

Lost in calibration: why people still do not calibrate their models, and why they still should – a case study from urban drainage modelling

Franz Tscheikner-Gratl, Peter Zeisl, Carolina Kinzel, Johannes Leimgruber, Thomas Ertl, Wolfgang Rauch and Manfred Kleidorfer

ABSTRACT

From a scientific point of view, it is unquestioned that numerical models for technical systems need to be calibrated. However, in sufficiently calibrated models are still used in engineering practice. Case studies in the scientific literature that deal with urban water management are mostly large cities, while little attention is paid to the differing boundary conditions of smaller municipalities. Consequently, the aim of this paper is to discuss the calibration of a hydrodynamic model of a small municipality (15,000 inhabitants). To represent the spatial distribution of precipitation, three distributed rain gauges were used for model calibration. To show the uncertainties imminent to the calibration process, 17 scenarios, differing in assumptions for calibration, were distinguished. To compare the impact of the different calibration scenarios on actual design values, design rainfall events were applied. The comparison of the model results using the different typical design storm events from all the surrounding data points showed substantial differences for the assessment of the sewers regarding urban flooding, emphasizing the necessity of uncertainty analysis for hydrodynamic models. Furthermore, model calibration is of the utmost importance, because uncalibrated models tend to overestimate flooding volume and therefore result in larger diameters and retention volumes.

Key words | calibration, hydrodynamic models, uncertainty

Franz Tscheikner-Gratl (corresponding author)
Peter Zeisl
Carolina Kinzel
Wolfgang Rauch
Manfred Kleidorfer
Unit of Environmental Engineering,
University of Innsbruck,
Technikerstrasse 13,
Innsbruck 6020,
Austria
E-mail: franz.tscheikner-gratl@uibk.ac.at

Johannes Leimgruber
Institute of Urban Water Management and
Landscape Water Engineering,
Graz University of Technology,
Stremayrgasse 10/I,
Graz 8010,
Austria

Thomas Ertl
Institute of Sanitary Engineering and Water
Pollution Control (SIG),
University of Natural Resources and Life Sciences,
Muthgasse 18,
Vienna 1190,
Austria

INTRODUCTION

The necessity of calibration of numerical models for technical systems is unquestioned from a scientific point of view. However, in engineering practice, often uncalibrated or not sufficiently calibrated models are used. This might result in under- or overestimation of design values and, consequently, in a waste of resources. In this paper, we discuss aspects of calibration of a hydrodynamic urban drainage model from a practical point of view. We show the impact on model results when input or calibration data are not available in the required quality and we give guidance as to how small and medium sized operators with limited financial and human resources can calibrate their urban drainage models.

In the literature, model calibration is often discussed as a question of finding the perfect calibration algorithm (Rauch & Harremoës 1999; Di Pierro *et al.* 2005), the best

objective function choice (Hauduc *et al.* 2011) or in context of accuracy, the accuracy of input and output data (Bertrand-Krajewski *et al.* 2003; Korving & Clemens 2005; Kleidorfer *et al.* 2009a; Dotto *et al.* 2014; Notaro *et al.* 2015). Modern approaches of calibration try to incorporate different sources of uncertainty to identify model parameters in a way that uncertainty bands of model predictions can be evaluated (Deletic *et al.* 2009; Freni *et al.* 2009b; Lindblom *et al.* 2011; Dotto *et al.* 2012). However, in practical engineering applications, model calibration is still highly influenced by questions of data availability (Freni *et al.* 2009a; Kleidorfer *et al.* 2009b) or data quality, despite several efforts to improve this situation for input (Einfalt *et al.* 2004) or calibration data (Gamerith *et al.* 2008; Dotto *et al.* 2011). For example, the Austrian guideline (ÖWAV-Regelblatt 19 2007), which describes how to predict combined sewer

overflow (CSO) volumes and defines the standards for CSO emissions in Austria (Kleidorfer & Rauch 2011), requires the calibration of a conceptual model based on just three rainfall events. Often, these very low requirements are still seen as a barrier. Hence, for a small municipality, the guideline leaves a loophole and allows model use without calibration.

Case studies in the scientific literature that deal with urban water management are mostly large, prestigious cities which have the financial and human resources to participate in research projects. They are selected because they have data over many years and/or the required infrastructure for further data collection and management. Owing to this, there is the risk that research outcomes are biased towards the large municipalities. Such case studies are not always representative for the actual situation of the living environment in a country. The difference is not only the present or absent data but also the system layout and the catchment land use. The network length per capita in smaller municipalities is longer, as the distances between houses are higher and the population density is lower. The imperviousness of the drained area is lower, pipe diameters are smaller and the diurnal pattern of wastewater generation is more pronounced. Also, pollutant characteristics in surface runoff are expected to be different due to a different land use and reduced traffic compared to large cities (Brombach *et al.* 2005).

Consequently, the aim of this paper is to discuss calibration of a hydrodynamic model for the representative example of a small Austrian municipality (15,000 inhabitants). We investigate different calibration scenarios (number of events, data uncertainties) and the impact of the spatial resolution of catchment information on the predictive performance of the model (expressed as Nash–Sutcliffe efficiency (NSE) of measured and predicted water levels). Finally, we compare these calibration results with another important source of uncertainty, the uncertainty in the design rainfall.

