DTU Library

Use of signatures for systematic diagnostic comparison of time series from urban drainage models and data

Pedersen, A. N.; Pedersen, J. W.; Borup, M.; Brink-Kjær, A.; Christiansen, L. E.; Mikkelsen, P. S.

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Use of signatures for systematic diagnostic comparison of time series from urban drainage models and data

A. N. Pedersen[1,2*], J. W. Pedersen[2], M. Borup[2], A. Brink-Kjær[1], L. E. Christiansen[3] & P. S. Mikkelsen[2]

[1]*VCS Denmark, Vandvaerksvej 7, 5000 Odense C, Denmark*
[2]*DTU Environment, Technical University of Denmark, Bygningstorvet, Bygning 115, 2800 Kgs. Lyngby, Denmark*
[3]*DTU Compute, Technical University of Denmark, Richard Petersens Plads, bygning 324, 2800 Kgs. Lyngby, Denmark*
**Corresponding author email: anp@vandcenter.dk*

Highlights

- Introduces the concept of hydrological signatures for validation of models in urban drainage.
- Based on an open-access study area with 10 years of data from multiple sensors.
- Shows that signatures are a powerful tool for diagnosing discrepancies between models and data.

Introduction

Digital Twins (DT) are multipurpose, physically-based models of the urban drainage system that can be used by utilities for planning, design and operation. In order to gain trust in these DTs, comparisons with observations from sensors are needed (Pedersen et al., in prep. a). The utility company, VCS Denmark, has for several years made automated daily runs with a DT of their urban drainage system and compared observations with simulated values. However, this has until recently been a visual comparison including a few metrics, which has been a large undertaking given the +200 sensors that are distributed throughout the system. In the coming years, a digital platform is therefore being developed that will automate the process of systematically comparing DT output with sensor data and provide diagnostic evaluations of where and how they disagree.

Model performance is often characterised by average "distance measures" between two time series such as NSE, statistical Likelihood, and R^2 (e.g. Bennett et al., 2013). These measures provide information about a model's degree of "wrongness" but say nothing about why and how the model is wrong. To improve the diagnostic power of model evaluation Gupta et al. (2008) introduced the concept of hydrological "signatures", which has since been widely used in the hydrological community. In this work, we introduce the concept of signatures to the urban drainage community and show how it can be the foundation of a systematic evaluation of DTs that can inform the user about what the DT can and cannot be used for.

Methodology

Definition of signatures

Signatures are quantitative metrics that target a specific property of a time series (peak level, slope of the rising limb, number of times a pump starts, etc.), which is related to a specific physical process (surface runoff, infiltration-inflow, basin emptying, etc.). A significant model-data discrepancy on a specific signature thus points directly to which process the model is failing to reproduce. Several rainfall-runoff signatures exist in the hydrological literature (Westerberg and McMillan, 2015), many of which can be borrowed by the urban drainage community. There is, however, still a need for experts to define additional signatures that are relevant in an urban context where systems are more non-linear than in natural systems.

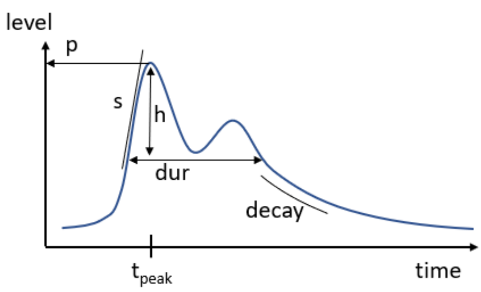

Figure 1. Identification of some signatures extracted from a time series of water levels.

Figure 1 illustrates an example of a few signatures that could be of relevance within urban drainage such as: slope of the rising limb *s*, the maximum level *h*, duration of the peak *dur*, the shape of the decay after the rain has stopped, and the time-to-peak t_{peak}. If surcharging situations is in focus, then the peak level and the duration of the peak will likely be the most relevant signatures.

Case area

The signature-based approach is evaluated on the unique open-access case area "Bellinge" (Pedersen et al, in prep. b). The area is located upstream of the city of Odense, Denmark, and has 10 years of observations from multiple sensors with no major urban developments affecting the urban drainage system in that period. The model in this study is a hydrodynamic Mike Urban model, which is also available in Pedersen et al. (in prep. b).

In order to reduce uncertainty, radar input from the utility's local X-band radar has been applied, but rain gauges could have been applied instead, however with higher uncertainty. The radar has worked from 2012 and up to 2019, however there has been miscommunication with the utility and the maintenance company of the radar, which resulted in poor quality from 2017 to 2019.

