected. However, it is anticipated that smart demand meters are increasingly used by water utilities as a result of the quick developments in the Internet of Things (Zheng et al., 2018; Creaco et al., 2019), and hence the applicability of the proposed method is only going to grow.

(ii) The proposed method assumes that observations from the WDS and FSS sensors (including smart demand meters) are accurate within the applications in this study. However, in reality, observation errors can exist due to the sensor malfunctions or signal transmission issues. Therefore, it is necessary to incorporate the potential observation errors into the modelling framework. Although that is an important study direction, it is beyond the scope of the paper and will be the focus of future work.

(iii) As shown in Equation (1), the time t_i implies that the nodal water consumption q_i estimated at time t using the smart demand meter or the real-time calibration method (section 2.3) should correspond to the manhole inflow at time $t+t_i$. The value of t_i can be dependent on the particular user properties, including the characteristics of the water supply area associated with the demand node, as well as the physical characteristics (e.g., length and slopes) of the connecting sewer pipes between users and the corresponding manholes. In this study, t_i is ignored as this value is typically small, ranging from several minutes up to 15 minutes for many cases (Wu et al., 2011). This assumption is considered valid in our study as the time resolution used for real-time FSS modelling in this paper is lower (i.e., 30 minutes).

(iv) In this study, a linear transfer function with a constant factor of k is proposed to describe the underlying relationship between the nodal water consumption and manhole inflows. While being simple in practical implementation, the transfer function as well as the k factor can be affected by not only the time delay t_i in Equation (1), but also the infiltration inflows and the properties of water users. More specifically, the transfer function may be different between the water users with or without smart demand sensors, and the k factor may temporally vary, or even change as a function of different water users. These influences need further consideration in future study along this research line.

2.2 Integrate the FSS and WDS models

Typically, FSS and WDS models are developed with the aid of the geographic information system (GIS) or the asset management system for the analyzed area (Behzadian and Kapelan, 2015; Huang et al., 2018). The details (e.g., locations and length) of various system components including pipes, tanks, valves and pumps can be taken from the GIS. This is often followed by system skeletonization in which many facilities (mainly small pipes) are removed or simplified without significantly affecting the hydraulic properties of the original full system (Huang et al., 2020). To enable the practical application of Eq. (1), it is important to build the connection between each demand node i, representing the water consumption in the WDS model, and the manhole, representing the facility to collect the sewages in the FSS model. Such a connection indicates that the consumption at the demand node i is received by its associated manhole. To this end, the WDS and FSS models are integrated within their development processes in this study.

Figure 3 illustrates the proposed integration method, where a water demand node in the WDS model is assigned to the manhole of the FSS model within a shortest distance from each other. The rationale behind this is that manholes are often built near the water users (nodes in the WDS model) to collect their sewerage discharges. Consequently, two cases are available as shown in Figure 3, which are (i) one demand node is assigned to a manhole, and (ii) multiple demand nodes are assigned to a single manhole. In addition to these two relatively simple cases, in practice, one demand node can be associated with multiple manholes, which is possible when this demand node represents many users. However, it is difficult to know the proportion of the total discharge associated with each relevant manhole. For the sake of simplicity, a single manhole with the minimum spatial distance to this demand node is selected to deliver the total discharge. While such a simplification can cause an unrealistic hydraulic status in a very small area relative to the original full system, its impacts on the overall results can be negligible. Since each demand node (say node i) in the WDS is assigned to a particular manhole in the FSS model, the water consumption of this node (q_i) is considered as the approximate manhole inflows (d_i). Their underlying flow relationship needs to be further accurately determined with the incorporation of the transfer factor k_i as shown in Equation (1).

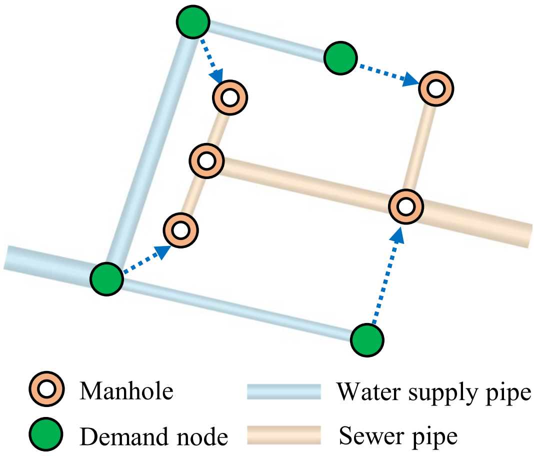

Fig. 3 Illustration of the proposed integration method for FSS and WDS model developments

It is noted that when a higher level of accuracy is needed for a practical application, individual water consumption and sewer connections could be identified if required. This will lead to a slight modification on the WDS and FSS model topologies, which can better reflect the flows of supplied drinking water and generated wastewater by different users.

2.3 Calibrate nodal water consumption based on historical observations

Based on the built connections between each WDS node and its corresponding FSS manhole as described in Section 2.2, nodal water consumption is the driver for triggering the real-time FSS modelling. In the WDSs, many smart demand meters can be available, providing near real-time water consumption data (if say every 15 minutes or 30 minutes) for WDS nodes (users), especially for water users with large demand. However, in practice, it may not be necessary to have smart meters installed at each demand node, and hence this study adopts a calibration method to enable the estimation of water consumption at the nodes without smart demand meters.

In this study, the numerical approach described in Zhang et al. (2018) is selected to calibrate the nodal water consumptions due to its demonstrated efficiency. The objective function of this adopted numerical method is formulated as the weighted sum of squared differences between the field-observed and model-simulated responses (pressures and flows) at monitoring points in the WDS within a particular time period (i.e., the time resolution of the monitoring data), i.e. as follows:

$$\text{Min}: f(\mathbf{q}) = \sum_{i=1}^{NH} w_h^i [H_i^o - H_i(\mathbf{q})]^2 + \sum_{j=1}^{NF} w_q^j [Q_j^o - Q_j(\mathbf{q})]^2 = \begin{bmatrix} \mathbf{H}^o - \mathbf{H}(\mathbf{q}) \\ \mathbf{Q}^o - \mathbf{Q}(\mathbf{q}) \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \mathbf{H}^o - \mathbf{H}(\mathbf{q}) \\ \mathbf{Q}^o - \mathbf{Q}(\mathbf{q}) \end{bmatrix} \quad (2)$$

where $\mathbf{q}$ is the vector of nodal water consumptions in the WDS, including known water consumption data at nodes with smart demand meters and unknown nodal water consumption data; NH and NF are the numbers of observed nodal pressures and pipe flows, respectively; w_h^i and w_q^j are the weighting factors for observed pressures at the i^{th} node and observed flows at the j^{th} pipe, respectively, where $w_h^i = 1/(H_i^o)^2$ and $w_q^j = 1/(Q_j^o)^2$ are used in this study following Kun et al. (2015) and Zhang et al. (2018). H_i^o and $H_i(\mathbf{q})$ are the observed and simulated pressures at the i^{th} node respectively; Q_j^o and $Q_j(\mathbf{q})$ are the observed and simulated flows at the j^{th} pipe respectively. Equation (2) can be expressed in the matrix form (see above) by using $\mathbf{H}^o = \left[H_1^o, H_2^o, ..., H_{NH}^o\right]^T$, $\mathbf{Q}^o = \left[Q_1^o, Q_2^o, ..., Q_{NF}^o\right]^T$, $\mathbf{H}(\mathbf{q}) = \left[H_1(\mathbf{q}), H_2(\mathbf{q}), ..., H_{NH}(\mathbf{q})\right]^T$, $\mathbf{Q}(\mathbf{q}) = \left[Q_1(\mathbf{q}), Q_2(\mathbf{q}), ..., Q_{NF}(\mathbf{q})\right]^T$, and $\mathbf{W} = diag([w_h^1, w_h^2, ..., w_h^{NH}, w_q^1, w_q^2, ..., w_q^{NF}])$.

Expanding Eq. (2) through the first-order Taylor series:

$$f(\mathbf{q}+\Delta\mathbf{q}) \approx \begin{bmatrix} \Delta\mathbf{H}-\mathbf{J}_H\Delta\mathbf{q} \\ \Delta\mathbf{Q}-\mathbf{J}_Q\Delta\mathbf{q} \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \Delta\mathbf{H}-\mathbf{J}_H\Delta\mathbf{q} \\ \Delta\mathbf{Q}-\mathbf{J}_Q\Delta\mathbf{q} \end{bmatrix} \quad (3)$$

where $\Delta\mathbf{H}=\mathbf{H}^o-\mathbf{H}(\mathbf{q})$ and $\Delta\mathbf{Q}=\mathbf{Q}^o-\mathbf{Q}(\mathbf{q})$ are the differences between the observed and simulated values of nodal pressures and pipe flows, respectively; $\mathbf{J}_H = \dfrac{\partial\mathbf{H}(\mathbf{q})}{\partial\mathbf{q}}$ and $\mathbf{J}_Q = \dfrac{\partial\mathbf{Q}(\mathbf{q})}{\partial\mathbf{q}}$ are the Jacobian matrix with details given in Zhang et al. (2018). Since Eq. (2) is a convex function (Kun et al., 2015), the minimum objective value of Eq. (2) can be obtained when its first-order derivative (Eq. (3)) equals to zero, that is:

$$\frac{df(\mathbf{q}+\Delta\mathbf{q})}{d\Delta\mathbf{q}} = -2\begin{bmatrix} \mathbf{J}_H \\ \mathbf{J}_Q \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \Delta\mathbf{H}-\mathbf{J}_H\Delta\mathbf{q} \\ \Delta\mathbf{Q}-\mathbf{J}_Q\Delta\mathbf{q} \end{bmatrix} = 0 \quad (4)$$