METHODS

Hydrodynamic model and calibration approach

For the assessment of the scenarios, the EPA Storm Water Management Model (SWMM) was used. It is a well-known dynamic rainfall–runoff model used for single event or long-term simulation of runoff quantity and quality from primarily urban areas (Gironás *et al.* 2010). It is widely

used for analysis, design and planning related to stormwater runoff, combined and sanitary sewers, and other drainage systems (Granata *et al.* 2016). Its numerous capabilities make it suitable for application to many hydrologic urban watersheds for a single event or continuous events. SWMM consists of several different blocks (subroutines) that can be simulated separately (Wang & Altunkaynak 2012). Infiltration is modelled for this paper using Horton's equation.

One challenge in developing an SWMM model for a large watershed is defining representative precipitation data. Rainfall data drive the model and produce runoff, which means that an accurate estimation of rainfall data determine the success of the modelling effort (Barco *et al.* 2008). Some parameters that characterize the hydrological behaviour of the watershed are difficult to evaluate reliably. Therefore, a calibration stage to estimate an optimal set of these parameters is essential (Granata *et al.* 2016). This calibration was carried through using the automated PC-SWMM calibration tool (James *et al.* 2002, 2010; James 2005). It is a genetic algorithm-based software tool for calibration of the SWMM modules (James *et al.* 2002).

Case study

The analysed case study in this paper is a small municipality with 15,000 inhabitants in Tyrol, Austria. It is characterized by alpine climatic conditions, implying cold winters and heavy convective precipitation during the summer period (Kleidorfer *et al.* 2014). In Austria, 68% of the population lives in such small municipalities with less than 50,000 inhabitants (Statistik Austria 2015), and the vast majority of water utilities in the world serve populations of less than 100,000 inhabitants (Alegre & Coelho 2012).

Figure 1 shows the sewer system of the case study, which leads to a wastewater treatment plant (WWTP). Historically, it is a combined system, which was adapted over time by disconnecting several settlements (i.e. subcatchments) for alternative drainage options. The actual network consists of only 14 km of combined sewers, 67 km of wastewater sewers and 24 km of stormwater sewers. These stormwater sewers have 28 outfalls into the receiving water bodies while, in comparison, only three CSOs exist. Furthermore, 11 pumping devices are necessary for small parts of the network (mainly house connections), due to the existing terrain. In total, a catchment area of 95.31 hectare is still connected to the combined sewer system. Also, wastewater sewers of adjacent communities are connected to the network (see Figure 1).

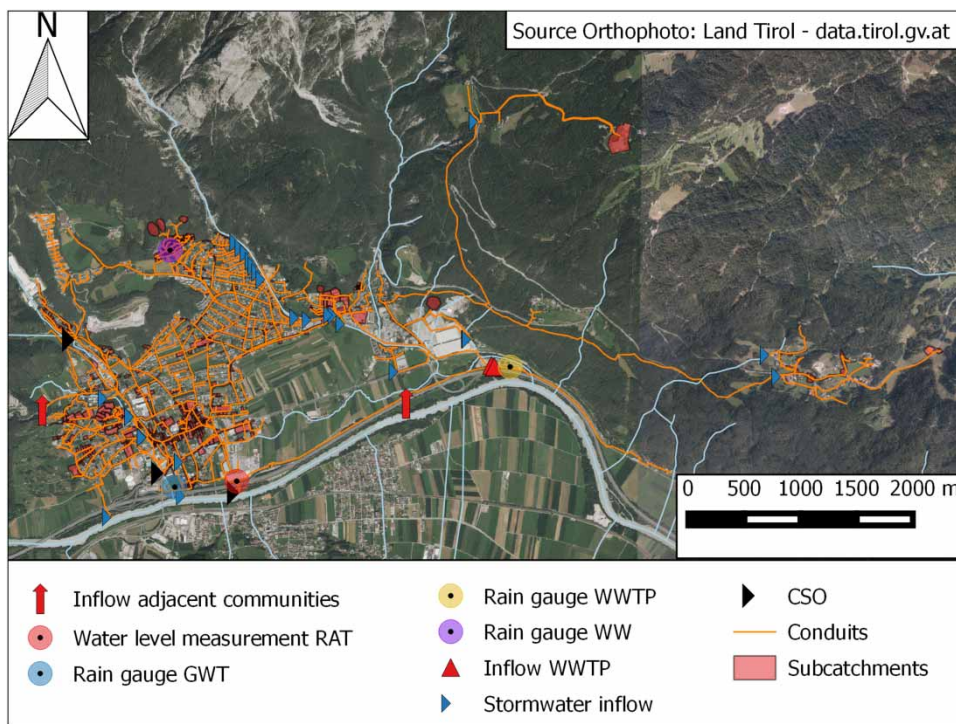


Figure 1 | Overview of the case study sewershed and the applied measurements.