Results and discussion

For this abstract, a simple visual evaluation of two signatures (peak water level and overflow duration) at three overflow locations are shown in Figure 2 with scatterplots.

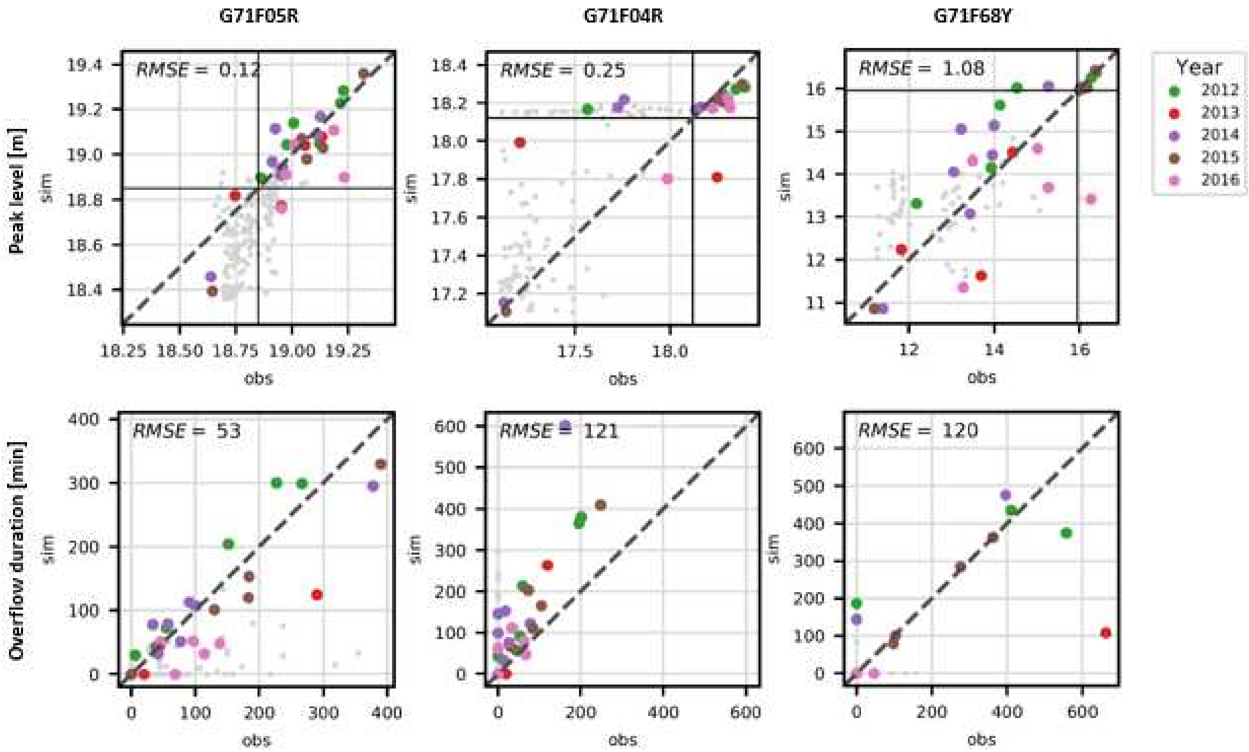

Figure 2. Results for the signatures peak level (p) and duration (dur) at three locations. G71F05R and G71F04R are upstream internal overflows that feed into G71F68Y, which is a downstream basin with overflow to a recipient. RMSE statistics in the plots are based on the coloured dots. Small grey dots refer to all events in the year 2016, while coloured dots represent 28 high-intensity rain events chosen based on local statistics for pluvial flooding based on rain gauges. The hashed line is the 1:1 diagonal. The thin black line illustrates the overflow weir level. The three locations are G71F05R Level inlet, G71F04R Level inlet 2 and G71F68Y Level PS when using the locations in Pedersen et al. (in prep. b).

At G71F05R and G71F04R, which are upstream internal overflows to G71F68Y, the peak-signature is simulated well by the DT for the high-intensity events (coloured dots). However, for all rain events in 2016 (grey dots) the model underestimates the peaks at G71F05R and significantly overestimates the peaks at G71F04R. Based on these results, the utility company assumes that there is a low point on the main pipe in G71F05R, and thus uses the signatures to learn about the performance of their real system. The overestimates at G71F04R could partly be because of a throttle pipe downstream from this location, where it is assumed that Mike Urban underestimates the flow through the throttle and thereby increasing the water level in G71F04R (Pedersen et al., in prep. b). The signatures of the downstream basin G71F68Y show a mix of the upstream effects with a larger scatter around the diagonal line than the other two locations.

