Robust model for estimating pumping station characteristics and sewer flows from standard pumping station data

Fencl, Martin; Grum, Morten; Borup, Morten; Mikkelsen, Peter Steen

Published in:
Water Science and Technology

Link to article, DOI:
10.2166/wst.2019.176

Publication date:
2019

Document Version
Peer reviewed version

Citation (APA):
Fencl, M., Grum, M., Borup, M., & Mikkelsen, P. S. (2019). Robust model for estimating pumping station characteristics and sewer flows from standard pumping station data. *Water Science and Technology*, *79*(9), 1739-1745. https://doi.org/10.2166/wst.2019.176

Robust model for estimating pumping station characteristics and sewer flows from standard pumping station data

Martin FENCL[1,2*], Morten GRUM[3], Morten BORUP[1], AND Peter Steen MIKKELSEN[1]

[1]Technical University of Denmark, Department of Environmental Engineering (DTU Environment), Urban Water Systems Section, 2800 Kgs. Lyngby, Denmark
[2]Czech Technical University in Prague, Department of Hydraulics and Hydrology, 166 29 Prague 6, Czech Republic
[3]WaterZerv, Copenhagen, Denmark

**Corresponding author*
E-mail: martin.fencl@cvut.cz,

ABSTRACT

Flow data represent crucial input for reliable diagnostics of sewer functions and identification of potential problems such as unwanted inflow and infiltration. Flow estimates from pumping stations, which are an integral part of most separate sewer systems, might help in this regard. A robust model and an associated optimization procedure is proposed for estimating inflow to a pumping station using only registered water levels in the pump sump and power consumption. The model was successfully tested on one month of data from a single upstream station. The model is suitable for identification of pump capacity and volume thresholds for switching the pump on and off. These are parameters which are required for flow estimation during periods with high inflows or during periods with flow conditions triggering pump switching on and off at frequencies close to the temporal resolution of monitored data. The model is, however, sensitive within the transition states between emptying and filling to observation errors in volume and on inflow/outflow variability.

KEYWORDS

Flow estimation, pumping station, level data, power consumption, urban hydrology.

INTRODUCTION

An efficient planning and evaluation of rehabilitation/replacement actions in sewer systems requires long-term flow data (Staufer et al., 2012), however, such data is often not available as flow monitoring is challenging and often costly in the harsh conditions occurring in sewers. Level sensors are more wide-spread in the systems, since they are cheaper and require less maintenance. For this reason, flow is often computed indirectly from level observations at e.g. overflow structures where the local hydraulic properties of the system allow for it (Ahm et al., 2016; Borup et al., 2016; Isel et al., 2014). Level sensors are also usually installed in pumping stations, and using data from these to compute reliable flow estimates will potentially facilitate better analysis and control of sewer systems (Carstensen et al., 1996; Löwe et al., 2016). Pumping stations consist of a pump sump (wet well) that collects wastewater inflow and a pump. The pumping is especially by smaller systems controlled by simple control rules. The pump switches on when the pump sump is filled to some predefined level and then switches off again when the pump sump has been emptied. Pumping stations are often equipped with two (or more) pumps. This enables to increase the pump capacity during higher

inflow rates and it increases the life span of the system as pumps usually alternate after each pumping cycle during low-flow conditions. The pump redundancy also enables maintaining one pump while operating the pumping station with a second pump and increases the resilience of the whole system against complete failure.

Pump sumps are usually equipped with level gauges which measure water level in the pump sump. Furthermore, the electrical power consumption of the pump is measured. These data are primarily intended for pumping control and diagnostics and can be accessed through a SCADA system (Olsson et al., 2005). They also represent valuable sources of proxy flow data, as the average inflow into the pump sump within the time step Δt can be estimated as:

$$Q_{in} = (\Delta V(h) + Qp * Tp)/\Delta t \quad (1)$$

where ΔV is the change of water volume in the pump sump, which is a function of measured water level h given by the pump sump geometry, Qp is the pump capacity and Tp is the duration of pumping (time when the pump is on). However, pumping stations are not primarily optimized for flow monitoring and data is often collected with insufficient temporal resolution. Thus, exact times of pumping are often unknown as well as the exact pump capacity and control rules which govern the switching of the pump on and off. This causes uncertainties in inflow estimation during high flow rates when the pump is constantly running for a long period or during flow conditions triggering pump switching on and off at frequencies close to the temporal resolution of monitored data. Independent flow observations would obviously be beneficial, but such data redundancy is very rare in the water sector, and making use of whatever non-ideal data from pumping stations that is available can be a first step before further investments in data acquisition for flow estimation are made.

