Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Nonlinear Grey-Box Identification with Inflow Decoupling in Gravity Sewers

Krisztian Mark Balla *,** **Carsten Skovmose Kallesøe** *,**
Christian Schou ** **Jan Dimon Bendtsen** *

* *Dept. of Electronic Systems, Aalborg University , Fredrik Bajers Vej 7c, DK-9220 Aalborg, Denmark (e-mail: {kmb,csk,dimon}@es.aau.dk)*
** *Grundfos, Poul Due Jensens Vej 7, DK-8850 Bjerringbro, Denmark (e-mail: {kballa,ckallesoe,cschou}@grundfos.com).*

Abstract: Knowing where wastewater is flowing in drainage networks is essential to utilize system storage, predict overflows and to optimize system operation. Unfortunately, flow in gravity-driven sewers is subject to transport delays, and typically influenced by significant disturbances entering the sewer pipes in the form of domestic, ground and rain inflows. Model-based optimal control of urban drainage requires knowledge about these inflows, even though it is often not feasible in operational setups. To this end, we propose a lumped-parameter hydrodynamic model with a bi-linear structure for identifying the transport delays, decouple periodic disturbances and to predict the discharged flow. Pumped inlet and discharged dry-weather flow is used to find the model parameters. Under mild assumptions on the domestic and groundwater inflows, i.e. disturbances, the decoupling capabilities of the identified model are presented. A numerical case study on an EPA Storm Water Management Model (EPA SWMM) and experimental results on a real network demonstrate the proposed method.

Keywords: Process identification, Transport delay, Disturbance parameters, Open hydraulics

1. INTRODUCTION

In sewers, waste water is collected and transported towards treatment plants, where contaminants are removed before releasing the water back to the environment (Butler and Davies [2006]). Flow routing in sewers is a complex task, since the network is characterized by large spatial dimensions, nonlinear dynamics, large flow variations and significant time delays. In this work, a nonlinear system identification approach is proposed to predict flow and delays based on the simplified Saint-Venant (SV) equations.

Gravity-driven flow in open channels is represented by a system of coupled partial differential equations (PDEs). Due to the complexity of these PDEs, simplified and linearized models are typically used in optimal control design. In Xu et al. [2011], the control effectiveness of reduced SV models in Model Predictive Control (MPC) has been studied. In Leirens et al. [2010], a linearized SV model has been proposed for pumped sewer networks. Moreover, linear cascade modelling is a common modeling approach in open water systems, e.g. in irrigation canals (Litrico and Fromion [2004]), and in inland waterways (Segovia et al. [2018]). Linearization, however, does not allow flow-dependent delays and the maximum allowed flow deviation from the steady-state solution is restricted.

In a previous paper, Kallesøe and Knudsen [2016], self-calibrating flow estimation has been developed for tracking the pump flow at the inlet, and for the discharged flow at the outlet of gravity pipes. This algorithm utilized information about wastewater pits and pump operation, hence applicable in any system configuration. However, prediction of flow dependant delays considering different disturbance inflows (i.e. domestic- and groundwater) has not been encountered yet. The current work utilizes this previously-established flow estimation algorithm as the source of training data, and proposes a modeling method for flow and delay prediction in long gravity sewers. While we do not explicitly address control strategies, the proposed identification approach is dedicated for the internal model of a predictive controller. The main compromise of establishing such a model is typically between complexity, accuracy and the computational burden (Lund et al. [2018]). In recent years, identification-based modeling has gained more attention, as data has become widely available at utilities. Yet, reports on Grey-Box modeling of open-channel water systems are relatively few (Su Ki Ooi and Weyer [2003], Weyer [2001]). Moreover, research on Grey-Box modelling in different domains, e.g. Sundar and Zlotnik [2019], considered large-scale natural gas networks where state and parameter estimation has been developed using data and the underlying graph of the network.

In many applications in this framework (e.g. irrigation canals, sewage networks, etc.), linear physical models are used. In contrast to these approaches, we propose a non-linear model structure which can describe a wide range of flows. The task of finding the correct physical parameters (e.g. length, shape, slope and friction) disappears with the proposed data-driven approach, thus enabling scalability to large systems with arbitrary structures.

Our approach is data-driven, yet we establish our proposed model structure based on physical considerations, familiar to those working in the water domain. In contrast to

Copyright lies with the authors

methods utilizing cross-correlation analysis e.g. for delay detection (Bjorklund and Ljung [2003]), and statistical blackbox models for disturbance prediction (Troutman et al. [2017]), we aim to preserve intuition by giving physical interpretation of the flows in the system.