The variation of signatures over time can be also be evaluated to identifying seasonality effects in certain processes and model components. Figure *3* splits the overflow duration signature for the (downstream) G71F68Y basin into dry summer months with high-intensity rain events (May-Sep, circles), and wet winter months with mostly long, low-intensity events (Oct-Apr, triangles). The model is clearly more suitable for simulating the summer period than the winter period. This is because the current model does not contain infiltration-inflow, nor does it represent the wetting of the surfaces during the wet winter period.

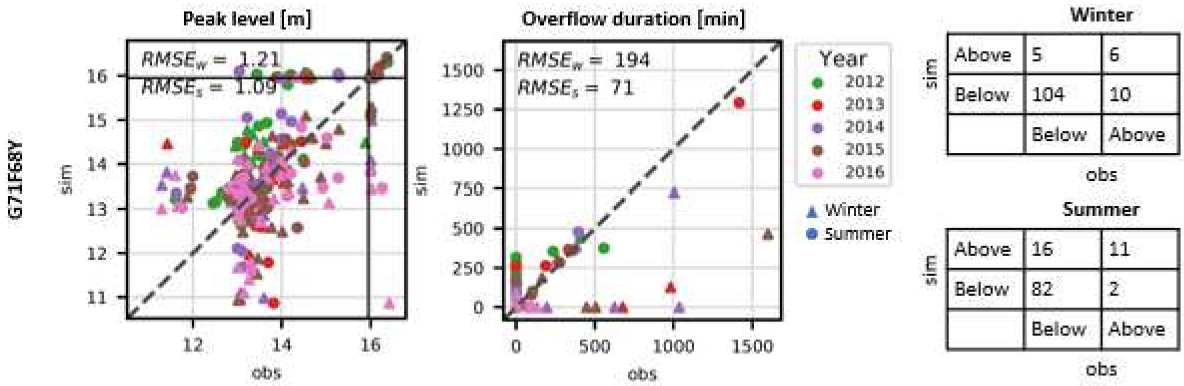

Figure 3: Signatures of peak level and duration of level above crest level. Circles are high intensity summer rain months (May-Sep) while triangles are seasons with long-term rainfall (winter, Oct-Apr). The events plotted in this graph are based on extracting data from all the periods with observed levels above 13 m in G71F68Y and running the same periods in a simulation. This gives 236 events in 5 years. The contingency tables to the right illustrate if the peak level is above crest level or below for the observation and simulation.

Conclusions and future work

Signatures are a powerful tool for diagnosing the cause of observed differences between two data series, e.g. from a model and a sensor. For the given case, two simple signatures showed that a part of the real, physical system did not function as the utility expected, that the model description near an overflow structure needs to be revisited, and that infiltration-inflow or difference in surface-characteristics needs to be explicitly included in the model to simulate rain events during winter months with long low intensity rains.

In the near future, the work will be expanded to include a systematic mapping between many different signatures and different objectives, such as surcharge evaluations, overflow estimations, infiltration-inflow, surface flooding, etc. The signature-based diagnosis of the digital twin will help the utility to better understand where to prioritize efforts in terms of optimizing the various components of the digital twin and along the way also to better estimate and accept the uncertainty of digital twins for different objectives.

References

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D. and Andreassian, V., (2013). Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Gupta, H. V., Wagener, T. and Liu, Y., (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrol. Process. 22, 3802–3813. https://doi.org/10.1002/hyp.6989

Pedersen, A.N., Borup, M., Brink-Kjær, A., Christiansen, L.E. and Mikkelsen, P.S., (in prep. a). Digital twins in urban drainage: Towards multi-purpose value-creation using models and sensors.

Pedersen, A.N., Pedersen, J.W., Vigueras-Rodriguez, A., Brink-Kjær, A., Borup, M. and Mikkelsen, P.S., (in prep. b). Bellinge: open data and models for community-wide urban drainage research.

Westerberg, I.K and, McMillan, H.K., (2015). Uncertainty in hydrological signatures. Hydrol. Earth Syst. Sci. 19, 3951–3968. https://doi.org/10.5194/hess-19-3951-2015