By solving Eq. (4), $\Delta\mathbf{q}$ can be obtained as follows:

$$\Delta\mathbf{q} = \left(\begin{bmatrix} \mathbf{J}_H \\ \mathbf{J}_Q \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \mathbf{J}_H \\ \mathbf{J}_Q \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{J}_H \\ \mathbf{J}_Q \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \Delta\mathbf{H} \\ \Delta\mathbf{Q} \end{bmatrix} \quad (5)$$

$$\mathbf{q}^{s+1} = \mathbf{q}^s + \Delta\mathbf{q}^s \quad (6)$$

where $s=0,1,...,S$ is the iteration number (S is the maximum allowed number of iterations). It is highlighted that the water consumption at nodes with smart demand meters are known within the entire calibration process and hence $\Delta\mathbf{q}$ is only considered for the nodes without smart meters. To ensure the estimated nodal water consumption values are practically meaningful, the domain knowledge has been incorporated within the calibration process in this study as shown below (Wu et al., 2010):

$$q_i^{s+1} = \begin{cases} q_i^{\min}, & \text{if } q_i^{s+1} < q_i^{\min} \\ q_i^{\max}, & \text{if } q_i^{s+1} > q_i^{\max} \\ q_i^{s+1}, & \text{others} \end{cases} \tag{7}$$

where $q_i^{\min} = (1-p)\times q_i^{initial}$ and $q_i^{\max} = (1+p)\times q_i^{initial}$ are the minimum and maximum allowed water consumptions at node i respectively; p is the percentage generally within 10%~20% in practice (Zhang et al., 2018); $q_i^{initial}$ is estimated using

$$q_i^{\text{initial}} = \frac{l_i}{L_T - L_M}(Q_T - Q_M) \tag{8}$$

where l_i is the length of the pipe associated with node *i*; L_T and L_M is the total pipe length of all nodes and the length of pipes associated with smart demand meters respectively; Q_T is the total water consumption of the WDS at a given time period (e.g., 30 minutes), which is estimated based on the flow meters installed at the outlet of the water treatment plants and volume changes in the tanks if available; Q_M is the sum of the water consumption values measured by the available smart demand meters within the WDS at a given time period.

The calibration process at each time period (i.e., the time resolution of the monitoring data, e.g. 30 minutes) is executed by iteratively updating $\Delta\mathbf{q}$ in Eq. (6) until the maximum value of vector $\|\Delta\mathbf{q}\|$ is smaller than a given threshold value ε (e.g. $\varepsilon = 0.1$). The entire calibration process is executed again once the monitoring data from sensors are updated, representing a real-time hydraulic calibration for the WDS. It is noted that the pipe resistance coefficients are not calibrated in a real-time manner as these values are not likely to change over a short time period (Kun et al., 2016).

2.4 Estimate the transfer factor *k* for each FSS manhole

As stated in Eq. (1), the nodal consumption data determined in Section 2.3 (q_i) cannot be directly taken as the manhole inflows (d_i) due to the inevitable loss during the transporting process within the facilities of the users (Behzadian and Kapelan, 2015). In this study, a transfer factor k_i is used to represent the proportion of water consumption used by node i that has been collected by its corresponding manhole. Such a factor can vary as a function of the properties of the water users, such as user types (commercial users or common resident users) and habits of water usages (Bailey et al., 2019). Therefore, the transfer factor needs to be calibrated for each demand node based on the nodal water consumption data and the sewer observations (e.g., sewer flow rates or water depth in the manholes) in the FSS. In this study, the transfer factor k_i associated with each demand node is considered to be approximately constant over time because the user properties are overall constant over a short time period (Bailey et al., 2019).

To calibrate the transfer factor $\mathbf{K} = [k_1, k_2, ...k_n]^T$ of the entire FSS with a total of n manholes with external inflows, the following objective function is defined,

$$\text{Min}: F(\mathbf{K}) = \sum_{t=T_w}^{T} \left(\sum_{i=1}^{M} [g(h_i^o(t)) - g(h_i^s(t))]^2 + \sum_{j=1}^{N} [g(f_j^o(t)) - g(f_j^s(t)]^2 \right) \tag{9}$$

$$[\mathbf{h}_i^s, \mathbf{f}_j^s] = [h_i^s(t_1), h_i^s(t_2), ..., h_i^s(T); f_j^s(t_1), f_j^s(t_2), ..., f_j^s(T)] = F_s(\mathbf{D}(T)) \tag{10}$$

$$\mathbf{D}(T) = \begin{bmatrix} d_1(t_1), d_2(t_1), \ldots d_n(t_1) \\ d_1(t_2), d_2(t_2), \ldots d_n(t_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ d_1(T), d_2(T), \ldots d_n(T) \end{bmatrix} \tag{11}$$

$$d_i(t) = k_i \times q_i(t) \tag{12}$$

$$k_i^{\min} \le k_i \le k_i^{\max},\ \ i = 1,2,\ldots,n \tag{13}$$

where T is the time period with observations used for FSS calibration; T_w is the warming-up time period for model setting up (Guo et al., 2020); *M* and *N* are the numbers of observed water depths at the manholes and flow rates in the sewer pipes, respectively; $h_i^o(t)$ and $f_j^o(t)$ are observed water depth at manhole *i* and observed flow rate at sewer pipe *j* at time *t* respectively; $h_i^s(t)$ and $f_j^s(t)$ are simulated water depth at manhole *i* and simulated flow rate at sewer pipe *j* at time *t* respectively; $g()$ is a linear function used to convert water depths and pipe flow rates into the same scale, thereby enabling both terms in the right side of Eq. (9) are approximately equivalent in terms of the objective function value.

$\mathbf{h}_i^s = [h_i^s(t_1), h_i^s(t_2), \ldots, h_i^s(T)]$ is a vector representing the simulated water depths of manhole *i* over the entire time period of *T*; $\mathbf{f}_j^s = [f_j^s(t_1), f_j^s(t_2), \ldots, f_j^s(T)]$ is a vector representing the simulated sewer flow rates of pipe *j* over the entire time period of *T*; $\mathbf{D}(T)$ is a $T \times n$ matrix, representing the inflows of all manholes across the total time period of *T*. The values of $\mathbf{h}_i^s$ and $\mathbf{f}_i^s$ are computed using $F_s(\mathbf{D}(T))$ as shown in Eq. (11). In this study a simulation package called Storm Water Management Model (SWMM, Rossman, 2010) is employed to calculate $\mathbf{h}_i^s$ and $\mathbf{f}_i^s$. In Eq. (12), $d_i(t)$ is the inflow rate of manhole *i* at time *t*, and $q_i(t)$ is the water consumption of node *i* at time *t* determined by real-time WDS modelling as described in Section 2.3. $k_i^{\min}$ and $k_i^{\max}$ are the minimum and maximum allowable values of k_i, which can be determined by engineering experience. In this study, $k_i^{\min} = 0.7$ is used for each demand node of the WDS, and $k_i^{\max} = 1.0$ is used for each WDS node with smart demand meters, but $k_i^{\max} = 1.3$ is used for WDS nodes without smart meters. This is because water consumptions of nodes without smart meters are calibrated using the method described in Section 2.3, and hence the identified values can inevitably deviate from the true water consumption values at a certain extent. To mitigate this potential impact, the maximum value of the transfer factor for these nodes is increased to 1.3. In this paper, an evolutionary algorithm (EA, Zheng et al., 2017) combined with the SWMM package is employed to solve the optimization problem defined in Eq. (9-13). While different EAs are available in literature, Borg (Hadka and Reed, 2013) is used in this study due to its well-demonstrated performance in dealing with complex water resources optimization problems, with more algorithm details in Section 3.2.

In the proposed FSS calibration method, manhole inflows are considered as the only calibration parameters due to their large temporal and spatial variations, with which the transfer factor k for each manhole can be estimated. It should be noted that Manning's roughness coefficients of the sewer pipes can also affect the hydraulics of the FSS. However, previous studies have shown that the impacts of the small to moderate variation in Manning's roughness coefficients of sewer pipes are limited (Rossman and Huber, 2017). In addition, the physical pipe properties (e.g., pipe ages and materials) that affect the Manning's roughness coefficient are unlikely to vary in a short time period (Zhang et al., 2018) and hence it is not considered within the real-time FSS modelling. It is highlighted that the values of

$\mathbf{K} = [k_1, k_2, ...k_n]^T$ are calibrated using a particular time period of historical water consumption data and in-sewer observations in an offline manner (carried out once as shown in Figure 2).

2.5 Model the FSS in real-time

It is noted that Phases 1-3 in Sections 2.2-2.4 are carried offline (in the offline module as shown in Figure 2), aimed to identify the transfer factors between the WDS nodal water consumptions and the FSS manhole inflows. This is followed by the real-time FSS modelling (real-time module of the proposed method in Figure 2) with the following steps.

Step 1: Collect the data from pressure sensors, flow meters and the available smart demand meters in the WDS at current time t,

Step 2: Estimate the water consumption for each WDS node without smart demand meters, q_i(t) in Equation (1), using the method described in Section 2.3 (Phase 2) based on the observations from Step 1.

Step 3: Update the manhole inflow d_i(t) based on q_i(t) and the identified transfer factor k in Phase 3 of the offline module using Equation (1).

Step 4: Run the FSS hydraulic model based on the manhole inflow d_i(t), producing the water depths and sewer flows for the entire FSS within the time resolutions (30 minutes in this study). This is followed by moving to Step 1 at $t = t + \Delta t$ where Δt is the time resolution of the FSS modelling (Δt =30 minutes in this study).