The network data were available in a geographic pipe information system including diameters and slopes, which could be easily and automatically included into the hydrodynamic model. However, the detailed information about special structures had to be taken from analogue or digital blueprints and implemented manually into the model. The characteristics of the subcatchments were taken from available maps (e.g. land register) and typical imperviousness values for the uncalibrated model were chosen depending on the land use (e.g. streets with 0.95, residential areas with 0.5 (Butler & Davies 2011)). The slopes (ranging from 0 up to 30%) of the subcatchments were determined from a digital elevation map.

For model calibration and validation, precipitation was measured over the period of 1 year (2014). The abbreviations for the measurement sites (GWT and WW, as shown in Figure 1) derive from geographical names. To represent the areal distribution of precipitation, three rain gauges (one totalizer – GWT and two tipping bucket rain gauges – WWTP and WW) were used, fulfilling the specifications of Quirnbach & Schultz (2002) of one site per 16 km² and a maximal distance to the catchment of 4 km. The temporal resolution of the totalizer is 1 min, while the tipping bucket rain gauge records each tip. The water level and the flow were measured in the main collector at a position

which allowed an integrated view of the entire catchment area. The installed flow meter was a combination of two sensors. At the bottom of the sewer a flow velocity sensor (flow velocity measurement using ultrasonic cross correlation, transit time and pressure for level measurement) was installed. At the top of the sewer, some distance upstream of the flow sensor an ultrasonic sensor for water level measurement was installed. The results of the flow measurement were repeatedly checked for their plausibility, determining snapshots of the flow velocity. The measurement of the flow velocity (and in consequence the flow volume) was hampered by a clogging of the sensor mounted at the pipe floor. As a result, only the water level was used for calibration (James *et al.* 2002, 2010; James 2005).

Catchment fragmentation analysis

Model parameters such as percentage of impervious area, infiltration rate or runoff coefficient are defined as mean values for subcatchments. A more detailed fragmentation into sub-areas (e.g. roof, street, garden) allows a more accurate definition of these values (João 2002). To examine this topic, a fraction of the entire network was used. The model used for the calibration consists of subcatchments of medium level of detail without a detailed classification

of imperviousness. To show the accuracy of the chosen configuration, to analyse the influence of the spatial fragmentation on the results and to choose the necessary spatial resolution for the entire case study, a part was modelled in three different levels of detail (see Figure 2). Based on the initial situation (medium model), the level is decreased to merely one subcatchment for the entire area (rough) and increased to land parcels with a detailed compilation of surfaces (detailed).

The medium resolution model is tailored to the inlet points of the sewer network and represents the configuration used for the calibration. It separates parcels and streets. The detailed model is based on an orthophoto and discretizes green space, paved area and rooftops. The imperviousness of the land use types is adjusted to the orthophoto and harmonized with the overall imperviousness of the medium model. As a reference model, the whole area is represented as one subcatchment and consists of the merged subcatchments of the medium model. The entire rain series, measured at the site GWT, was used for the simulations.

Regarding the catchment fragmentation, we can observe that the difference in runoff rate behaviour between the detailed and medium model is very small (see Figure 3). The detailed model generates 1.6% (average) less peak runoff and runoff volume in total. This is because the overall imperviousness of the detailed model is 1.8% less than in the medium model. Only at the beginning of each rainfall event, the detailed model runoff rate slightly prevails (compare with marking in Figure 3) caused by the punctual higher imperviousness. In the course of a rain event, hydrographs continue in parallel. An explanation for this is that the inlet points in the network are the same in both models.

As expected, the reference model with one catchment shows a much slower reaction for occurring rainfall events. Short peaks are cut off up to 45% (first peak) with decreasing effect the longer the rain event takes (11% at

the last peak in Figure 3) and drainage lasts longer. This results from the initial abstraction resulting in a lack of highly paved areas and therefore a higher retention capacity. The shape of the curve differs considerably from the other two models; however, the overall runoff volume is only 2.5% less compared to the medium and 0.9% compared to the detailed model.

Considering the results of this examination, a medium resolution model was chosen which represents a good approximation of the investigation area. The runoff calculation does not benefit from the detailed model whereas the low resolution causes large deviations in the shape of the hydrograph.

Calibration scenarios

For the calibration of the hydrodynamic model, the ten most intense rainfall events which occurred in the measurement period were chosen (see Figure 4). This was done to identify the rainfall events most probably used for calibration according to the Austrian guidelines (ÖWAV-Regelblatt II 2009). In order to find relevant and independent rain events with significant intensity, those rain events were chosen from the database of measured rain events in 2014 which surpassed an event threshold of 3 mm for all three rain gauges with an inter-event time of 24 h.