In this article, we suggest a pumping station inflow model with pump capacity and control rules as model parameters (we further refer to these parameters as pumping station characteristics). Furthermore, we propose a procedure to fit the model using pumping station data only. This enables us to estimate the pump capacity as well as the control rules, i.e. parameters needed for reliable pump sump inflow estimation during periods with high flows. The proposed methodology in its current form assumes constant pump capacity and is therefore only suitable for pump sumps where the pump capacity can be regarded as being independent from the water level in the pump sump.

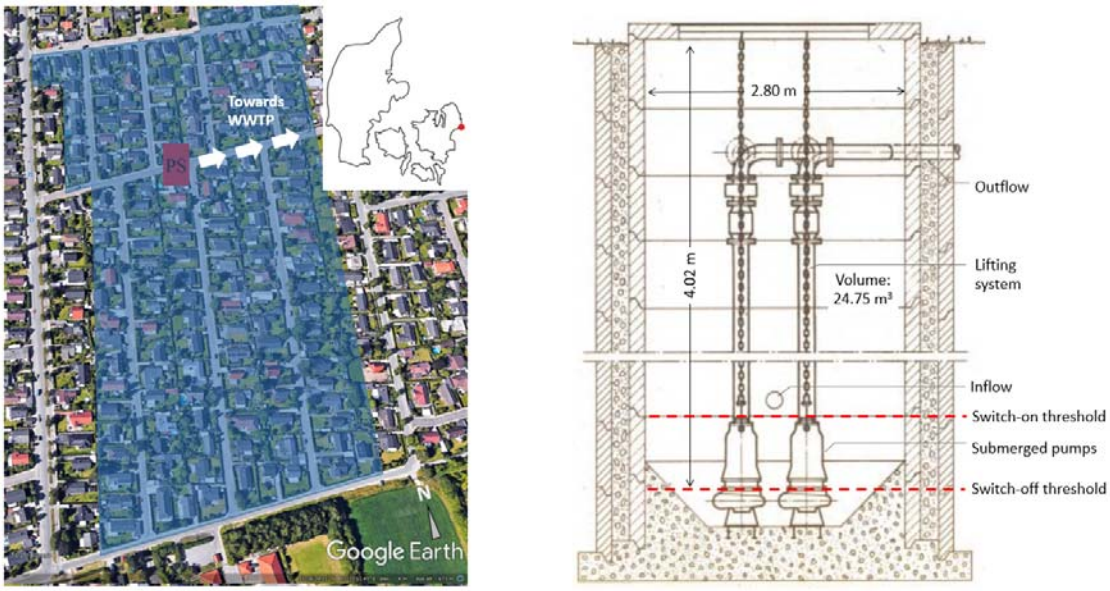

Figure 1. Subcatchment of the pumping station (left) and layout and geometry of the pumping station (right) with depicted level thresholds for switching the pump on and off.

MATERIAL AND METHODS

Taarnby catchment – data specification

The pumping station used in this study is a part of a combined sewer system in a small urban catchment located in the coastal town Taarnby, Denmark (Fig. 1, left) and characterized by flat terrain with altitudes not exceeding 10 m above sea level. The pumping station serves a subcatchment with approx. 400 inhabitants and a pipe length of about 2 km. The volume of the pump sump is 24 m^3, but it is mostly operated within the approximate range 2-5 m^3(Fig. 1, right). The pumping station has a duplex pumping system with two pumps each having a nominal pump capacity 9 l/s. The pumps alternate after each pumping cycle during dry weather flows and run simultaneously when the inflow rate into the pump sump exceeds the capacity of a single pump. In this investigation, we concentrate on dry weather flows and do not distinguish between the two pumps, implicitly assuming the pumps have the same characteristics.