The remainder of the paper is structured as follows. Section 2 provides an overview of the system, whereupon in Section 3 we review the PDE model for gravity sewers and propose a model for the disturbance flows. In Section 4, we formulate the estimation as a least squares problem in the form of a Nonlinear Program (NLP), while in Section 5, we present numerical and experimental results. Finally, Section 6 sums up the contributions of the work. Throughout the paper, all quantities mentioned are real. We use boldface letters for sets, such as $\boldsymbol{s} = \{s_1, ..., s_n\}$, as well as for vectors $\boldsymbol{x} = [x_1, ..., x_n]^T \in \mathbb{R}^n$.

2. SEWER SYSTEMS OVERVIEW

We consider networks with transport lines over long distances. The layout of such network is shown in Fig. 1.

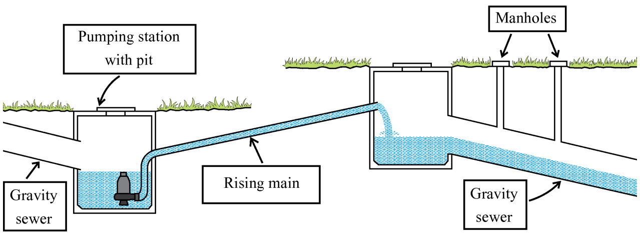

Fig. 1. Topology of a pumped sewer network.

The sewage is first collected at a pumping station. The pumping station consists of a small storage tank (pit) and one or more pumping units. The collected water is then pumped through a rising main, whereupon it enters a long gravity-driven sewer channel. Typically, pumps operate in combination and deliver flow at a fixed rate governed by local on/off controllers (see Schütze et al. [2002]). When the pump state is off and the pit volume reaches the maximum threshold, the pumps turn on. Then, the pumped flow arrives to the next pit in line with a delay, thereby forcing the next station to turn on, and so forth.

Exogenous inflows, i.e. disturbances, enter gravity sewers in the form of domestic waste, rainfall run-off and through leakages allowing groundwater to infiltrate into the channel. These disturbances are characterized by specific flow patterns, shown in Fig. 2.

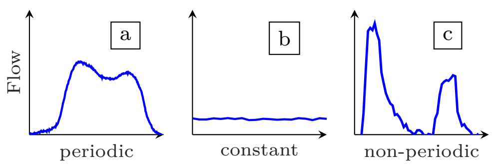

Fig. 2. Disturbances occurring in sewer networks.

Households are common sources of inflows in sewer networks. Due to the large spatial dimension of sewers, domestic waste, collected from urban areas, may enter the network anywhere along the gravity pipes. Domestic waste is typically characterized by a diurnal pattern which has an inherent periodicity of 24 hours with peak points in the morning and in the afternoon (see Pattern *a* in Fig. 2.). Groundwater is present when rain runs off slowly and thereby water accumulates in the ground. In this case, water may infiltrate gravity channels through leakages with a nearly constant flow (see Pattern *b* in Fig. 2.).

The third type of disturbance flow is due to rainfall run-off. These discharges are disregarded and we rather focus on domestic and groundwater infiltration. Indeed, this is a common practice (see Courdent et al. [2018]), as typically the domestic and groundwater flows are estimated first, using inflow measurements from dry-weather periods.

3. MODELING

3.1 Flow Model

Unsteady, one-dimensional water flow in gravity sewers can be computed accurately by the well-known Saint-Venant PDEs (Schütze et al. [2002])

$$\frac{\partial A_{x,t}}{\partial t} + \frac{\partial Q_{x,t}}{\partial x} = q_{x,t}, \tag{1}$$

$$\frac{\partial Q_{x,t}}{\partial t} + \frac{\partial}{\partial x}\left(\frac{Q_{x,t}^2}{A_{x,t}}\right) + gA_{x,t}\left(\frac{\partial h_{x,t}}{\partial x} + S_f + S_b\right) = 0, \tag{2}$$

where $Q_{x,t}$ denotes the flow inside the channel, $q_{x,t}$ is the disturbance inflow, $A_{x,t}$ is the wetted channel area and $h_{x,t}$ is the water level. Moreover, $Q_{x,t}, q_{x,t}, A_{x,t}$ and $h_{x,t}$ are functions from $(0, L) \times \mathbb{R}_+$ to $\mathbb{R}_+$. The gravitational acceleration is $g \in \mathbb{R}_+$, furthermore we assume that $S_b \in \mathbb{R}_+$ bed slope and $S_f \in \mathbb{R}_+$ friction slope parameters are independent of x and t, which is a fair assumption if the slope variance is small (Schütze et al. [2002]).

For simplicity, we assume kinematic waves, meaning that we neglect the first three terms in (2) (Singh [2001]), which results in a balance between friction and gravitational forces. For determining S_f in (2), we utilize Manning's equation (Schütze et al. [2002]). Then, (2) simplifies to

$$S_b = S_f(Q_{x,t}, h_{x,t}) = \frac{n^2 Q_{x,t}^2}{A_{x,t}^2 R_{x,t}^{\frac{4}{3}}}, \tag{3}$$

where $R = \frac{A}{P}$ is the hydraulic radius, $P \in \mathbb{R}_+$ is the wetted perimeter and $n \in \mathbb{R}_+$ is the Manning coefficient. Simplification with the Manning formula restricts the flows to be one-directional, meaning that the phenomena of backwater effect is not considered. Backwater occurs when the channel is overloaded, thus water surcharges. This is typically negligible in large gravity lines, transporting waste water over long distances (Singh [2001]).