To exemplify the uncertainties imminent to the calibration process, 17 scenarios – differing in assumptions for calibration – were used. The first scenario was the base scenario without calibration (UK_00). Then, a calibration for each rainfall event took place individually (KR_01 to KR_10 – compare with Figure 4). Then the model was calibrated using the entire rain series (KG_11). This rain series consists of the 10 rainfall events and artificially made intermediate dry weather periods which should ensure that the events do not interact in terms of remaining water in the network. According to DWA-A II 8 (2006), a dry weather period

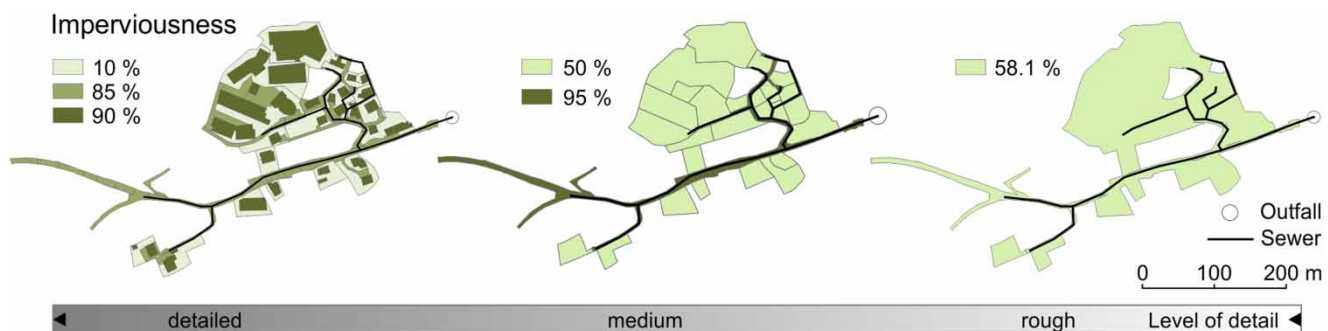


Figure 2 | Three different levels of detail of an urban drainage model (high-low).

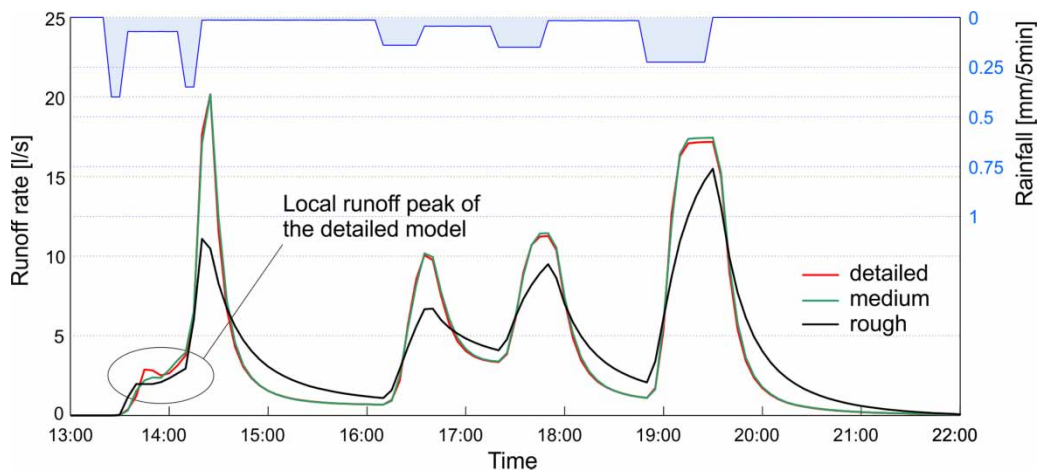


Figure 3 | Runoff at the outfall for the detailed, medium and rough model.

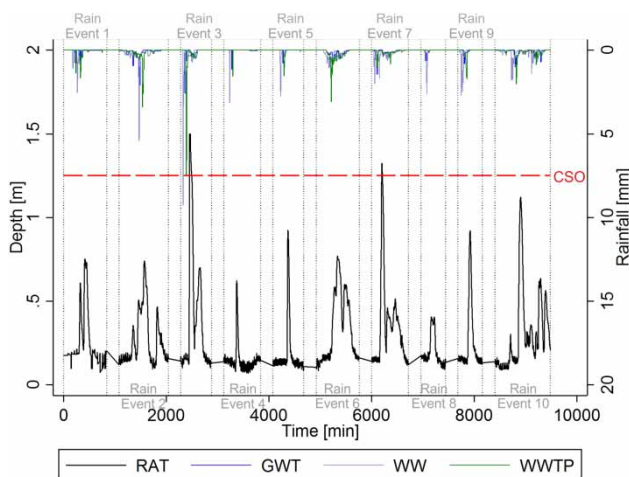


Figure 4 | Measurements used to calibrate the model. RAT: water depth in the sewer; GWT, WW, WWTP: rain gauges.

of 4 h was chosen, considering that the maximum flow time is smaller. By including these 4 h intervals, the overall length of the rain series summed to 158 h. For these calibrations, the data from all three rain measurements were used by assigning the data from the closest rain gauge to each catchment. For the next scenarios, the rain series was used for calibration with only one rain gauge – GWT (KE_12), WWTP (KE_13) and WW (KE_14) – respectively. Furthermore, the model was calibrated twice under the assumptions of a systematic 30% error of the depth monitoring data (KF_15 for +30% and KF_16 for –30%) to show the influence of measured calibration data uncertainties (Deletic *et al.* 2012).