2.6 Metrics used for performance evaluation

Five statistical metrics are used in this paper to evaluate the performance of the proposed method in simulating the FSS hydraulic variables. They are the absolute percentage error (*APE*), the mean absolute percentage error (*MAPE*), the coefficient of determination (R^2), the Nash-Sutcliffe model efficiency (*NSE*), and the Kling-Gupta Efficiency (*KGE*). These five metrics are selected due to their wide applications in assessing the model performance within the water resources domain (Mu et al., 2020). The *APE* between the i^{th} observation Y_i and its corresponding simulation $\hat{Y}_i$ is defined as

$$APE = \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\% \tag{14}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\% \tag{15}$$

where n is the total number of data points. As shown in Eq. (14) and (15), a lower value of *APE* or *MAPE* indicate an overall better model performance. The R^2 is a goodness-of-fit measure for linear regression models, which can be mathematically described as (Gujarati et al., 2009):

$$R^2 = \frac{\left(\sum_{i=1}^{n} \left(Y_i - \tilde{Y} \right) \left(Y_i - \bar{Y} \right) \right)^2}{\sum_{i=1}^{n} \left(Y_i - \tilde{Y} \right)^2 \sum_{i=1}^{n} \left(Y_i - \bar{Y} \right)^2} \tag{16}$$

where $\bar{Y}$ represents the mean of the observations and $\tilde{Y}$ is the mean of the simulations. A large value of R^2 represents a better model performance. The *NSE* is defined as follows (Nash and Sutcliffe, 1970), with a larger value implying a better model performance:

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} \tag{17}$$

The *KGE* metric is mathematically described as follows (Knoben et al., 2019):

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \tag{18}$$

where r is the linear correlation between observations and simulations; σ_{sim} and σ_{obs} are the standard deviation in simulations and observations, respectively; μ_{sim} and μ_{obs} are the mean of simulations and observations, respectively. A large value of *KGE* means that the simulations can match observations better, with *KGE*=1 representing the best model performance.

3. Case studies

3.1 Case study description

Two real-world FSSs in China, the Benk network (BKN) and the Xiuzhou network (XZN), are selected as case studies to demonstrate the utility of the proposed method. These two FSSs are selected as their associated WDSs have good coverage of monitoring sensors, especially the smart demand meters. In addition, BKN and XZN respectively represent scales of a relatively small region and a town, aimed to demonstrate the utility of the proposed method in handling the FSSs with different complexity levels.

BKN consists of one outlet, 64 manholes and 64 sewer pipes (Figure 4), delivering the wastewater for the users with water supplied by a WDS (referred to as WDS-BNK). WDS-BNK is composed of one reservoir, 65 nodes and 93 pipes, as well as one flow meter, three pressure sensors and 40 smart water demand meters (Figure 4), providing approximately 4,800 m^3 of water per day. As shown in Figure 4, one in-sewer flow meter and three water depth sensors with a 30-minute time resolution have been installed in BKN, with an average discharge of about 4,100 m^3/day. The dotted arrow lines in Figure 4 represent the receiving manhole for each demand node determined based on the spatial distances. Observations from the WDS-BNK and BNK sensors are recorded for consecutive 31 days without rainfall or snowfall events in winter with a 30-minute time resolution.

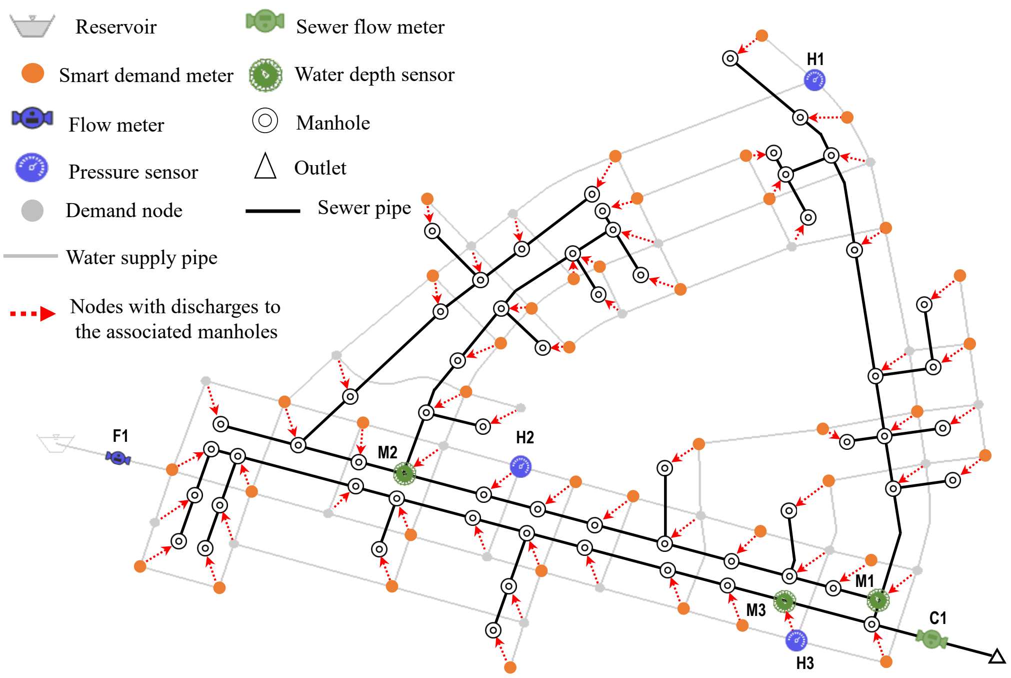

Fig. 4 The layout and sensor locations of the BKN case study and its corresponding WDS-BKN

The XZN system is a large-scale complex FSS in Jiaxing City, with a total length of approximately 86 km and an average discharge of about 21,500 m^3/day. The layout of the

XZN network is shown in Fig. 5, consisting of one outlet, 1,214 manholes and 1,214 sewer pipes. As shown in Fig. 5, three flow meters and eight water depth sensors have been installed in this FSS. The WDS that supplies water demands for this area (refereed as WDS-XZN) has one reservoir, one pump station, 1,119 nodes and 1,137 water consumption pipes as shown in Fig. 6. In the WDS-XZN network, five flow meters, eight pressure sensors and 525 smart demand meters are deployed as illustrated in Fig. 6. The WDS-XZN network supplies approximately 23,150 m^3 per day for a population about 107,500 living in this area within the Jiaxing City. As the same for the BKN network, the data from the WDS-XZN and XZN sensors are recorded for consecutive 31 days without rainfall or snowfall events in winter with a 30-minute time resolution.

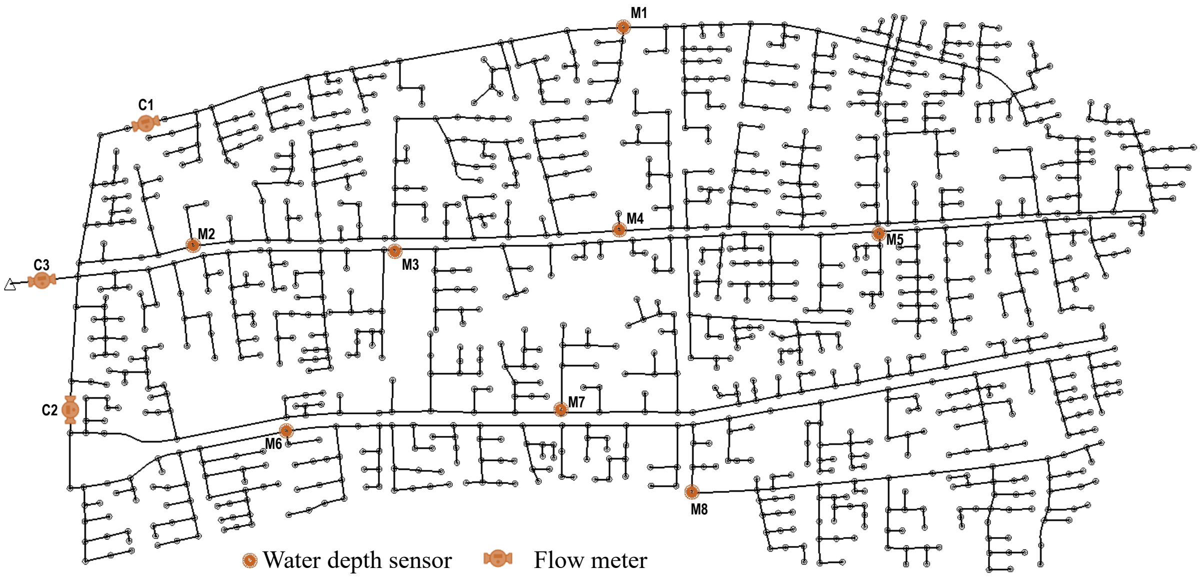

Fig. 5 The layout and sensor locations of the XZN case study

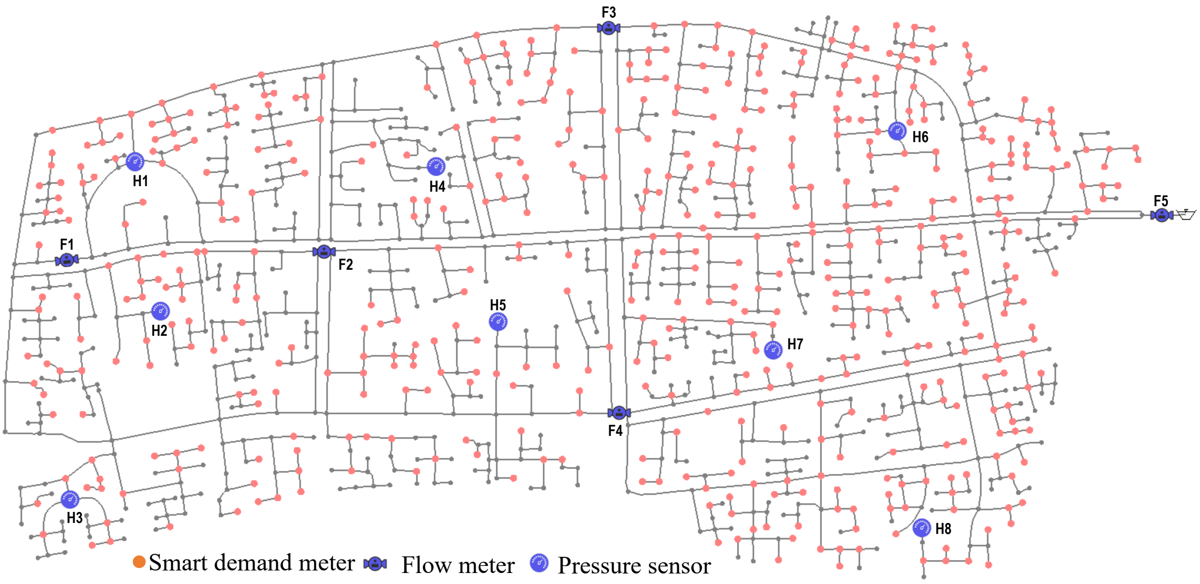

Fig. 6 The layout and sensor locations of the WDS-XZN

3.2 Implementation of the proposed method

The EPANET2.0 and SWMM5.1 (Rossman, 2000, 2010) were used as WDS and FSS hydraulic simulation model respectively in this study. For both case studies, historical data of the first 17 consecutive days from WDS sensors with a 30-minute time resolution were used to estimate the water consumptions of nodes without smart meters. This led to a total of 816 $(17\times24\times2)$ time periods with nodal water consumptions to be calibrated for each WDS. These estimated nodal water consumption data were subsequently used to identify the transfer factors *k* of the FSS based on sewer observations at the first 17 days.