We derive the model for rectangular cross section and argue that semi-filled circular sewers are reasonably well-approximated by a rectangular shape, shown in Fig. 3.

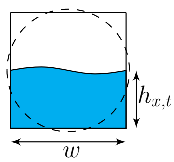

Fig. 3. Rectangular channel, where $w \in \mathbb{R}_+$ is the width.

Hence, the hydraulic radius of a channel is parametrized by the width w and approximated by

$$R_{x,t} = \frac{A_{x,t}}{P_{x,t}} \approx \frac{w h_{x,t}}{2h_{x,t} + w}, \tag{4}$$

where a linear area-level relation is used. The independent variables remaining in the SV equations are flow and level on the domain $(0, L) \times \mathbb{R}_+$, given by

$$w\frac{\partial h_{x,t}}{\partial t} + \frac{\partial Q_{x,t}}{\partial x} = q_{x,t}, \tag{5}$$

$$Q_{x,t} = \frac{\sqrt{S_b}}{n}\frac{(w h_{x,t})^{\frac{5}{3}}}{(2h_{x,t} + w)^{\frac{2}{3}}}. \tag{6}$$

The equation in (6) comes from (3) and (4). The coupled PDEs in equations (5) and (6) describe kinematic waves traveling through open channels. Note, that semi-fillness of the channel is an assumption which does not always hold. However, the formulation presented here does not hold for fully-filled flow conditions.

3.2 Discretized Model

We now formulate the physical model in a form more amenable to system identification. The channel is partitioned into N_x equal-sized, non-overlapping δx segments of length, while $h_{x,t}$, $Q_{x,t}$ and $q_{x,t}$ are approximated as piecewise constant functions of x. The spatial discretization of a channel at time t is shown in Fig. 4.

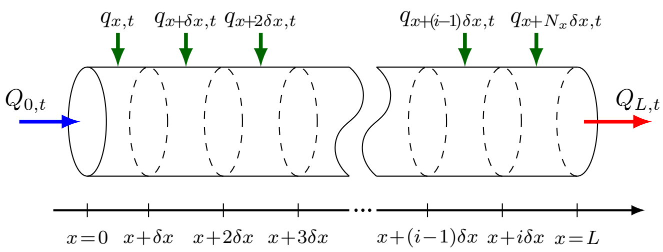

Fig. 4. Channel divided into N_x sections, where $i \in \{1, 2, ..., N_x\}$.

We use backward Euler discretization for the spatial and forward Euler for the time coordinate. The left boundary (upstream) is defined at $x = 0$ and the right boundary (downstream) at $x = L$. Discretizing equations (5) and (6), the model may be recast as

$$h_{x,t+\delta t} = h_{x,t} + \alpha q_{x,t} + \beta_1(Q_{x-\delta x,t} - Q_{x,t}), \tag{7}$$

$$Q_{x,t} = \beta_2 \frac{h_{x,t}^{\frac{5}{3}}}{(h_{x,t} + \gamma)^{\frac{2}{3}}}, \quad \forall\, x \in (0, L) \tag{8}$$

where δt is the sampling time and we define the parameters

$$\alpha \triangleq \frac{\delta t}{w}, \quad \beta_1 \triangleq \frac{\delta t}{w\delta x}, \quad \beta_2 \triangleq \frac{\sqrt{S_b} w^{\frac{5}{3}}}{2^{\frac{2}{3}} n}, \quad \gamma \triangleq \frac{w}{2}, \tag{9}$$

where $\alpha, \beta_1, \beta_2, \gamma \in \mathbb{R}_+$. Note that the time and spatial steps, δt and δx, are part of the parameters α and β_1. This implies that (7) and (8) are affected by the choice of the sampling time and the section size. Indeed, δt and δx affects the dynamics by introducing distortion in the traveling wave (see Singh [2001]). Next, we insert the section flows $Q_{x,t}$ from (8) into (7), so we get

$$h_{x,t+\delta t} = h_{x,t} + \alpha q_{x,t} + \beta_1\beta_2\Bigg[\frac{h_{x-\delta x,t}^{\frac{5}{3}}}{(h_{x-\delta x,t} + \gamma)^{\frac{2}{3}}} - \frac{h_{x,t}^{\frac{5}{3}}}{(h_{x,t} + \gamma)^{\frac{2}{3}}}\Bigg] \tag{10}$$