One model parameter that was adapted during calibration was the flow time on the surface of each

subcatchment. In SWMM, this is related to the rectangular shape of the catchment and is expressed by the parameter ‘width’. This parameter influences the shape of the hydrograph on each subcatchment. Another important parameter is the fraction of imperviousness in a subcatchment, which is directly related to the amount of stormwater runoff that is generated from rainfall. The pipe roughness (the third parameter used) in the sewer pipes influences the flow velocity in the pipes and, consequently, the shape of the hydrograph and the water depth in the pipes. Other model parameters (losses, surface roughness, etc.) were not changed as a sensitivity analysis showed a negligible impact on model results. Using parameters influencing different qualities of the hydrograph puts forward the possibility that an excellent match in peak value and timing could be compromised by pure volume match. To minimize this risk, an analysis of the general hydraulic behaviour over the entire measurement campaign (2014) has been made. We compared the manholes that suffered overflows during this period as well as the inflow at the WWTP using the available measurement there (daily inflow) with our model results.

The following ranges of physically reasonable boundaries of model parameters, which have to be considered during calibration, were assumed to be:

- minimum width of 1 m for the subcatchments;
- fraction imperviousness between 0 and 100%;
- pipe roughness – Manning’s n between 0.01 and 0.02.

The minimum width is caused by local minima of subcatchments (in our case rural roads). Due to the fact that the model used for the calibration applies subcatchments without a detailed classification of imperviousness, the

parameter space is chosen to allow a compensation of inaccuracy in area data, especially for the scenarios simulating measurement errors. Manning roughness cannot be smaller than about 0.01 (which corresponds to smooth turbulent flow conditions) and should not be bigger than 0.02 for corrugated concrete sewer pipes (Butler & Davies 2011).

The NSE was chosen as the objective function for model calibration (Nash & Sutcliffe 1970). NSE is a measure to compare time series. It ranges from $-\infty$ to 1 (perfect match). The calibration caused notable deviations from the uncalibrated model. For example, for KG_11, the imperviousness decreased by 31%, the conduit roughness by 7% and the width by 67%.

Scenario comparison

To compare the impact of the different calibration scenarios on actual design values, design rainfall events of type Euler II with a duration of 120 min were used (see, as an example, Figure 5). This is a typical design storm event used in Germany and Austria (DWA-A 118 2006). It is constructed based on measured rainfall data for different required return periods. In typical engineering applications, a design rainfall event for a specific return period between 1 and 10 years (depending on the land use) is used and pipe diameters are chosen in a way that no flooding may occur for that return period. This means that differences in model results for the different design rainfall events directly influence required investments and flood risk. The rainfall intensities of these design storms depend on the data in the Austrian rainfall database (eHYD) (Weilguni 2009). In practice, the consulting engineer is advised to use rainfall intensities of one of these data points in the case study

area (see Figure 5). In order to analyse the uncertainty related to the rainfall input, design rainfall events are constructed from all of the 16 nearest data points. Additionally, the difference of the calibration scenarios was compared by using one continuous year of measured rainfall records from the nearest available measurement site of the Austrian Central Institute for Meteorology and Geodynamics (ZAMG). The performance of these scenarios was assessed using the SWMM software tool (Gironás et al. 2010). This approach using virtual cones to store flooding volume is not sufficient to provide a detailed flood risk assessment. For this task, particularly in a steep catchment, a dual drainage model (Djordjević et al. 1999) would be more appropriate. However, as it is not the aim of this study to provide detailed flood risk maps, the indicator of 'ponded volume' can be used to assess uncertainties related to input data and calibration scenarios.

RESULTS AND DISCUSSION

The differences of the quality of calibration (expressed by NSE) depending on the modeling scenario can be seen in Figure 6. For the calculation of the NSE of the rainfall series (all 10 events), only the rainfall events are compared and not the dry weather period in between to avoid any distortion of the results. As minimal acceptable quality for calibration, a $NSE > 0.5$ is assumed after consulting local authorities. It can be seen that for two rainfall events (2 and 5), a calibration is not possible, i.e. an acceptable NSE cannot be reached. This can happen, for example, if the recorded rainfall at the rain gauges is not representative. Other causes could be unrecorded cross connections between subcatchments, incorrect

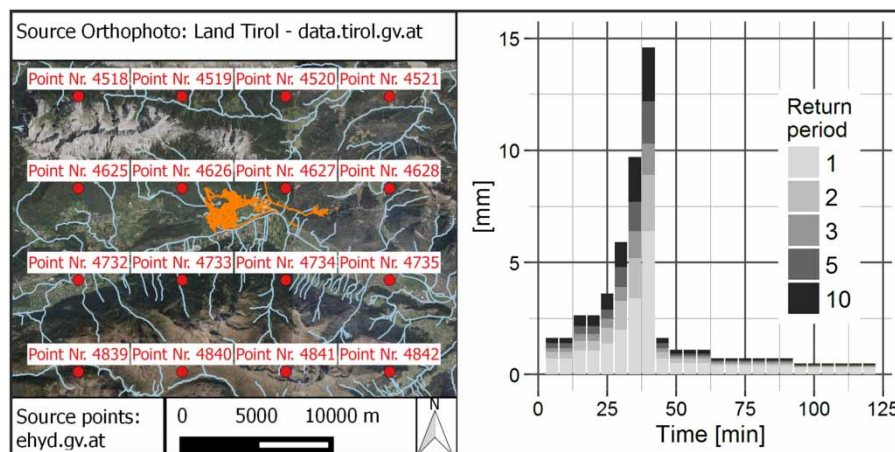


Figure 5 | Data points of Austrian rainfall database for design storm events and rainfall intensities for design rainfall event at point 4626.