The WDS and FSS sewer observations of the remaining 14 days ($14\times24\times2$ data points used for model validation) were used to run the real-time FSS models with a 30-minute time resolution. In other words, the first set of WDS observations at the validation period (the last 14 days) was considered as the observations at time *t* in the real-time module of Figure 2 (Δt

=30 minutes), followed by the execution of the four steps in Section 2.5.

For the nodal water consumption calibration, the termination error was set as $\max(\|\Delta\mathbf{q}\|) \leq 0.1$ (Eq. 6), the maximum allowed iterations was $S = 100$ (Eq. 6), and the adjustment range of nodal water consumptions was $p = 20\%$ for each WDS (Eq. 7). For the WDS-BNK (Figure 4), observations of the first 17 days from two pressure sensors (H1 and H3) and the flow meter F1 were used for calibration, and the records of pressure sensor H2 were used for validation. For the WDS-XZN (Figure 6), observations of the first 17 days from H1, H3, H4, H6 and H8 pressure sensors, as well as F2, F3, F4 and F5 flow meters were utilized for model calibration, and the records of H2, H5, H7 and F1 were used for validation. The first three days were considered as the warming-up time period for the FSS model setting up as stated in Eq. (9), i.e., T_W=3 days. The observations of the next 14 days were used for FSS model calibration, and the remaining observations of 14 days were utilized for validating the performance of the real-time FSS models. The linear scale function $g()$ in Eq. (9) for each case study is defined as

$$g(x) = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{17}$$

where x represents the observed or simulated values at monitoring points; $x_{\min}$ and $x_{\max}$ are the lower and upper bounds, respectively. These two parameters for each monitoring point are determined by analyzing historical observation data over 14 days (i.e., the calibration time period) in this paper.

The evolutionary algorithm Borg (Hadka and Reed, 2013) was selected to solve the proposed calibration problem defined in Eq. (9-13) due to its great performance in handling complex urban water resources and engineering optimization problems (Reed et al., 2013, Zheng et al., 2016). The initial population size of Borg applied to BKN and XZN case studies were 500 and 1,000 respectively, and the maximum allowable iterations are 50,000 for both case studies. The default values of the other parameters of Borg were used in this study as they have been validated through various applications (Wang et al., 2014). Five Borg runs with different random number seeds were applied to each case study, and the results showed that the final optimization results were overall similar across different runs. Therefore, the results of a typical Borg run were presented to enable discussions for each of the two FSS case studies.

4. Results and discussions

4.1 Calibration results of WDS nodal water consumptions

For each FSS case study, nodal water consumptions of its associated WDS need to be calibrated at each time period, resulting in a total of 816 calibration runs using the calibration method as described in Section 2.3. The resultant time consumption was approximately 25 seconds and 10 minutes for WDS-BKN and WDS-XZN systems, respectively, on a PC with a 2.60-GHz Intel Core i9-7980XE and 2 GB of RAM. Fig. 7 shows the density plot of the errors between observations and simulations at the monitoring locations for both case studies. It is seen that, for the WDS-BKN case, more than 90% of the absolute errors (*AEs*) is less than 0.30m for each pressure monitoring point (including the H2 sensor used for model validation), with the maximum *AE* being 0.32m across the three pressure monitoring points. In terms of flow, more about 93% of absolute percentage errors (*APEs*) are smaller than 1.5%, with the maximum *APE* being 2.40% as shown in Fig. 7(a, b). For the WDS-XZN (Fig. 7(c, d)), the differences between the simulated and observed pressure values at the eight monitoring locations are negligible (including H2, H5 and H7 used for validation), with all *AEs* being lower than 0.4m. Relative to pressure, the deviations between the flow simulations and observations are slightly larger (Figure 7(d)), with the majority of *APEs* smaller than 5% and the maximum *APE* being 9.8% (F1 used for validation).

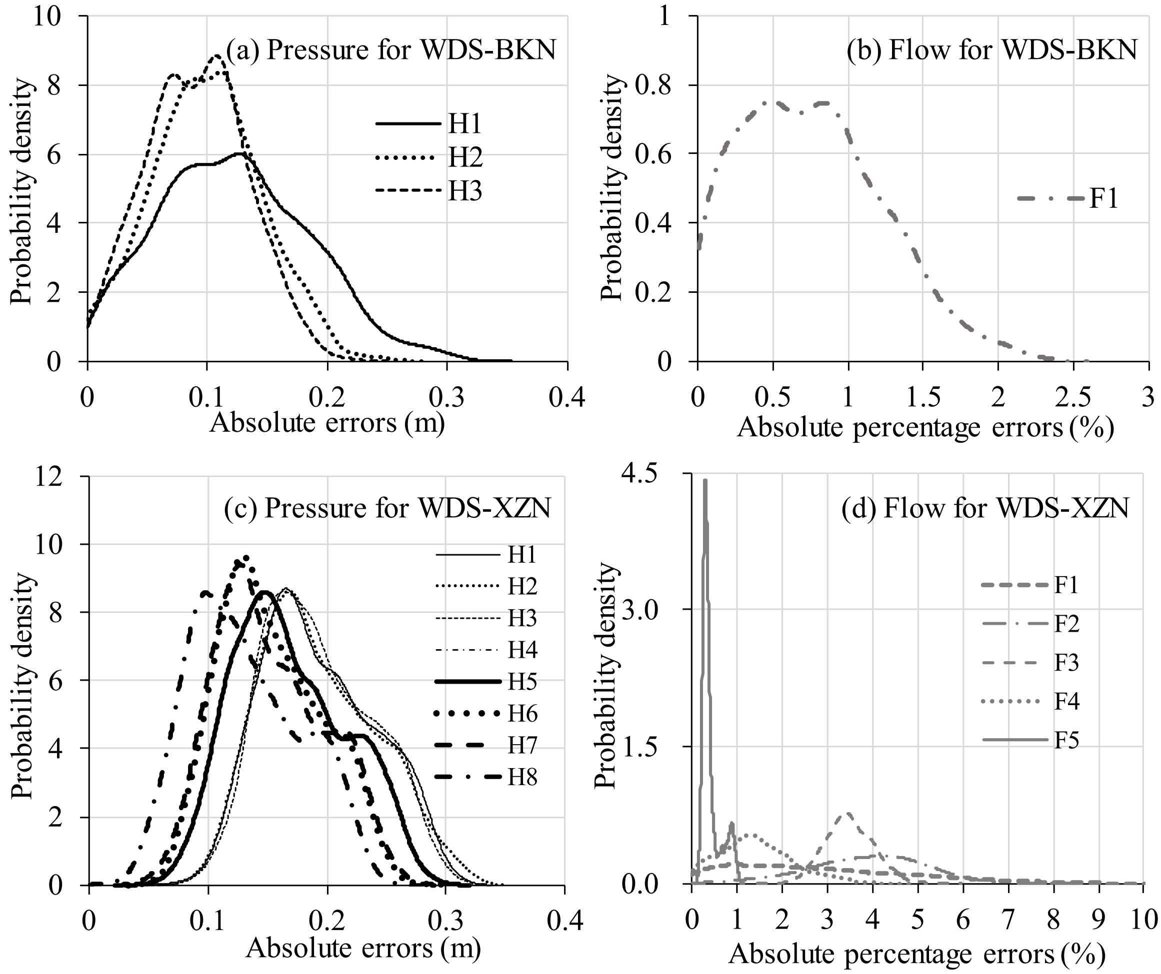

Fig. 7 The probability density distribution of the errors between observations and simulations for all monitoring locations of the WDS-BKN and WDS-XZN

To further demonstrate the quality of calibration results, the criteria defined in Walski et al. (2003) were used to verify the simulation accuracies. As stated in Walski et al. (2003), a satisfactory WDS model calibration should ensure 85% of pressure errors within ±0.2m, 100% of pressure errors within ±0.5 m, trunk main flow errors (flows more than 10% of the total demands) within ± 5%, and the other flow errors within ± 10%. The calibration results of the two WDSs satisfied these criteria, implying that the calibration was successful as the resultant nodal water consumptions can reproduce the overall hydraulics of the WDS. Figure 8 presents the nodal water consumptions over the 31 days with a resolution of 30-minute for the two WDSs ((a) for the WDS-BKN and (b) for WDS-XZN), where the grey solid lines represent the calibrated nodal water consumptions and the orange dotted lines indicate the nodal water consumptions measured by smart demand meters. Despite some variations, all the nodal water consumptions exhibited an overall similar trend for both WDSs, with two peak demand periods occurring at each demand node as shown in Fig. 8, which matches well with the typical water use properties (Zhang et al., 2018).