The exact pumping capacity and level resp. volume thresholds for switching the pump on and off are unknown. Level data and electrical power consumption data are sampled at irregular time steps of $\Delta t \approx 5$ min. One month of observations from May to June 2017 is used for this study. This period was characterised by dry weather with only very few light rain events.

Pumping station inflow model

The pumping station inflow model should be capable of estimating flow from pump sump data during both filling (no pumping) and emptying periods (pumping). Fig. 2 (left) shows a time series of volume and electrical power consumption observations during dry weather. The pump sump is being filled for about 5 time steps (≈ 25 min) and emptied within one or two time steps (5-7 min). A positive electrical power consumption I indicates times when the pump is on. The state of the pump between these records is however not known and the exact time of pumping is thus unknown as well. For example, the first pumping on the figure could last only five minutes from 18:30 to 18:35, but also almost 15 minutes from any time after 18:25 (last record with no pumping) to any time before 18:40 (first record with no pumping).

We combine the electric power consumption data (I) and volume data (V) from consecutive time steps k and $k-1$ to estimate the exact pumping times. For this, four pump states (A-D) are defined (Fig. 2):

A: pump sump filling: $I_{k-1} = I_k = 0$
B: transition between filling and emptying period: $I_{k-1} = 0 \wedge I_k > 0$
C: pump sump emptying: $I_{k-1} > 0 \wedge I_k > 0$
D: transition between emptying and filling period: $I_{k-1} > 0 \wedge I_k = 0$

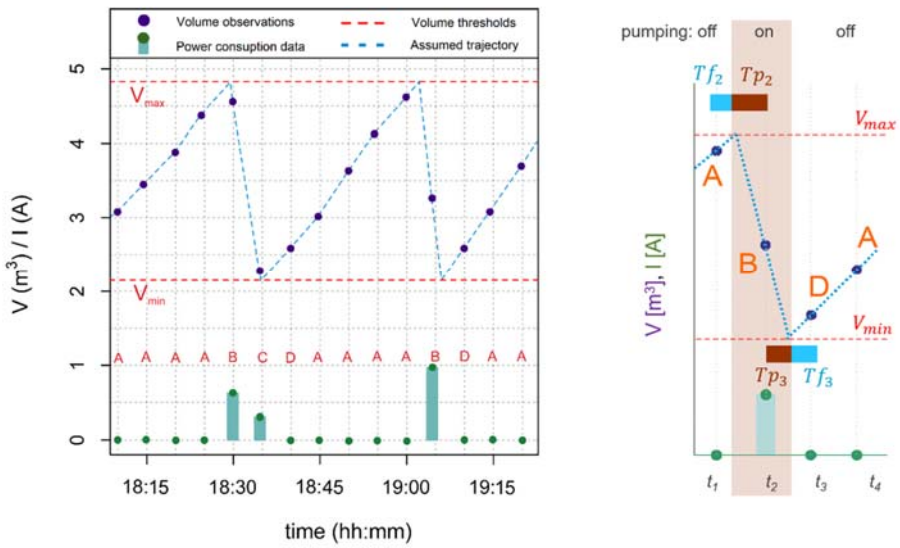

Figure 2. Left: Volume and power consumption data obtained from the most upstream pumping station in the Taarnby catchment. Red labels above power consumption records denote pump states A-D. Blue, dashed lines were added manually to the figure to indicate differences between continuous and discrete volume time series sampled at raw resolution. Right: Theoretical model for estimating exact pumping times as described by equations (2-5).