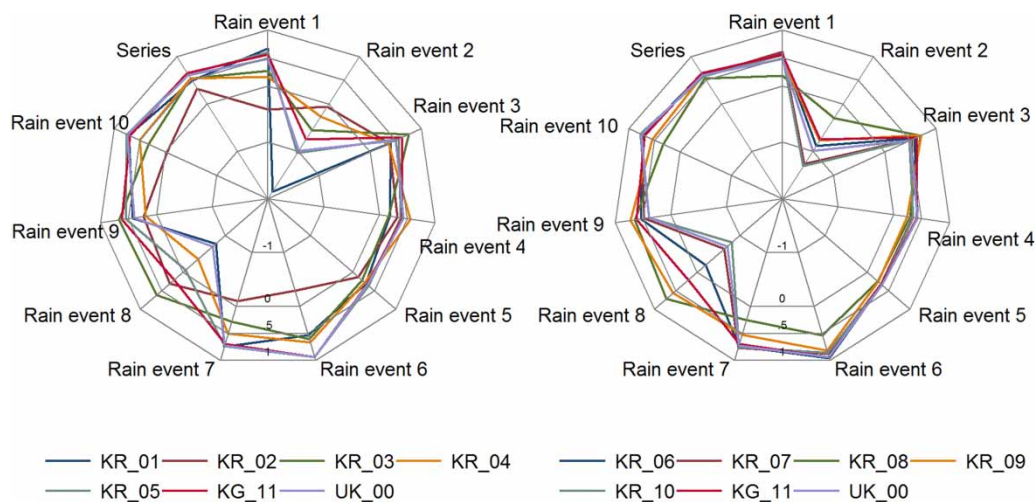


Figure 6 | NSE of the different calibration scenarios depending on the rainfall event.

measured flow depth and blockages and backwater. The uncalibrated model (UK_00) is above this limit for three out of 10 rainfall events as well as for the entire rain series. Two calibration scenarios (KR_01 and KR_02) show a worse performance than the uncalibrated model. For KR_02, this is not surprising, as this event also cannot be predicted by any of the calibration scenarios. Hidden uncertainties, e.g. an unrepresentative measurement of rainfall, prevent the model from predicting this event well. Theoretically, other sources of uncertainty such as an ill-posed model structure or measurement errors could lead to such a behaviour. However, this is not likely in this case as calibration was successful for the other events. The remaining calibration scenarios exhibit a better performance than the uncalibrated model. Four scenarios achieve a NSE which is above 0.5 for four out of 10 rainfall events (KR_03, KR_04, KR_08 and KR_10) and in one scenario (KR_07) five out of 10 events can be predicted and in four scenarios (KR_05, KR_06, KR_09 and KG_11) six events and the entire rain series can be predicted at an acceptable level.

To compare the scenarios, the absolute distance of the NSE of every model (UK_00, KG_11 and KR_01 to KR_10) for every rainfall event compared to the individually best NSE for each rainfall event (which is always the scenario calibrated on that rainfall event) was calculated and summed to get a total distance. This verifies that scenario KG_11 (calibration on all 10 events) shows the best overall performance with a distance of 1.9 in comparison to the next best scenario, KR_09 with 2.2, followed by KR_06 (2.4) and KR_05 (2.5), while the uncalibrated model UK_00 sums to a value of 3.34.

The calibration of the scenarios KE_12 to KE_14 (consideration of all rainfall events but only one rain gauge) behaved differently. The calibration for the rainfall measurement at the site GWT was successful ($NSE = 0.776$), but for the other two sites only an NSE of 0.444 (WWTP) and 0.238 (WW) could be reached using the entire rain series for calibration omitting the artificial dry weather periods. In this case, the GWT site seemed to represent the actual rainfall the best of the three sites. This could be caused either by the usage of a totalizer instead of tipping bucket rain gauges or the more central location of the site in comparison to the other two. This shows the importance of a careful selection of the position for the rainfall measurement. The spatial precipitation distribution, especially in alpine regions, is heavily dependent on the topography (Mikovits *et al.* 2015a). The simple spatial interpolation of rainfall measurements (as shown here or using Voronoi-polygons (Muthusamy *et al.* 2015)) delivers different results than, for example, the usage of radar data (Einfalt *et al.* 2004; Fencel *et al.* 2013).

The scenarios KF_15 ($NSE = 0.717$) and KF_16 ($NSE = 0.346$) were calibrated using the three parameters mentioned before, showing that a systematic error can be balanced by calibration. They could be calibrated to a higher NSE (i.e. systematic errors could be compensated by adaption of model parameters), but the parameters would have to be chosen beyond plausibility limits.