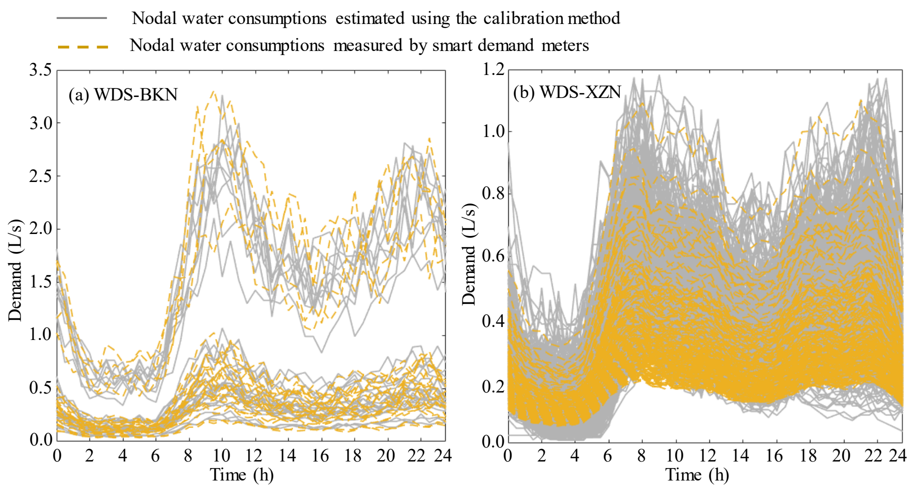

Fig. 8 Nodal water consumptions of the two WDSs at a typical day with a 30-min time resolution

4.2 Estimated transfer factor values

Figure 9 shows the distribution of the probability density of the identified transfer factor *k* values for all manholes of the BKN and XZN based on the historical data over the first 17 calibration days (observations of the first three days were used as model setting-up). Such an optimization (Section 2.4) took 4.86 and 56 hours respectively based on the same computing platform as mentioned above. It can be seen that the majority of *k* values is within the range of 0.7~1.0 for the BKN and XZN, with a mean value of 0.83 and 0.92 respectively, meaning that around 83% and 92% of the total water consumptions have been collected by the FSS of BKN and XZN in this area, respectively. This demonstrates that the calibrated *k* values for all manholes were overall practically meaningful (Behzadian and Kapelan, 2015).

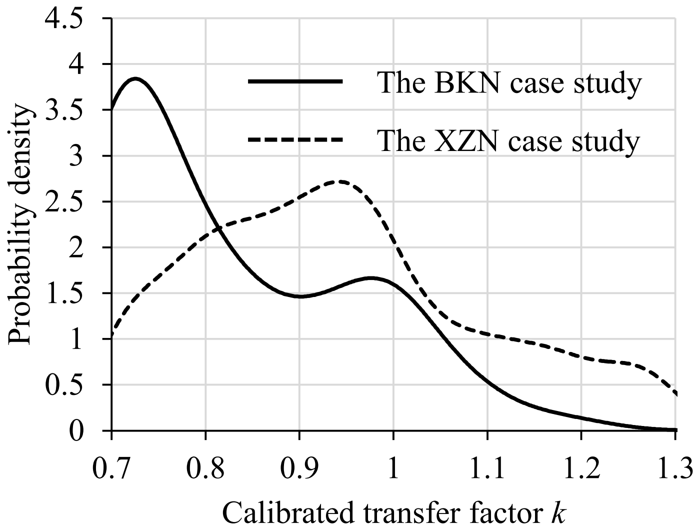

Fig. 9 The density probability distribution of the identified transfer factor *k* for the BKN and XZN case study

It is noted that around 10% and 28% calibrated *k* values were greater than 1 as shown in Figure 1. Such values were only allowed for the WDS nodes without smart demand sensors, and hence their nodal water consumptions were estimated using the calibration method described in Section 2.3. While Fig. 7 showed that the calibration results can reproduce the overall hydraulics of the WDS at the monitoring locations, the calibrated nodal water consumptions might inevitably deviate from the true values at a certain extent (Zhang et al., 2018). To mitigate this potential impact, the value of *k* for the FSS manholes associated with WDS nodes without smart demand meters was allowed to have a range between 0.7 and 1.3, as previously stated. This led to that a proportion of *k* values were greater than 1 as shown in Fig. 9.

Fig. 10 shows the FSS calibrated results (the first 17 days) corresponding to the transfer factor values presented in Figure 9. It is seen that the simulated flows in C1 in the small BKN case study matched well with the observations (Fig. 10(a)), where all *APE* values were lower than 5.0% and the mean *APE* value was 1.16%. For the XZN case study (Fig. 10(e)), the maximum and the mean *APE* values between simulations and observations within the calibration period at the C3 monitoring location were 13.68% and 3.02% respectively. Therefore, it can be deduced that the simulations matched well with the observations for such a large XZN case study. While the *APE* values at the period with relatively low sewer flows were relatively large, their corresponding absolute errors (*AE*s) were overall low as shown in Figure 10 (c,f). For example, the maximum *AE* value was 1.88 L/s for the BKN case study with an average flow of 48.47 L/s in C1 (Fig. 10(c)). Similarly, the maximum *AE* value was 30.51 L/s for the XZN case study with an average flow of 255.10 L/s in C3 (Fig.10(f)).

Tables 1 and 2 present the values of the performance metrics applied to the simulations and observations at monitoring locations for both case studies. As shown in this table, for the BKN case study, the averaged values of *MAPE*, R^2, *NSE* and *KGE* over four different monitoring locations within the calibration period are 3.61%, 0.99, 0.94 and 0.94 respectively. For the large XZN case study, the averaged values of *MAPE*, R^2, *NSE* and *KGE* over 11 different monitoring locations within the calibration period are 4.98%, 0.98, 0.89 and 0.93 respectively. This implied that FSS calibration (aimed to estimate the transfer factor) was overall successful.

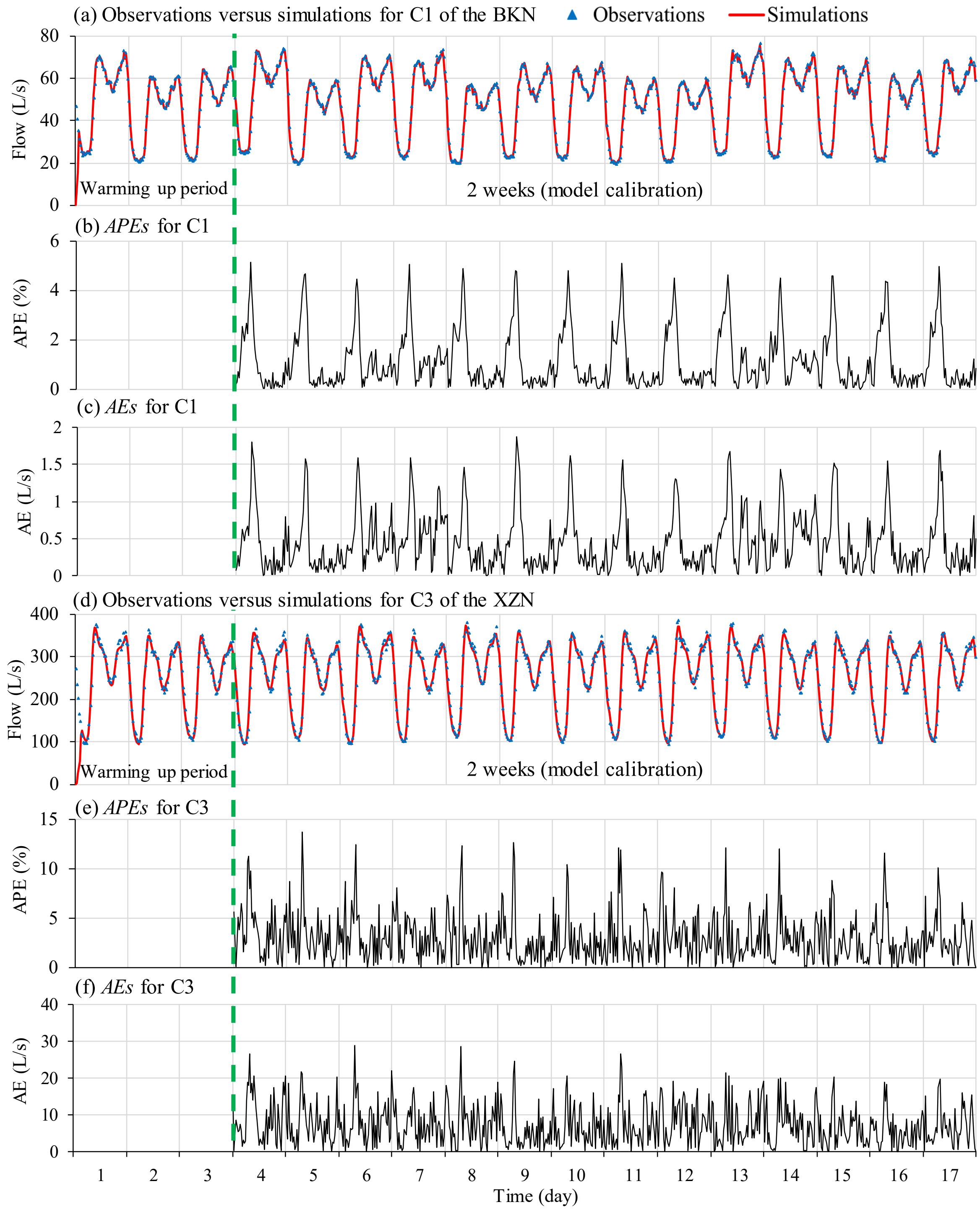

Fig. 10 Observations versus simulations, as well as the APE (%) and AE(L/s) values for C1 in the BKN, and C3 in the XZN within the calibration period (the first 17 days), where C1 and C3 are shown in Fig. 4 and 5 respectively

Table 1 Values of the performance metrics applied to the simulations and observations within the validation period for the BKN case study

Sensor ID	Calibration period				Validation period			
	MAPE(%)	R^2	*NSE*	*KGE*	*MAPE*	R^2	*NSE*	*KGE*
M1	3.18	0.99	0.97	0.93	3.20	0.99	0.96	0.93
M2	2.06	0.99	0.97	0.93	2.12	0.99	0.97	0.92
M3	8.05	0.99	0.84	0.89	8.04	0.99	0.84	0.88
C1	1.16	0.99	0.99	0.99	1.15	0.99	0.99	0.98
Average	**3.61**	**0.99**	**0.94**	**0.94**	**3.63**	**0.99**	**0.94**	**0.93**

Table 2 Values of the performance metrics applied to the simulations and observations within the validation period for the XZN case study