The state equation in (10) is parametrized by α, β_1, β_2 and γ. In order to reduce the number of parameters, and thus avoid non-identifiability, we attempt to restructure (10) by removing γ from the denominator of the nonlinear expression. Therefore, we define new states such that

$$h_{x,t} \triangleq \gamma z_{x,t}, \quad \forall\, x \in (0, L), \tag{11}$$

where $z_{x,t}$ are the scaled equivalents of the physically measurable water levels $h_{x,t}$, i.e. the transformed states. Now let us define a mapping $g : \mathbb{R}_+ \to \mathbb{R}_+$

$$g : (z_{x,t}) \mapsto \frac{z_{x,t}^{\frac{5}{3}}}{(z_{x,t} + 1)^{\frac{2}{3}}}, \quad \forall\, x \in (0, L). \tag{12}$$

Utilizing the state transformation and the nonlinear mapping g, we recast the state equation in (10). This yields as a set of coupled bi-linear difference equations, describing the states in the N_x partitioned sections along the channel.

$$\begin{aligned}
z_{0,t+\delta t} &= z_{0,t} + \tilde{q}_{0,t} + \theta_1 Q_{0,t} - \theta_1\theta_2 g(z_{\delta x,t}),\\
&\;\;\vdots\\
z_{x,t+\delta t} &= z_{x,t} + \tilde{q}_{x,t} + \theta_1\theta_2\big(g(z_{x-\delta x,t}) - g(z_{x,t})\big), \qquad (13)\\
&\;\;\vdots\\
z_{L,t+\delta t} &= z_{L,t} + \tilde{q}_{L,t} + \theta_1\theta_2\big(g(z_{L-\delta x,t}) - g(z_{L,t})\big),
\end{aligned}$$

where we defined $\tilde{q}_{x,t} \triangleq \frac{\alpha}{\gamma} q_{x,t}$ as scaled disturbance flows. Furthermore, the task of parameter estimation has been reduced to find the parameters $\theta_1, \theta_2 \in \mathbb{R}_+$ and the unknown disturbances $\tilde{q}_{x,t}$. The parameters are given by

$$\theta_1 \triangleq \frac{\beta_1}{\gamma}, \quad \theta_2 \triangleq \beta_2\gamma. \tag{14}$$

We consider the upstream boundary flow $Q_{0,t}$ (hereinafter Q_{in}), as the control input. The output of the control model is the downstream flow at the boundary $x = L$, which we hereinafter call Q_{out}. The output equation is then

$$Q_{\text{L,t}} = \theta_2 g(z_{L,t}), \tag{15}$$

which is the reformulated Manning equation in (8).

3.3 Disturbance model

In this application, we consider periodic domestic waste flows coming from urban areas and groundwater infiltration which we assume to be constant in time. Here, a Fourier series model is presented to estimate the periodic disturbance signals $\tilde{q}_{x,t}$ entering into the dynamic flow model in (13). It is assumed that the infiltration of groundwater is uniformly distributed along the partitioned channel sections, i.e. groundwater enters each section with the same amplitude. Moreover, it is assumed that we know where residential areas are connected by pipelines, i.e. where domestic waste water enters the channel. This is a fair assumption, as typically the piping layout of the infrastructure is stored in, e.g. a GIS (Geographic Information System) database at most utilities. Hence, the scaled disturbances entering the i^{th} section at time t are given by

$$\begin{aligned}\tilde{q}_{i,t}(\lambda_0, \boldsymbol{\lambda}) &\triangleq \tilde{q}_i^{\text{gnd}}(\lambda_0) + \tilde{q}_{i,t}^{\text{dom}}(\boldsymbol{\lambda}) \qquad (16)\\ &\triangleq \lambda_0 + \underbrace{\lambda_1 + \sum_{j=1}^{k}\big(\lambda_{1j}cos(j\omega t) + \lambda_{2j}sin(j\omega t)\big)}_{f_{i,t}(\boldsymbol{\lambda})}\end{aligned}$$

where the set of parameters regarding domestic flows are $\boldsymbol{\lambda} \triangleq \{\lambda_1, \lambda_{11}, \lambda_{21}, ..., \lambda_{1k}, \lambda_{2k}\} \in \mathbb{R}$, and $\lambda_0 \in \mathbb{R}_+$ represents constant groundwater flows. The angular frequency ω corresponds to a period of one day and $k \geq 2$ is the number of frequency terms in the truncated Fourier series. Furthermore, $f_{i,t}(\boldsymbol{\lambda})$ is a family of functions parametrised by $\boldsymbol{\lambda}$, and for each $\boldsymbol{\lambda}$, $f_{i,t} : (1, N_x) \times \mathbb{R}_+ \to \mathbb{R}$.