The pumping times (Tp) within state A equal zero and within state C equal the time step size Δt. The pumping times within state B are estimated from volume data assuming constant inflow Q_{in_k} during the time step Δt_k (Fig. 2, right). The period before the threshold V_{max} is reached (Tf_k) and the period the pump is on (Tp_k) can then be expressed as follows:

$$Tf_k = (V_{max} - V_{k-1})/Q_{in_k} \quad (2)$$

$$Tp_k = (V_{max} - V_k)/(Qp - Q_{in_k}) \quad (3)$$

where V_{k-1} is the last volume record when the pump was off, V_k the first record with pump being on and thus labelled as B, and Qp is the pump capacity.

Similarly, the pumping time during state D is estimated by expressing the period during pumping (Tp_k) and after it has stopped (Tf_k) using the volume threshold V_{min}:

$$Tp_k = (V_{k-1} - V_{min})/(Qp - Q_{in_k}) \quad (4)$$

$$Tf_k = (V_k - V_{min})/Q_{in_k} \quad (5)$$

Parameter estimation

The pump capacity Qp and volume thresholds V_{max} and V_{min} are regarded as model parameters and are considered constant. The model assumption of constant pump capacity is a simplification which is further discussed in the discussion section. The model is fitted to the inflow Q_{inA} estimated during state A using eq. (1) and considered constant during subsequent states B-D. The assumption of constant inflow is discussed in the Discussion section. The three parameters are optimized separately. The pump capacity Qp is optimized by minimizing the cost function L_{Qp}:

$$L_{Qp} = \sum_{k \in C} \left| Q_{inA_k} - Q_{inC_k} \right| \tag{6}$$

where Q_{inC} is the inflow estimated from eq. (1) during the state C. The V_{max} threshold is optimized for state B by minimizing the cost function L_{Vmax}:

$$L_{Vmax} = \sum_{k \in B} \left| \widehat{\Delta t_k} - \widetilde{\Delta t_k} \right| \tag{7}$$

where $\widetilde{\Delta t_k}$ is the observed duration of time step k, i.e. the time difference between two consecutive records, and $\widehat{\Delta t_k}$ is the duration of time step k estimated as:

$$\widehat{\Delta t_k} = Tp_k + Tf_k \tag{8}$$

where Tp_k and Tf_k are obtained from eqs. (2) and (3). The V_{min} threshold is optimized similarly as for V_{max} by minimizing the cost function L_{Vmin} defined for state D:

$$L_{Vmin} = \sum_{k \in D} \left| \widehat{\Delta t_k} - \widetilde{\Delta t_k} \right| \tag{9}$$

where $\widehat{\Delta t_k}$ is calculated according to eq. (8) using Tp_k and Tf_k obtained from (4) and (5). The optimization is performed by a combination of *golden section search* and *successive parabolic interpolation* (Brent, 1973) implemented in the `optimize()` function available within the statistical computing language R (R Core Team, 2017).

Performance evaluation

The exact values of pump capacity and volume thresholds are unknown and we, unfortunately, do not have independent observations of water levels and in-/out-flows to check these values. This is however a common situation for many pumping stations, and the developed estimation procedure thus has great practical value. The correctness of the suggested procedure is therefore checked indirectly. First, the proposed model is fitted to the whole dataset and the parameter space of the model is inspected to verify if it converges to well-defined minima. Second, the estimated pump capacity is compared to its nominal value and the V_{max} and V_{min} thresholds, which are completely unknown, are compared to all observed water volumes at the beginning of the transition states B and D, respectively. If the volume thresholds are identified reliably, most of the observed volumes should fall into the range bounded by V_{min} and V_{max}. Third, the total volume of inflow during the whole observation period is compared to the total outflow (pumped) volume. The inflow volume (V_{in}) is calculated as:

$$V_{in} = \sum_{k \in \{A,B,C,D\}} \left(Q_{inA_k} * \widehat{\Delta t_k} \right) \tag{10}$$

where Q_{inA} is calculated directly according to eq. (1) for time steps without pumping (A), and during pumping (B, C, D) is considered the same as the last value of Q_{inA} before pumping. The outflow (pumped) volume (Vp) is calculated from the estimated pump capacity Qp and pumping times Tp:

$$Vp = \sum_{k \in \{B,C,D\}} (Qp * Tp_k) \tag{11}$$

Finally, the pumping model is fitted separately for daily subsets of data (n = 31) to investigate the robustness and stationarity of the model parameters when optimized under different flow conditions.