Sensor ID	Calibration period				Validation period			
	MAPE (%)	R^2	*NSE*	*KGE*	*MAPE (%)*	R^2	*NSE*	*KGE*
M1	7.85	0.98	0.78	0.87	7.79	0.97	0.77	0.87
M2	6.49	0.97	0.83	0.90	6.81	0.96	0.80	0.88
M3	6.82	0.98	0.81	0.89	6.43	0.96	0.81	0.89
M4	8.09	0.98	0.77	0.87	8.08	0.97	0.77	0.87
M5	2.83	0.98	0.95	0.96	3.33	0.96	0.94	0.96
M6	7.09	0.97	0.79	0.89	7.07	0.96	0.79	0.88
M7	3.56	0.98	0.93	0.91	4.23	0.96	0.91	0.91
M8	3.10	0.97	0.95	0.92	3.45	0.96	0.94	0.91
C1	2.89	0.99	0.99	0.99	3.20	0.99	0.99	0.99
C2	3.00	0.99	0.99	0.99	3.57	0.98	0.98	0.99
C3	3.02	0.99	0.99	0.99	3.62	0.98	0.98	0.99
Average	**4.98**	**0.98**	**0.89**	**0.93**	**5.23**	**0.97**	**0.88**	**0.92**

4.3 Performance of the real-time FSS modelling

Results in the calibration period demonstrated that the proposed method was capable of identifying suitable transfer factors that can match well simulations and observations at the monitoring locations. This section validated the performance of the real-time FSS models driven by the WDS consumption data in modelling the sewer hydraulics and such a performance evaluation was conducted using the observations from the 17th to the 31st days (i.e., validation period). The steps of the real-time FSS modelling were presented in Figure 2 (real-time module). Figures 11 and 12 show the observations versus observations of the monitoring locations every 30 minutes within the validation period for both case studies.

It is seen from Figure 11 that the sewer flow and the water depth simulations matched well with the observations within the validation period at the four monitoring locations (C1, M1, M2 and M3) in the BKN case study. More specifically, the maximum flow *APE* value was 4.91%, and the maximum absolute error of water depth was 0.7 cm across M1, M2 and M3 locations. Similarly, the differences between the simulation and observations for C1, C2, M1 and M5 monitoring locations were also matched very well for the XZN case study as shown in Fig. 12. For this large FSS, the maximum flow *APE* value was 13.45% and the maximum absolute error of water depth was 1.4 cm (similar observations can be made for other monitoring locations).

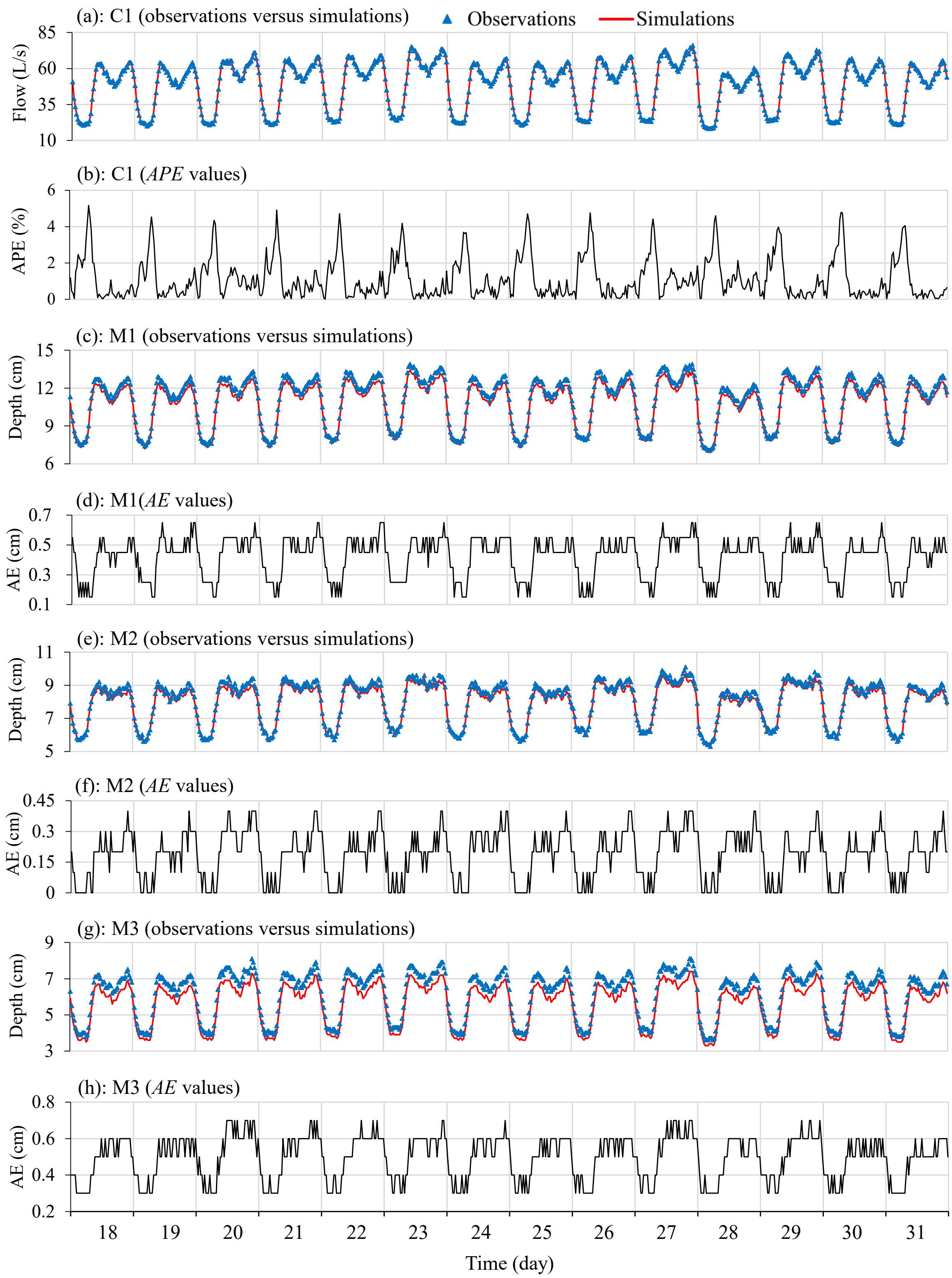

Fig. 11 Observations versus simulations, as well as the APEs or AEs for the four monitoring locations (shown in Fig. 4) within the validation period of the BKN case study

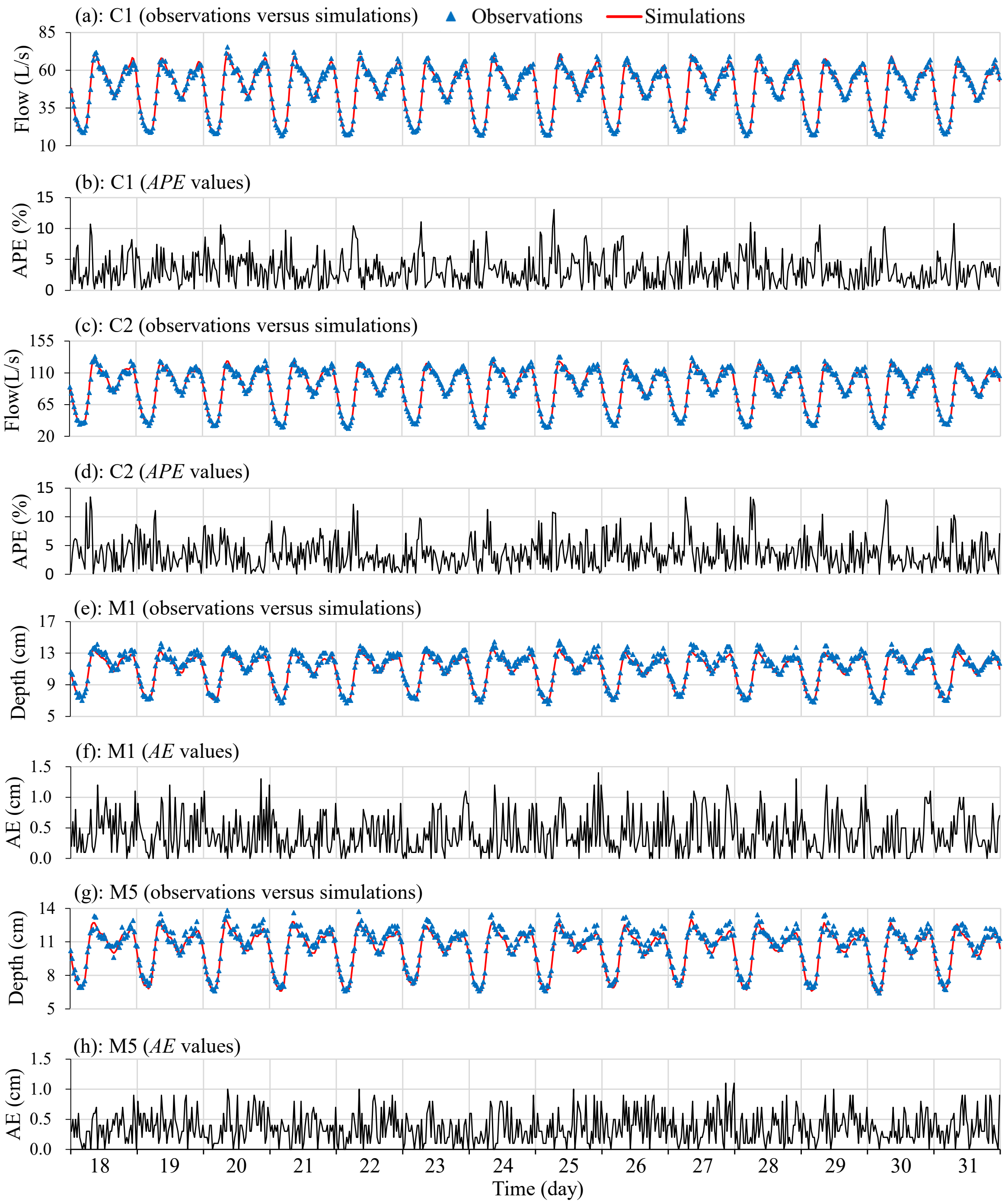

Fig. 12 Observations versus simulations, as well as the APEs or AEs for the four monitoring locations (shown in Fig. 5) within the validation period of the XZN case study

The values of performance metrics applied to the observations and simulations within the validation period for both case studies are also presented in Tables 1 and 2 respectively. As shown in these two tables, the averaged values of *MAPE*, R^2, *NSE* and *KGE* over four different monitoring locations within the validation period are 3.63%, 0.99, 0.94 and 0.93 respectively for the BKN case study. The averaged values of *MAPE*, R^2, *NSE* and *KGE* over four different monitoring locations within the validation period are 5.23%, 0.97, 0.88 and 0.92 respectively for the XZN case study. Overall, the performance of the FSS models within the validation period was similar or slightly worse than the calibration period for both case studies (see Tables 1 and 2). This indicated that (i) there was a low likelihood of over-fitting within the calibration process due to the similar performance between the calibration and validation period, and (ii) the real-time FSS models driven by WDS water consumption data were effective in accurately simulating the sewer hydraulics at a high time resolution (very 30 minutes).