Note, that the model describing the scaled disturbances does not correspond to the real domestic and groundwater flows. In order to calculate the real disturbances $q_{i,t}$ from the scaled estimates $\tilde{q}_{i,t}$, recall that the disturbance flows are scaled by the model parameters, such that

$$\tilde{q}_{i,t} \triangleq \frac{\alpha}{\gamma} q_{i,t}, \quad \forall\, i \in (0, N_x). \qquad (17)$$

Then, using the estimated parameters defined in (14) and the physical model parameters in (9), it is seen that

$$\sum_{i=1}^{N_x} q_{i,t}\delta x = \sum_{i=1}^{N_x} \frac{\tilde{q}_{i,t}}{\theta_1} = \sum_{i=1}^{N_x} \frac{\alpha}{\beta_1} q_{i,t} = q_{i,t}\delta x N_x, \qquad (18)$$

where we use θ_1 to calculate $q_{i,t}$. The last term in (18) represents all disturbance flows along the total pipe length.

Now, let us consider the function $f_{i,t}(\boldsymbol{\lambda})$ in (16). From this, we form the vector $\boldsymbol{f}_t(\boldsymbol{\lambda})$, such that the function at index i is $f_{i,t}(\boldsymbol{\lambda})$, i.e. $\boldsymbol{f}_t : \mathbb{R}_+ \to \mathbb{R}^{N_x}$. Using the vector representation and the relation shown in (18), we decompose the domestic and groundwater flows such that

$$\sum_{i=1}^{N_x} q_{i,t}\delta x = \underbrace{\frac{\lambda_0}{\theta_1} N_x}_{q^{\text{gnd}}(\lambda_0,\theta_1)} + \underbrace{\frac{1}{\theta_1}\boldsymbol{b}^T \boldsymbol{f}_t(\boldsymbol{\lambda})}_{q^{\text{dom}}(\boldsymbol{\lambda},\theta_1)}, \qquad (19)$$

where q^{gnd} is the ground infiltration flow summed over all pipe sections and q^{dom} is the total domestic inflow. The vector $\boldsymbol{b} \in \{0,1\}^{N_x}$ has ones in its entries regarding section points where urban areas are connected, and zeros with no connection. The number of parameters that we need to identify is $4+2k$, depending on order of the Fourier series.

4. SYSTEM IDENTIFICATION

The system identification is given as a constrained nonlinear least squares problem, where samples, at time t_i, $i = \{0, ..., N_t\}$, of the pumped inlet flows Q_{in} and the discharged Q_{out} flows are known and estimated a priori. Let $\boldsymbol{\theta} \triangleq \{\theta_1, \theta_2\} \in \mathbb{R}_+$ denote the set of system parameters and $\boldsymbol{\Lambda} \triangleq \{\lambda_0, \boldsymbol{\lambda}\} \in \mathbb{R}$ denote the parameters corresponding to the disturbance flows. Then, the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\Lambda}$ and the initial states $\boldsymbol{z}(t_0)$ are found by solving the NLP

$$\begin{pmatrix}\boldsymbol{\theta}^*\\ \boldsymbol{\Lambda}^*\\ \boldsymbol{z}^*(t_0)\end{pmatrix} = \underset{\boldsymbol{\theta},\boldsymbol{\Lambda},\boldsymbol{z}(t_0)}{\operatorname{argmin}} \sum_{i=0}^{N_t} \big(Q_{out}(t_i) - \hat{Q}_{out}(t_i)\big)^2 \qquad (20)$$

s.t.

$$\boldsymbol{z}(t_{i+1}) = \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\Lambda}}(\boldsymbol{z}(t_i), Q_{in}(t_i)), \qquad (21)$$

$$\hat{Q}_{out}(t_i) = H_{\theta_2}(z_L(t_i)), \qquad (22)$$

$$\boldsymbol{0} \leq \boldsymbol{z}(t_i) \leq \overline{\boldsymbol{z}}, \qquad (23)$$

$$\boldsymbol{0} \leq \boldsymbol{\theta} \leq \overline{\boldsymbol{\theta}}, \qquad (24)$$

where $\boldsymbol{z}(t_i) \in \mathbb{R}^{N_x}$ is the vector of states in (13) and the dynamics in (21) are represented by $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\Lambda}}(\boldsymbol{z}(t_i), Q_{in}(t_i)) : \mathbb{R}_+ \to \mathbb{R}^{N_x}$. The function $H_{\theta_2}(z_L(t_i)) : \mathbb{R}_+ \to \mathbb{R}_+$ in equation (22) represents the scalar output where z_L corresponds to the downstream boundary state. The constraints in (23) and (24) impose bounds on the transformed state variables and parameters, respectively.