RESULTS

The model was first optimized for the whole experimental period. Fig. 3 shows the parameter space of pump capacity Qp evaluated for states C and volume thresholds V_{max} and V_{min} evaluated for states B and D. All the parameters converge to well-defined minima.

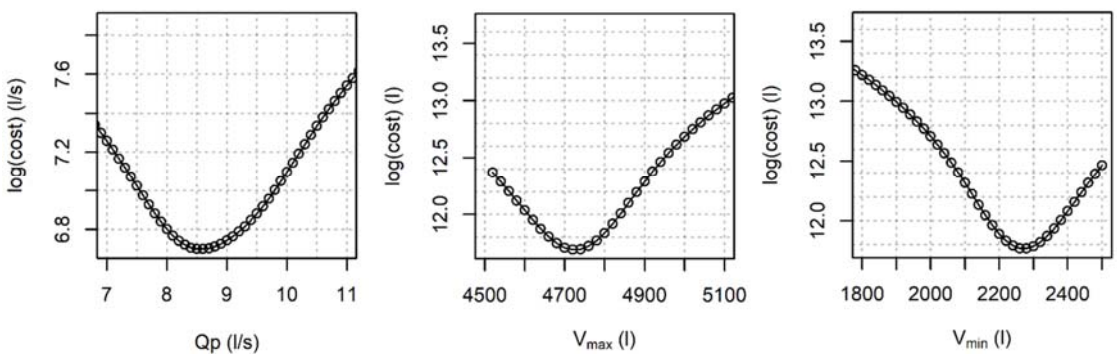

Figure 3. Cost functions of all three model parameters exhibit clear minima.

The optimal value for pump capacity was found to be 8.6 l/s, which is close to the nominal pump capacity. The maximal and minimal volume thresholds for switching the pump on and off were found to be 4.72 m^3 and 2.27 m^3. The volumes observed at the beginning of transition time steps classified as B (i.e. during which pumping started) and D (i.e. during which pumping stopped) are shown on Fig. 4. The V_{max} threshold corresponds to the 95.4% quantile of observed volumes for state B and the V_{min} threshold to the 1.7% quantile for state D, i.e. 4.6% of the state B observations are higher than V_{max} and 1.7% of the state D observations are lower than V_{min}. The records outside of the range given by thresholds are, however, close to the thresholds and may thus be explained either by inaccuracy of the level control floats switching the pump on and off, or by observation errors. The maximal volume observed (4.86 m^3) corresponds to a water level 2.2 cm above the estimated V_{max} threshold and the minimal observed volume (2.16 m^3) corresponds to a water level 1.7 cm below the estimated V_{min} threshold.