The real-time model was able to offer a great opportunity to enable the comparison between the simulations and observations at monitoring locations at a very high time resolution (every 30 minutes in this paper), followed by a warning trigger if large deviations between the simulations and observations were observed. More specifically, a threshold can be determined by long-term historical data for each monitoring location as did in Qi et al. (2018). If the deviations between the simulations and observations at a particular monitoring location go beyond the specified range, a warning can be triggered efficiently. It should be highlighted that since the real-time FSS model developed using the proposed method has already accounted for the inflow variation caused by the change in water consumption, the false warning rate is expected to be significantly reduced. Therefore, the proposed real-time FSS model can be a useful tool for the development of an efficient warning system, aimed to detect the potential hydraulic issues (e.g., leaks and illicit inflows) for the FSSs.

In addition to providing accurate simulations at the monitoring locations, the proposed method was also able to produce real-time simulations for the manholes and sewer pipes without monitoring sensors. While the accuracies of these simulations cannot be directly evaluated due to the unavailability of observations, it can be anticipated that they can reasonably represent the true hydraulics of the manholes and sewer pipes without monitoring sensors. This was because the real-time FSS model was driven by the water consumption data from the water distribution system, where nodal water consumptions were either measured by smart demand meters or estimated with the aid of an intensive sensor (pressure and flow sensors) coverage. As shown in Figure 13, water depths of 10 manholes near M5 sensor of the XZN case study over a typical day within the validation period exhibited a similar and reasonable trend. These accurate hydraulic simulations at the manholes and pipes without monitoring sensors can be useful to enable the efficient localization of leaks, deposits or illicit inflows, through comparing the simulations with the sampled observations from the field survey.

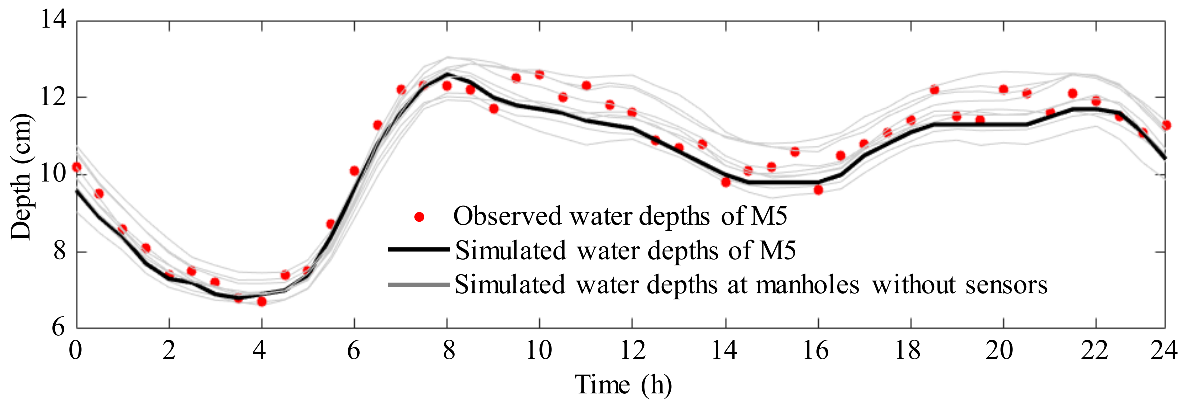

Fig. 13 Water depth simulations and observations of M5, as well as the water depth simulations of 10 manholes near M5 without sensors in the XZN case study in 18th day within the validation period

5. Conclusions

This paper proposes a novel method to develop a real-time foul sewer system (FSS) model driven by water consumption data from its associated water distribution system (WDS) that often has a large number of sensors such as pressure sensors, flow meters and smart demand meters. Within the proposed method, the FSS and the WDS models are integrated to build physical connections between water consumption nodes and their corresponding manholes based on spatial distances. This is followed by a proposal of an optimization approach to identify the transfer factor k between nodal water consumptions and FSS manhole inflows according to historical observations. Subsequently, real-time nodal water consumption data are acquired using an efficient calibration approach based on the dense sensors in the WDS. Finally, these nodal water consumption data combined with the identified k values drive the FSS real-time modelling.

Two real FSS case studies, the smaller BKN with 64 sewer pipes and 64 manholes and the large ZXN case study with 1214 sewer pipes and 1214 manholes have been used to test/validate and demonstrate the proposed method. The results obtained demonstrate that the proposed method can produce real-time predictions of water depths and flows that are in good agreements with the corresponding observations at monitoring locations. The evidence for this can be found in the high mean values of R^2, *NSE* and *KGE* metrics obtained across different monitoring locations, which are 0.99, 0.94 and 0.93 of the small BNK case study, and 0.97, 0.88 and 0.92 for the large XZN case study, respectively. In addition to providing accurate simulations at the monitoring locations, the proposed method is expected to produce reasonable real-time simulations for the manholes and sewer pipes without monitoring sensors. This deduction is based on that the real-time FSS model is driven by the WDS water consumption data that are either measured by smart demand meters or estimated based on a large number of sensors (pressure and flow sensors). This implies that the "equifinality" problem can be successfully addressed by using the proposed method. Therefore, the developed real-time FSS model offers an important tool to facilitate effective and efficient foul sewer system management and operation.

Finally, it is acknowledged that the proposed method is developed ignoring a number of uncertainties that exist in reality. These include potential inaccuracies of WDS and FSS sensor measurements (e.g., smart demand meters, water depth sensors, flow sensors), the potential impacts of the ignorance of the water travelling time within the user property and the influence of the variation in Manning's coefficients of the sewer pipes. These uncertainties need to be more systematically investigated in a future study.

Acknowledgement

This work is funded by the National Natural Science Foundation of China (Grant No. 51922096), and the Excellent Youth Natural Science Foundation of Zhejiang Province, China (LR19E080003).

References

Ahm, M., Thorndahl, S., Nielsen, J.E. and Rasmussen, M.R. (2016) Estimation of combined sewer overflow discharge: a software sensor approach based on local water level measurements. Water Science and Technology 74(11), 2683-2696.

Bailey, O., Arnot, T.C., Blokker, E.J.M., Kapelan, Z., Vreeburg, J. and Hofman, J. (2019) Developing a stochastic sewer model to support sewer design under water conservation measures. Journal of Hydrology 573, 908-917.

Banik, B.K., Di Cristo, C., Leopardi, A. and de Marinis, G. (2017) Illicit intrusion characterization in sewer systems. Urban Water Journal 14(4), 416-426.

Beheshti, M. and Saegrov, S. (2019) Detection of extraneous water ingress into the sewer system using tandem methods - a case study in Trondheim city. Water Science and Technology 79(2), 231-239.

Behzadian, K. and Kapelan, Z. (2015) Modelling metabolism based performance of an urban water system using WaterMet(2). Resources Conservation and Recycling 99, 84-99.

Behzadian, K., Kapelan, Z., Venkatesh, G., Brattebø, H. and Sægrov, S. (2014) WaterMet2: A tool for integrated analysis of sustainability-based performance of urban water systems. Drinking Water Engineering Sciences 7, 63-72.

Black, J. and Endreny, T. (2006) Increasing stormwater outfall duration, magnitude, and volume through combined sewer separation. Journal of Hydrologic Engineering 11(5), 472-481.

Breinholt, A., Grum, M., Madsen, H., Thordarson, F.O. and Mikkelsen, P.S. (2013) Informal uncertainty analysis (GLUE) of continuous flow simulation in a hybrid sewer system with infiltration inflow - consistency of containment ratios in calibration and validation? Hydrology and Earth System Sciences 17(10), 4159-4176.

Broekhuizen, I., Leonhardt, G., Marsalek, J. and Viklander, M. (2020) Event selection and two-stage approach for calibrating models of green urban drainage systems. Hydrology and Earth System Sciences 24(2), 869-885.

Bruen, M. and Yang, J.Q. (2006) Combined hydraulic and black-box models for flood forecasting in urban drainage systems. Journal of Hydrologic Engineering 11(6), 589-596.

Burke, C., Molzahn, R., Pherson, P. and Couts, P. (1986) An evaluation of a sanitary sewer system using a computer model. Tunnelling and Underground Space Technology - TUNN UNDERGR SPACE TECHNOL 1, 153-161.

Creaco, E., Campisano, A., Fontana, N., Marini, G., Page, P.R. and Walski, T. (2019) Real time control of water distribution networks: A state-of-the-art review. Water Research 161, 517-530.

Creaco, E., Signori, P., Papiri, S. and Ciaponi, C. (2018) Peak Demand Assessment and Hydraulic Analysis in WDN Design. Journal of Water Resources Planning and Management 144(6).

Di Pierro, A., Hankin, C. and Wiklicky, H. (2005) Quantitative static analysis of distributed systems. Journal of Functional Programming 15, 703-749.

Du, K., Long, T.-Y., Wang, J.-H. and Guo, J.-S. (2015) Inversion Model of Water Distribution Systems for Nodal Demand Calibration. Journal of Water Resources Planning and Management 141(9).