In the above NLP, we assumed that the number of states are fixed, i.e. N_x is given. Instead, we introduce N_x as an auxiliary variable in the model. Hence, we carry out estimations multiple times on equivalent models but with different grid sizes as explained in the algorithm below

Algorithm 1 Model evaluation for different N_x

Input: Q_{in}, Q_{out}, $\overline{\boldsymbol{z}}$, $\overline{\boldsymbol{\theta}}$
repeat at every iteration k = 1,2,... N_x
  Initialize: $\boldsymbol{\theta}$, $\boldsymbol{\Lambda}$, $\boldsymbol{z}(t_0)$
  Solve: $\min_{\boldsymbol{\theta},\boldsymbol{\Lambda},\boldsymbol{z}(t_0)} \sum_{i=0}^{N_t} \big(Q_{out}(t_i) - \hat{Q}_{out}(t_i)\big)^2$
    s.t. constraints
until $RMSE(\hat{Q}_{out}^{k-1}) < RMSE(\hat{Q}_{out}^{k})$
Output: $\boldsymbol{\theta}^*$,$\boldsymbol{\Lambda}^*$,$\boldsymbol{z}^*(t_0)$, N_x

The model is evaluated for each trial of N_x using Root Mean Squared Errors (RMSE) and the algorithm is terminated at the lowest N_x. The estimation accuracy is shown in Fig. 5. for a selection of k iteration steps.

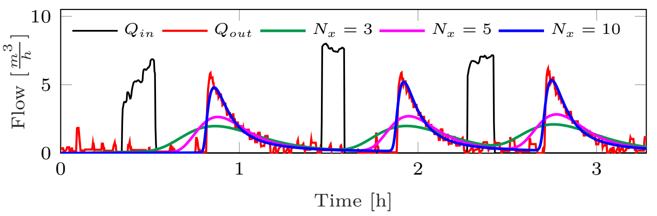

Fig. 5. Model accuracy for different N_x tested on real data.

As seen, increasing N_x above a threshold does not increase accuracy significantly. To evaluate the convergence of Algorithm 1, the RMSE at each iteration is calculated.

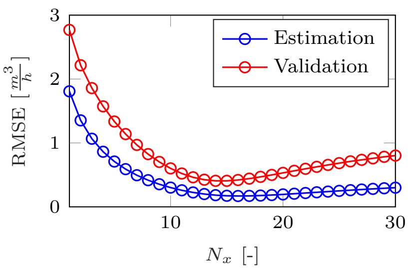

Fig. 6. Estimation and validation evaluated against N_x.

As seen in Fig. 6., there is an optimal selection for N_x where model accuracy is the highest for the provided training and validation datasets.

5. RESULTS

5.1 Numerical results

Results of applying the method are first presented on a numerical case study in the EPA SWMM(Rossman [2015]) simulation software. The test model is depicted in Fig. 7.

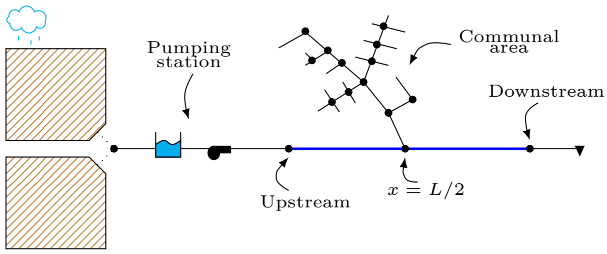

Fig. 7. Schematics of the EPA SWMM simulation model.

In this network, a single sewer line is considered, transporting the sewage from a pumping station to an outlet point, representing either the next pumping station or the treatment plant. We consider an urban area discharging to the transportation line at $x = L/2$. The pumped inlet flows Q_{in} enter the sewer at the upstream and we observe the discharged flows Q_{out} at the downstream, indicated in Fig. 7. Moreover, measurement noise is added to the simulated $Q_{out}(t_i)$ flows with the property of $n \sim \mathcal{N}(0, 0.2)$.

In simulation, we attempt to mimic the behavior of a real scenario, where the wastewater pit collects non-periodic runoff water with a variety of rain intensity, forcing the pump to turn on for different time duration. The validation of the identified model is shown in Fig. 8.

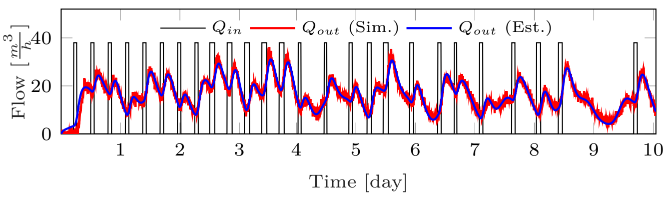

Fig. 8. Discharged downstream flow prediction.

$N_x = 4$ sections resulted in good model accuracy. Note, that the discharged flow shown in Fig. 8. consists of the delayed non-periodic pumped flows Q_{in} and the periodic disturbance inflows q. In EPA SWMM, we can access the q periodic disturbances for validating our results. The disturbance q entering the network at $x = L/2$ and then discharged at the end of the channel are shown in Fig. 9.