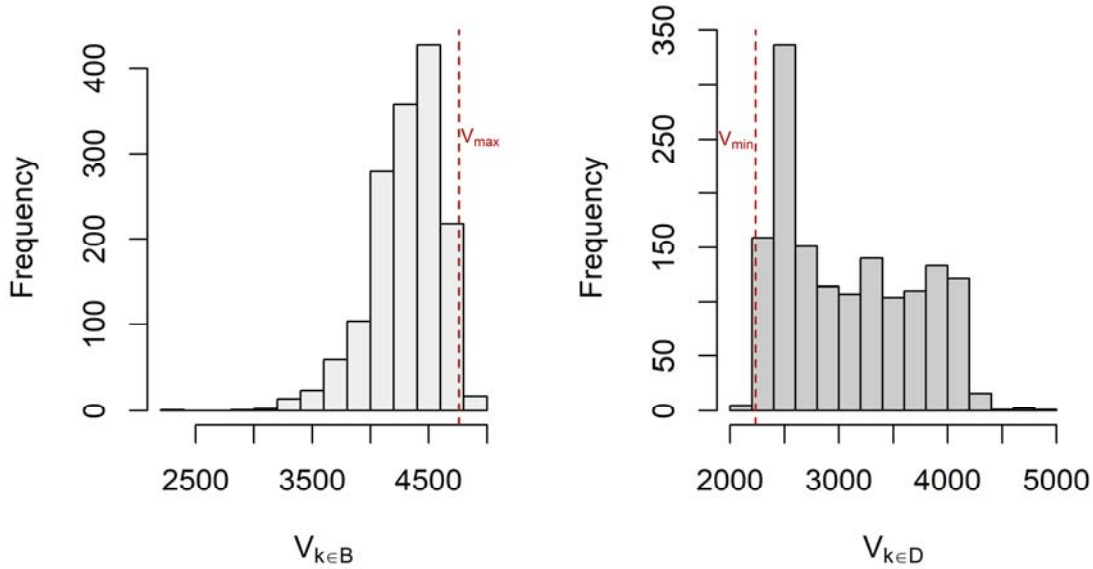

Figure 4. Observed volumes at the beginning of the transition time step during which the pump switches on (left, state B) and off (right, state D).

The total inflow volume V_{in} is 4867 m^3, and the pumped volume V_p is 4957 m^3, i.e. 1.8% higher. The pumped volume estimated without the pumping model (eqs. 2-5) is 5731 m^3, i.e. 17.7% larger than V_{in}. The close match of V_{in} and V_p indicates that the suggested model and optimization procedure is conceptually correct and improve the description of the pumping process when pumping data have low temporal resolution.

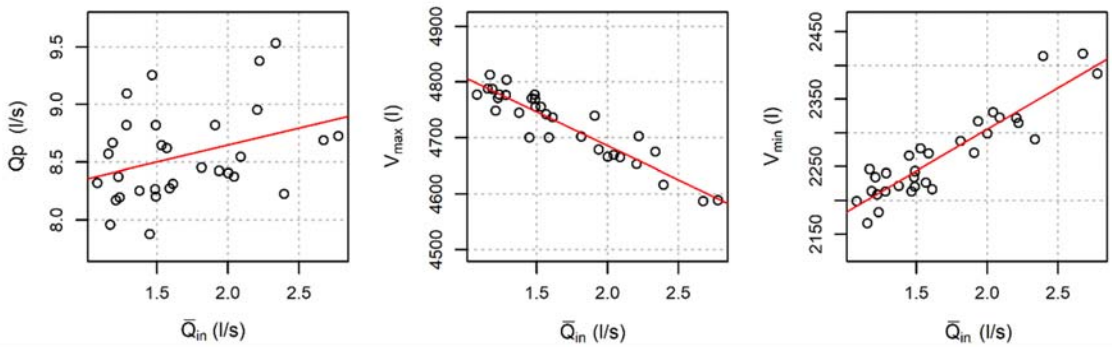

Figure 5. Pump capacity (left) and volume thresholds (centre, right) found for daily subsets and compared to average daily inflows. Linear trend lines show relations between average daily inflows ($\bar{Q}_{in}$) and the estimated parameter values.

The parameter obtained for daily subsets vary noticeably although their mean is almost identical with the parameter values obtained for the whole dataset (difference does not exceed 0.1 %). The standard deviation in pump capacity (Qp) is 0.39 l/s, and in V_{max} as well as in V_{min} it is 0.06 m^3. The magnitude of the estimated parameters is clearly dependent on the inflow rates estimated for daily data subsets (Fig. 5). The link between the mean daily inflow and the pump capacity is relatively mild (Pearson's r = 0.35), whereas the V_{max} and V_{min} magnitudes are strongly linked to daily inflows having Pearson's correlation coefficients 0.91 and -0.93, respectively. The dependence of the estimated parameters on the average inflows may partly reflect the behaviour of the real system, but it is more likely due to estimation procedure and the assumptions behind it, such as the constant pumping rate.

DISCUSSION

The suggested procedure is suitable for estimating pumping station characteristics, however inflows estimated during states B and D are sensitive to the observation errors in volume, the inflow variability and also the possible outflow (pumping rate) variability. The violation of the assumption of constant inflow and constant pumping rate during emptying periods does however not substantially influence the estimated pumping station characteristics when the variably in inflow/outflow is random.