The comparison of the models using the 120 min Euler II design storm events (as used by Mikovits *et al.* (2015b)) from the rainfall intensity database eHYD (see Figure 5)

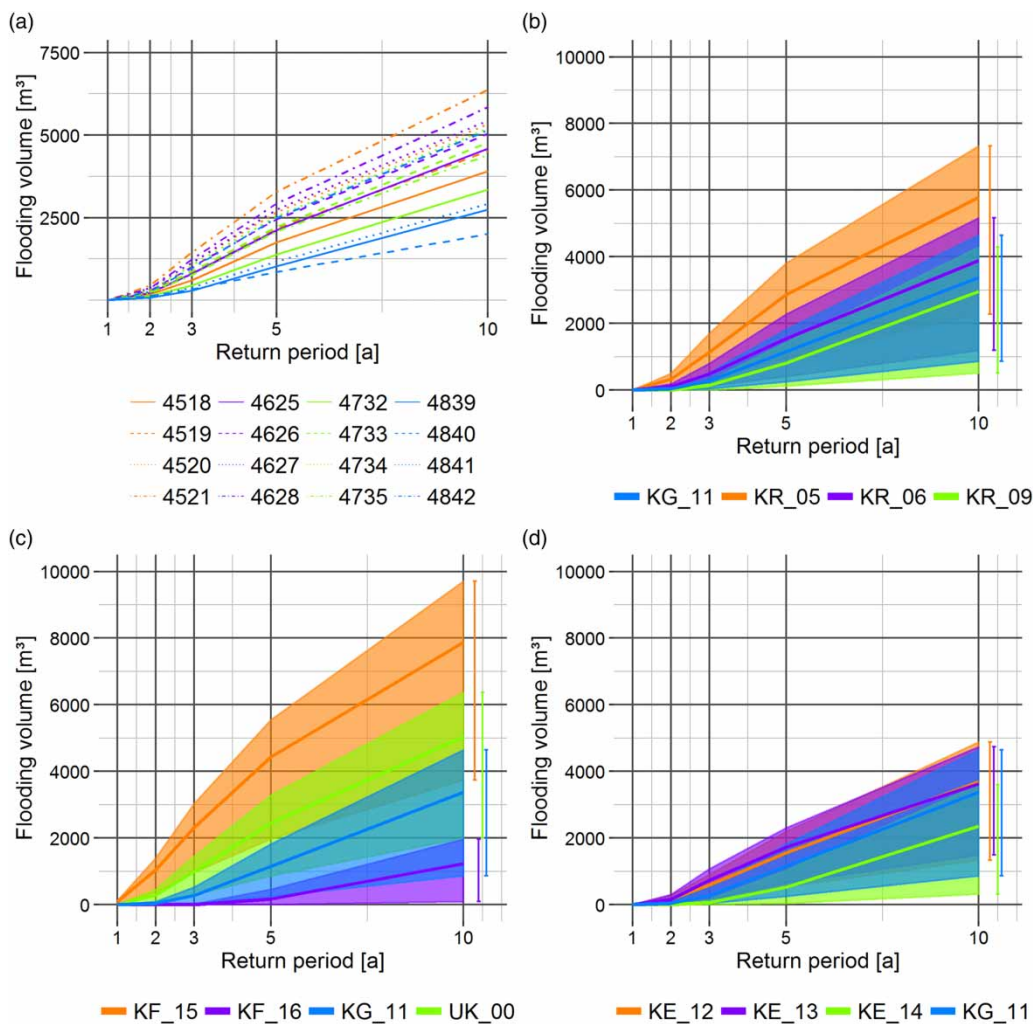


Figure 7 | Results of the hydrodynamic simulations. Flooding volumes of different design storms from data points for the uncalibrated model UK_00 (a) and the range of different calibration scenarios over these rain data highlighting the design storm event 4626 used by the operator (b)–(d).

is shown in Figure 7. It can be seen that the design storm event derived from data point 4521, with a maximum rainfall intensity of 15.7 mm/5 min for a return period of 10 years, causes the highest flooding volume while the rainfall event from data point 4840, with a maximum rainfall intensity of 14.0 mm/5 min for a return period of 10 years, elicits the lowest. Figure 7(a) shows variations in flooding volume for different return periods and data points for the uncalibrated model (UK_00). While all data points result in a negligible flooding volume for a return period of 1 year, for a return period of 2 years (which is a design value for residential areas (ÖWAV-Regelblatt II 2009)) differences of up to 80% of the maximum flooding volume (455 m³ in data point 4521) already occur between the different data points. This trend is perpetuated until the return period of 10 years where the maximum flooding

volume (data point 4521 with 6,373 m³) is more than three times higher than the minimum (data point 4840 with 2,010 m³). Similar differences can be seen between the calibration scenarios.

Figure 7(b) shows the four model scenarios which showed the best overall performance regarding the absolute distance to the optimal NSE. The range of the different rain data is shown as the background filling, while for highlighting the difference between the calibration scenarios, the design rainfall of data point 4626 is used (solid line). While data point 4627 is nearer to the WWTP the main part of the subcatchments is nearer to point 4626. It can be seen that with a return period of 2 years, flooding already occurs in the models KR_05 and KR_06, while KG_11 and KR_09 start flooding at a return period of 3 years.