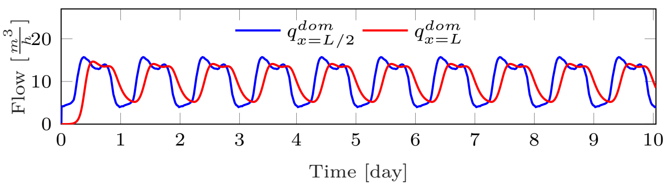

Fig. 9. Upstream and discharged domestic waste flows.

The domestic waste flow arrives at the downstream with a flow-dependent delay. Using the estimated parameter θ_1 and the disturbance model defined in Section 4, the decoupled domestic disturbance flow $q^{dom}_{x=L/2}$ yields as shown in Fig. 10.

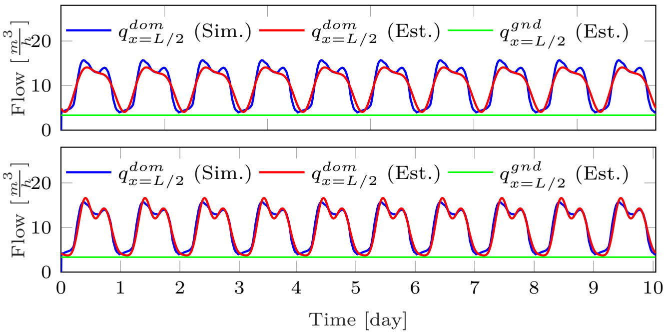

Fig. 10. Disturbance decoupling with $2nd$ (above) and $4th$ order Fourier series (below).

Fig. 10. shows that we are able to estimate the periodic domestic waste flows and the constant groundwater infiltration at the point where they enter the system, by using the disturbance model and the identified model parameters. The disturbances q^{dom} and q^{gnd} in Fig. 10. correspond to the decoupled flows in (19). With higher order Fourier series, the estimation is more precise, however at the cost of increasing the number of parameters.

5.2 Experimental results

We also present results of applying the system identification method on a real world case study. The available data is flow estimation, extracted from a sewer network, operated by Provas A/S, located in Gram, Denmark. The pipe layout of the drainage network is shown in Fig. 11.

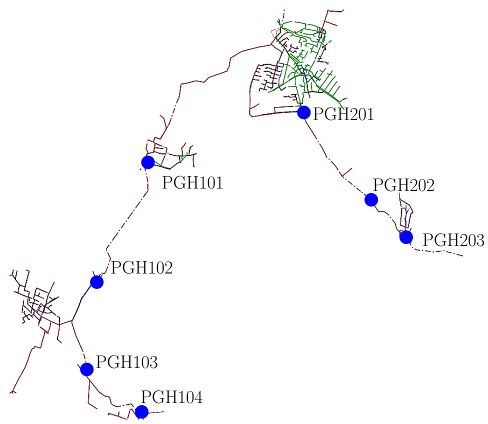

Fig. 11. A segment of a combined sewer network, where blue dots denote waste water pumping stations.

This particular segment of the network consists of seven pits with corresponding pumping stations. The estimation data has been sampled at 1 Hz and gathered from the gravity sewers connected by PGH103-104 and PGH202-203 pumping stations. (For detailed explanation of the flow estimation method utilized in this work, consult Kallesøe and Knudsen [2016]). In the two test scenarios, urban areas are not connected, therefore our tests have been restricted to groundwater detection. The model validations are shown for the two tests in Fig. 12.

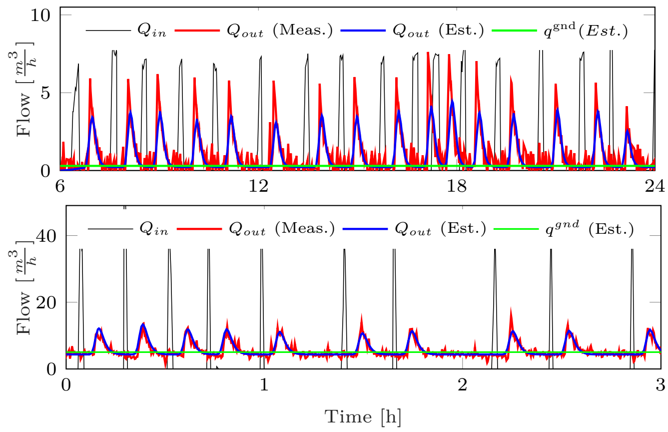

Fig. 12. Model validation on experimental data. Gravity sewer between PGH103-104 stations above, and between PGH202-203 stations below.

The estimation data covers two days in both cases. In the graph above, groundwater infiltration is approximately zero, meaning that between each pump cycle the discharged flow becomes zero, thus the channel dries out. However in the graph below, groundwater infiltration is significant, meaning that initial water level estimation is necessary. The model error yielded sufficiently small with $N_x = 7$ channel sections in both cases. As shown, the model describes the flow-dependent delays accurately under significant ground water infiltration and under a large variety of pumped input flows.