The variability of inflows is random in case of upstream pumps, nevertheless our further investigations (not presented here) revealed that this does not always hold for downstream pumps. In the specific case where downstream pumping is triggered by short pulses of high flows from upstream pumps, the inflow rates during the last time step before pumping can be substantially larger than the average inflow rates during pumping. Such systematic deviation then influences the estimated pump capacity and thus makes the estimated flows more uncertain. To take such effects into account the model would have to be extended. A probabilistic formulation of inflow might improve the results as demonstrated e.g. by (Leonhardt et al., 2014) on a combined sewer overflow structure with a storage tank. Reformulating the model into a stochastic grey-box (state space) model (Juhl et al., 2016) may be another feasible approach to describe uncertainties in inflow.

The violation of the assumption of constant pump capacity may systematically influence the parameter estimates. The pump capacity in reality depends also on the total hydraulic head, which is lower for higher water levels. The difference in water levels for the V_{min} and V_{max} thresholds is in our case only approx. 40 cm, nevertheless, pressure heads for pumping stations in flat terrains are typically only a few meters and thus even relatively small change in hydraulic head may influence the pump capacity noticeably. The pump capacity is estimated for state C during which the water levels are lower than during state B and on the other hand higher than during state D. The underestimated pump capacity during state B results in an overestimated V_{max} threshold. This overestimation is more pronounced during periods of higher inflow rates where emptying of the pump sump is slower and thus water levels at the end of state B are on average higher than during periods of lower inflow rates. In contrast, overestimation of the pump capacity during state D results in underestimation of V_{min} threshold and this underestimation is more pronounced for lower inflow rates. This could explain the strong link between the estimated V_{min} and V_{max} thresholds and the average daily inflow rates. The extension of the model to consider pump pressure head, as e.g. (Carstensen and Harremoës, 1999) did in their model of a storage tunnel, may further improve the description of pumping and thus also the reliability of the estimated pumping station characteristics and the estimated inflows. Such extension will, however, require data of higher temporal resolution to be able to identify further parameters describing the relation between pump capacity and water level in the pump sump.

The absence of independent flow measurements does not enable us to draw conclusion on the exact accuracy of the proposed method. The indirect evaluation of the model including the estimated pumping station characteristics using the inflow/outflow balance is here based on estimated pumping times and pump capacity. Total pumping times are often monitored together with the total electrical consumption for the purpose of pump diagnosis and might therefore also be used for indirect evaluation of a model like the one proposed here. We did not, however, have access to such data for the current study.

Leaky or defect non-return valves is a frequently occurring phenomenon in pump sumps and this can result in a substantial amount of pumped water flowing back into the pump sump after a pumping cycle. This does not affect the proposed methods' ability to estimate pump sump characteristics, since the method does not distinguish between the sources of inflowing water. It does, however, affect the interpretation of the inflow time series obtained using the method since the estimated inflow in that case represents both return flow and water entering the right way into the pump sump.

Further work on validating the method and extending it to pumping stations in series may benefit from an improved data set including (i) total pumping time and electricity consumption for each of the two pumps in a pumping station, (ii) level and electrical power consumption data recorded at a sufficiently high resolution to identify the pump switching more exactly as well as any systematically increased flows just after pumping that may indicate defect non-return valves, and (iii) direct flow observations allowing the results to be independently confirmed.

CONCLUSIONS

A model for inflows to pumping stations was proposed and an optimization procedure was suggested to estimate the pump capacity and volume thresholds for switching the pump on and off. The model also enables to estimate timing of the pump switching, which is often unknown due to suboptimal sampling frequency of water level and electrical power consumption data. This enables us to identify more exactly duration of the pumping in each pumping cycle. Furthermore, a way of evaluating the performance of the model without independent flow observations was suggested. The model performance was successfully tested on one month of data from a pumping station operated within a small urban catchment with a separate sewer system in Taarnby, Denmark.