Eren, B. and Karadagli, F. (2012) Physical Disintegration of Toilet Papers in Wastewater Systems: Experimental Analysis and Mathematical Modeling. Environmental Science & Technology 46(5), 2870-2876.

Garda, A., Castillo, F., Binet, G., Litrico, X. and Gil, A. (2016) Needs and potential of unmanned vehicles in sewers. Houille Blanche-Revue Internationale De L Eau (1), 24-29.

Gujarati, D.N. (2009) Basic econometrics. Tata McGraw-Hill Education.

Guo, D., Zheng, F., Gupta, H. and Maier, H.R. (2020) On the Robustness of Conceptual Rainfall-Runoff Models to Calibration and Evaluation Data Set Splits Selection: A Large Sample Investigation. Water Resources Research 56(3).

Hadka, D. and Reed, P. (2013) Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. Evolutionary Computation 21(2), 231-259.

He, G., Zhang, T., Zheng, F. and Zhang, Q. (2018) An efficient multi-objective optimization method for water quality sensor placement within water distribution systems considering contamination probability variations. Water Research 143, 165-175.

Huang, D., Liu, X.H., Jiang, S.Z., Wang, H.C., Wang, J.Y. and Zhang, Y.K. (2018) Current state and future perspectives of sewer networks in urban China. Frontiers of Environmental Science & Engineering 12(3), 16.

Huang, Y., Zheng, F., Duan, H.-F., Zhang, T., Guo, X. and Zhang, Q. (2019) Skeletonizing Pipes in Series within Urban Water Distribution Systems Using a Transient-Based Method. Journal of Hydraulic Engineering 145(2).

Irvine, K., Maryc, R., Vermette, S., Bakert, J. and Kleinfelder, K. (2011) Illicit discharge detection and elimination: Low cost options for source identification and trackdown in stormwater systems. Urban Water Journal 8(6), 379-395.

Joseph-Duran, B., Ocampo-Martinez, C. and Cembrano, G. (2014) Hybrid modeling and receding horizon control of sewer networks. Water Resources Research 50(11), 8497-8514.

Kapelan, Z. (2002) Calibration of water distribution system hydraulic models.

Khu, S.T., di Pierro, F., Savic, D., Djordjevic, S. and Walters, G.A. (2006) Incorporating spatial and temporal information for urban drainage model calibration: An approach using preference ordering genetic algorithm. Advances in Water Resources 29(8), 1168-1181.

Kleidorfer, M., Leonhardt, G. and Rauch, W. (2012) Identifiability analysis in conceptual sewer modelling. Water Science and Technology 66(7), 1467-1474.

Knoben, W.J.M., Freer, J.E. and Woods, R.A. (2019) Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrology and Earth System Sciences 23(10), 4323-4331.

Koch, M.W. and McKenna, S.A. (2011) Distributed Sensor Fusion in Water Quality Event Detection. Journal of Water Resources Planning and Management 137(1), 10-19.

Korving, H. and Clemens, F. (2005) Impact of dimension uncertainty and model calibration on sewer system assessment. Water Science and Technology 52(5), 35-42.

Lepot, M., Makris, K.F. and Clemens, F. (2017) Detection and quantification of lateral, illicit connections and infiltration in sewers with Infra-Red camera: Conclusions after a wide experimental plan. Water Research 122, 678-691.

Li, T., Zhang, W., Feng, C. and Shen, J. (2014) Performance assessment of separate and combined sewer systems in metropolitan areas in southern China. Water Science and Technology 69(2), 422-429.

Lim, J.S., Kim, J., Friedman, J., Lee, U., Vieira, L., Rosso, D., Gerla, M. and Srivastava, M.B.

(2013) SewerSnort: A drifting sensor for in situ Wastewater Collection System gas monitoring. Ad Hoc Networks 11(4), 1456-1471.

Liu, Y., Tugtas, A.E., Sharma, K.R., Ni, B.-J. and Yuan, Z. (2016) Sulfide and methane production in sewer sediments: Field survey and model evaluation. Water Research 89, 142-150.

McCall, A.-K., Bade, R., Kinyua, J., Lai, F.Y., Thai, P.K., Covaci, A., Bijlsma, L., van Nuijs, A.L.N. and Ort, C. (2016) Critical review on the stability of illicit drugs in sewers and wastewater samples. Water Research 88, 933-947.

Mu, L., Zheng, F., Tao, R., Zhang, Q. and Kapelan, Z. (2020) Hourly and Daily Urban Water Demand Predictions Using a Long Short-Term Memory Based Model. Journal of Water Resources Planning and Management 146(9).

Nash, J.E. and Sutcliffe, J. (1970) River Flow Forecasting Through Conceptual Models: Part 1. — A Discussion of Principles. Journal of Hydrology 10, 282.

Qi, Z., Zheng, F., Guo, D., Maier, H.R., Zhang, T., Yu, T. and Shao, Y. (2018) Better Understanding of the Capacity of Pressure Sensor Systems to Detect Pipe Burst within Water Distribution Networks. Journal of Water Resources Planning and Management 144(7).

Rauch, W., Bertrand-Krajewski, J. L., Krebs, P., Mark, O., Schilling, W., Schütze, M., and Vanrolleghem, P. A. (2001). Mathematical modelling of integrated urban drainage systems. Keynote paper. In Second International Conference on Interactions between sewers, treatment plants and receiving waters-INTERURBA II, Lisbon/Portugal, 89-106.

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R. and Kollat, J.B. (2013) Evolutionary multiobjective optimization in water resources: The past, present, and future. Advances in Water Resources 51, 438-456.

Rokstad, M.M. and Ugarelli, R.M. (2015) Evaluating the role of deterioration models for condition assessment of sewers. Journal of Hydroinformatics 17(5), 789-804.

Rossman, L.A. (2000) users manual. National Risk Management Research Laboratory.

Rossman, L.A. (2010) Storm Water Management Model User's Manual.

Rossman, L.A. and Huber, W. (2017) Storm water management model reference manual volume II–hydraulics. US Environmental Protection Agency, 190.

Sambito, M., Di Cristo, C., Freni, G. and Leopardi, A. (2020) Optimal water quality sensor positioning in urban drainage systems for illicit intrusion identification. Journal of Hydroinformatics 22(1), 46-60.

Sara C. Troutman, Nancy G. Love, Branko Kerkez. (2020) Balancing water quality and flows in combined sewer systems using real-time control. Environmental Science: Water Research & Technology 6 (5), pages 1357-1369.

Savic, D.A., Kapelan, Z.S. and Jonkergouw, P.M.R. (2009) Quo vadis water distribution model calibration? Urban Water Journal 6(1), 3-22.

Schilperoort, R., Hoppe, H., de Haan, C. and Langeveld, J. (2013) Searching for storm water inflows in foul sewers using fibre-optic distributed temperature sensing. Water Science and Technology 68(8), 1723-1730.

Schütze M, Campisano A, Colas H, Schilling W and Vanrolleghem, P. A. (2002). Real-time control of urban wastewater systems-where do we stand today? Global Solutions for Urban Drainage, 1-17.

Schütze, M., Campisano, A., Colas, H., Vanrolleghem, P., and Schilling, W. (2003). Real-time control of urban water systems. In International Conference on Pumps, Electromechanical Devices and Systems Applied to Urban Water Management PEDS, 22-25.

Seco, I., Schellart, A., Gomez-Valentin, M. and Tait, S. (2018) Prediction of Organic Combined Sewer Sediment Release and Transport. Journal of Hydraulic Engineering 144(3).

Sweetapple, C., Astaraie-Imani, M. and Butler, D. (2018) Design and operation of urban wastewater systems considering reliability, risk and resilience. Water Research 147, 1-12.

Talaiekhozani, A., Bagheri, M., Goli, A. and Khoozani, M.R.T. (2016) An overview of principles of odor production, emission, and control methods in wastewater collection and treatment systems. Journal of Environmental Management 170, 186-206.

Walski, T.M., Chase, D.V., Savic, D.A., Grayman, W., Beckwith, S. and Koelle, E. (2003) Advanced water distribution modeling and management.

Wang, Q., Guidolin, M., Savic, D. and Kapelan, Z. (2015) Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front. Journal of Water Resources Planning and Management 141(3).

Wu, Z.Y., Farley, M., Turtle, D., Kapelan, Z., Boxall, J., Mounce, S., Dahasahasra, S., Mulay, M. and Kleiner, Y. (2011) Water loss reduction, Bentley Institute Press.

Wu, Z.Y., Sage, P. and Turtle, D. (2010) Pressure-Dependent Leak Detection Model and Its

Application to a District Water System. Journal of Water Resources Planning and Management 136(1), 116-128.

Xie, X., Zhang, H. and Hou, D. (2017) Bayesian Approach for Joint Estimation of Demand and Roughness in Water Distribution Systems. Journal of Water Resources Planning and Management 143(8).

Zhang, Q., Zheng, F., Duan, H.-F., Jia, Y., Zhang, T. and Guo, X. (2018) Efficient Numerical Approach for Simultaneous Calibration of Pipe Roughness Coefficients and Nodal Demands for Water Distribution Systems. Journal of Water Resources Planning and Management 144(10).

Zheng, F., Tao, R., Maier, H.R., See, L., Savic, D., Zhang, T., Chen, Q., Assumpcao, T.H., Yang, P., Heidari, B., Rieckermann, J., Minsker, B., Bi, W., Cai, X., Solomatine, D. and Popescu, I. (2018) Crowdsourcing Methods for Data Collection in Geophysics: State of the Art, Issues, and Future Directions. Reviews of Geophysics 56(4), 698-740.

Zheng, F., Zecchin, A.C., Maier, H.R. and Simpson, A.R. (2016) Comparison of the Searching Behavior of NSGA-II, SAMODE, and Borg MOEAs Applied to Water Distribution System Design Problems. Journal of Water Resources Planning and Management 142(7).

Zheng, F., Zecchin, A.C., Newman, J.P., Maier, H.R. and Dandy, G.C. (2017) An Adaptive Convergence-Trajectory Controlled Ant Colony Optimization Algorithm With Application to Water Distribution System Design Problems. Ieee Transactions on Evolutionary Computation 21(5), 773-791.