6. CONCLUSION

The presented paper focused on detecting and decoupling periodic and constant disturbance flows from the total discharge in gravity sewers. To this end, a data-driven identification method has been proposed based on physical models. The method has been tested in simulation and on data from a real network. The implementation has shown that the identification is feasible and that the estimated models predict flow and transport delays with high accuracy. The main advantage of the data-driven aspect of the modeling is that the method becomes scalable to a variety of networks, having different structures and physical parameters. Additionally, using a physical model carries an advantage of restricting the parameter space.

In our future work, we focus on stability and identifiability. Note, that the bound of the physical parameters have been chosen in a heuristic fashion. Furthermore, the state transformation resulted in loss of insight regarding water levels. Including the above considerations, we consider utilizing the models in an MPC framework.

ACKNOWLEDGEMENTS

The authors would like to thank to Provas A/S for letting us use data from their network. This project was supported by Innovation Fund Denmark (Ref. *9065-00018A*).

REFERENCES

Bjorklund, S. and Ljung, L. (2003). A review of time-delay estimation techniques. In *Proc. of 42nd CDC*, 2502–2507. Maui, USA. doi:10.1109/cdc.2003.1272997.

Butler, D. and Davies, J.W. (2006). *Urban Drainage*. Spon Press. doi:10.1016/s1462-0758(00)00017-0.

Courdent, V., Grum, M., and Mikkelsen, P.S. (2018). Distinguishing high and low flow domains in urban drainage systems 2 days ahead using numerical weather prediction ensembles. *Journal of Hydrology*, 556, 1013–1025. doi:10.1016/j.jhydrol.2016.08.015.

Kallesøe, C.S. and Knudsen, T. (2016). Self calibrating flow estimation in waste water pumping stations. In *2016 European Control Conference (ECC)*, 55–60. Aalborg. doi:10.1109/ECC.2016.7810263.

Leirens, S., Giraldo, J.M., Negenborn, R.R., and De Schutter, B. (2010). A pattern search method for improving the operation of sewer systems. *12th IFAC Symposium on Large Scale Systems: Theory and Applications*, 43(8), 591–596. doi:10.3182/20100712-3-FR-2020.00096.

Litrico, X. and Fromion, V. (2004). Simplified Modeling of Irrigation Canals for Controller Design. *Journal of Irrigation and Drainage Engineering*, 130(5), 373–383. doi:10.1061/(asce)0733-9437(2004)130:5(373).

Lund, N.S.V., Falk, A.K.V., Borup, M., Madsen, H., and Steen Mikkelsen, P. (2018). Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management. *Critical Reviews in Environmental Science and Technology*, 48(3), 279–339. doi:10.1080/10643389.2018.1455484.

Rossman, L. (2015). Storm Water Management Model User's Manual. *US Environment Protection Agency.*

Schütze, M.R., Butler, D., and Beck, M.B. (2002). *Modelling, Simulation and Control of Urban Wastewater Systems*. Springer. doi:10.1007/978-1-4471-0157-4.

Segovia, P., Rajaoarisoa, L., Nejjari, F., Duviella, E., and Puig, V. (2018). Input-Delay Model Predictive Control of Inland Waterways Considering the Backwater Effect. In *2018 IEEE Conference on Control Technology and Applications*, 589–594. doi: 10.1109/CCTA.2018.8511553.

Singh, V.P. (2001). Kinematic wave modelling in water resources: A historical perspective. *Hydrological Processes*, 15(4), 671–706. doi:10.1002/hyp.99.

Su Ki Ooi and Weyer, E. (2003). Closed loop identification of an irrigation channel. In *Proceedings of the 40th IEEE Conference on Decision and Control*, 4338–4343. Orlando, Florida. doi:10.1109/cdc.2001.980883.

Sundar, K. and Zlotnik, A. (2019). Dynamic State and Parameter Estimation for Natural Gas Networks using Real Pipeline. In *CCTA 2019 - 3rd IEEE Conference on Control Technology and Applications*. doi: 10.1109/CCTA.2019.8920430.

Troutman, S.C., Schambach, N., Love, N.G., and Kerkez, B. (2017). An automated toolchain for the data-driven and dynamical modeling of combined sewer systems. *Water Research*, 126, 88 – 100. doi: https://doi.org/10.1016/j.watres.2017.08.065.

Weyer, E. (2001). System identification of an open water channel. *Control Engineering Practice*, 33(15), 265–270. doi:10.1016/S0967-0661(01)00099-5.

Xu, M., van Overloop, P.J., and van de Giesen, N.C. (2011). On the study of control effectiveness and computational efficiency of reduced Saint-Venant model in model predictive control of open channel flow. *Advances in Water Resources*, 34(2), 282–290. doi: 10.1016/j.advwatres.2010.11.009.