The model is suitable for estimating characteristics of upstream pumping stations which are not systematically influenced by inflow pulses from upstream pumping. Estimation of pumping station characteristics is valuable for pumping diagnostics. Reliable identification of pump sump capacity is also valuable for estimating inflows into pump sumps during periods with high flow rates when the pump is on for a long period. The model is, however, sensitive within the transition states between emptying and filling to observation errors in volume and on inflow/outflow variability. This sensitivity does not substantially influence the estimation of pumping station characteristics when the variably in inflow/outflow is random. The inflow variability can be considered random for upstream pumping stations, however, the model will require explicit consideration of upstream inflow pulses to be applicable also for downstream pumps. Further development will concentrate on considering a variable hydraulic head and a robust quantification of errors affecting the inflow estimation.

ACKNOWLEDGEMENTS

This work was financially supported by the Danish EPA through the INIGO project (MST-141-01424) and by the European Regional Development Fund (RFH-15-011) and the Capital Region Denmark (15006962) through the WISE project and by the Ministry of Education, Youth and Sports of the Czech Republic through institutional support for long-term conceptual development. The authors thank Jörg Rieckermann, EAWAG and Jan K. Møller, DTU Compute for rewarding discussions and Tårnby Utilities who provided the case data used.

REFERENCES

Ahm, M., Thorndahl, S., Nielsen, J.E., Rasmussen, M.R., 2016. Estimation of combined sewer overflow discharge: a software sensor approach based on local water level measurements. Water Sci. Technol. J. Int. Assoc. Water Pollut. Res. 74, 2683–2696. https://doi.org/10.2166/wst.2016.361

Borup, M., Grum, M., Linde, J.J., Mikkelsen, P.S., 2016. Dynamic gauge adjustment of high-resolution X-band radar data for convective rain storms: Model-based evaluation against measured combined sewer overflow. J. Hydrol. 539, 687–699. https://doi.org/10.1016/j.jhydrol.2016.05.002

Brent, R.P., 1973. Algorithms for minimization without derivatives. Englewood Cliffs, N.J. : Prentice-Hall.

Carstensen, J., Harremoës, P., 1999. Gray-Box Modeling Approach for Description of Storage Tunnel. J. Hydraul. Eng. 125, 17–25. https://doi.org/10.1061/(ASCE)0733-9429(1999)125:1(17)

Carstensen, J., Harremoës, P., Strube, R., 1996. Software sensors based on the grey-box modelling approach. Water Sci. Technol., Sensors in Wastewater Technology 33, 117–126. https://doi.org/10.1016/0273-1223(96)00164-3

Isel, S., Dufresne, M., Bardiaux, J.B., Fischer, M., Vazquez, J., 2014. Computational fluid dynamics based assessment of discharge-water depth relationships for combined sewer overflows. Urban Water J. 11, 631–640. https://doi.org/10.1080/1573062X.2013.806561

Juhl, R., Møller, J.K., Madsen, H., 2016. ctsmr - Continuous Time Stochastic Modeling in R. ArXiv160600242 Stat.

Leonhardt, G., D'Oria, M., Kleidorfer, M., Rauch, W., 2014. Estimating inflow to a combined sewer overflow structure with storage tank in real time: evaluation of different approaches. Water Sci. Technol. J. Int. Assoc. Water Pollut. Res. 70, 1143–1151. https://doi.org/10.2166/wst.2014.331

Löwe, R., Vezzaro, L., Mikkelsen, P.S., Grum, M., Madsen, H., 2016. Probabilistic runoff volume forecasting in risk-based optimization for RTC of urban drainage systems. Environ. Model. Softw. 80, 143–158. https://doi.org/10.1016/j.envsoft.2016.02.027

Olsson, G., Nielsen, M., Yuan, Z., Lynggaard-Jensen, A., Steyer, J.-P., 2005. Instrumentation, Control and Automation in Wastewater Systems. IWA Publishing.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Staufer, P., Scheidegger, A., Rieckermann, J., 2012. Assessing the performance of sewer rehabilitation on the reduction of infiltration and inflow. Water Res. 46, 5185–5196. https://doi.org/10.1016/j.watres.2012.07.001