Differences between the calibrated (KG_11) and the uncalibrated model (UK_00) can be in the same magnitude as differences between different return periods. This can be observed in Figure 7(c). It shows that for a return period of 2 years flooding already occurs in UK_00 whereas none of the rainfall events causes flooding in KG_11. What can also be seen in Figure 7(c) are the effects of measurement errors (scenarios KF_15 and KF_16) on the flooding values and, consequently, on required pipe diameters, retention volume and required construction costs. The influences of the different rain measurements (see Figure 7(d)) are smaller in comparison but are still recognizable.

This can be further analysed using a rainfall measurement of an adjacent rain gauge of the Austrian Central Institute for Meteorology and Geodynamics (ZAMG). Figure 8 shows the total flooding volume, total runoff and CSO volume for the different calibration scenarios aggregated over 1 year of measured rainfall. To highlight the differences between the scenarios, the deviation (in absolute percentages) of all calibration scenarios related to the best calibration KG_11 is shown in Figure 9. As expected, the scenario KF_15, which was calibrated assuming a +30% systematic error on the water level measurement, is the farthest off from KG_11. This systematic error caused a difference of 283% for the flooding volume, 152% for the CSO volume and 82% for the total runoff. Further, it can be seen that calibration using only one rainfall event (here KR_01, KR_05, KR_07 and KR_10) can be ineffective or even worse than not calibrating. Also, the usage of only one rain gauge for

calibration can lead to high deviations. This can be explained by the observations of Muthusamy *et al.* (2015), who encountered deviations of 21.6% between rain gauge measurements in catchment areas less than 10 km².

CONCLUSION

The comparison of the three different levels of detail in the representation of the subcatchments shows that they cause little effect on total runoff volume but substantial differences in the shape of the hydrograph. This is due to the similar overall imperviousness in the models. The runoff calculation does not benefit from the detailed model and therefore does not justify the extra work for the model preparation. Since the rough model, besides the runoff rate deviations, provides less scope for calibration, the medium resolution model was used for the investigation.

Furthermore, the investigations show that model calibration is of utmost importance. It was demonstrated that the uncalibrated model tended to overestimate flooding volume, thus resulting in larger diameter pipes and retention volumes. Although a conservative model could make sense in terms of safety and resilience, it should be a conscious choice of the modeller (e.g. by adding safety factors to pipe sizing and retention volume estimation) and not driven by calibration uncertainties. By calibration of the model parameters, only one type of uncertainty can be concretized while there is still substantial uncertainty in the

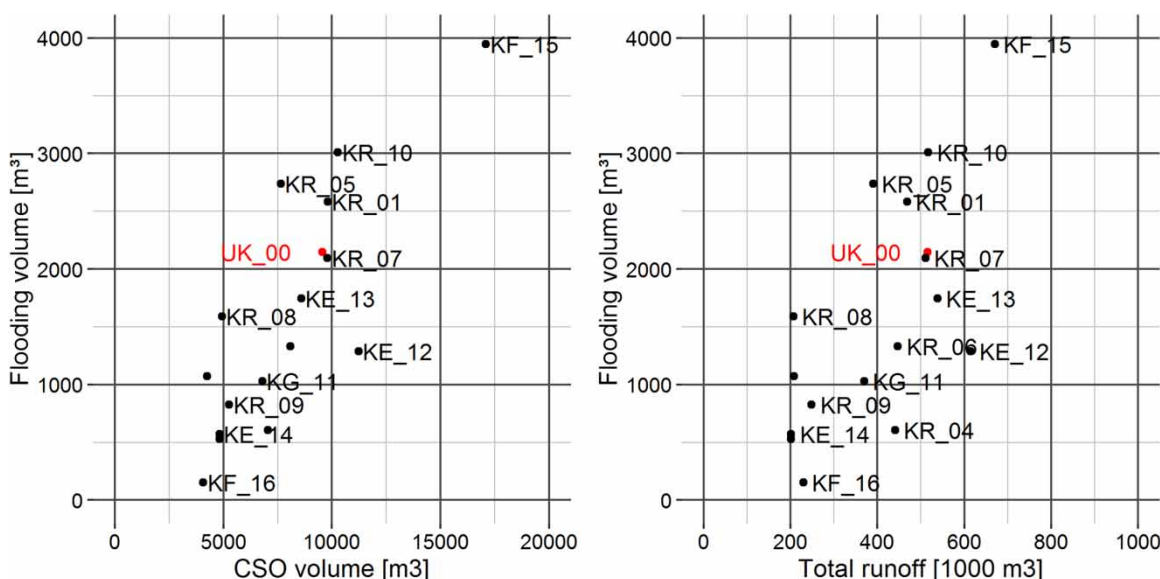


Figure 8 | Performance of the calibration scenarios using a measured rain series for the year 2014 from the nearest measurement of the Austrian Central Institute for Meteorology and Geodynamics (ZAMG).

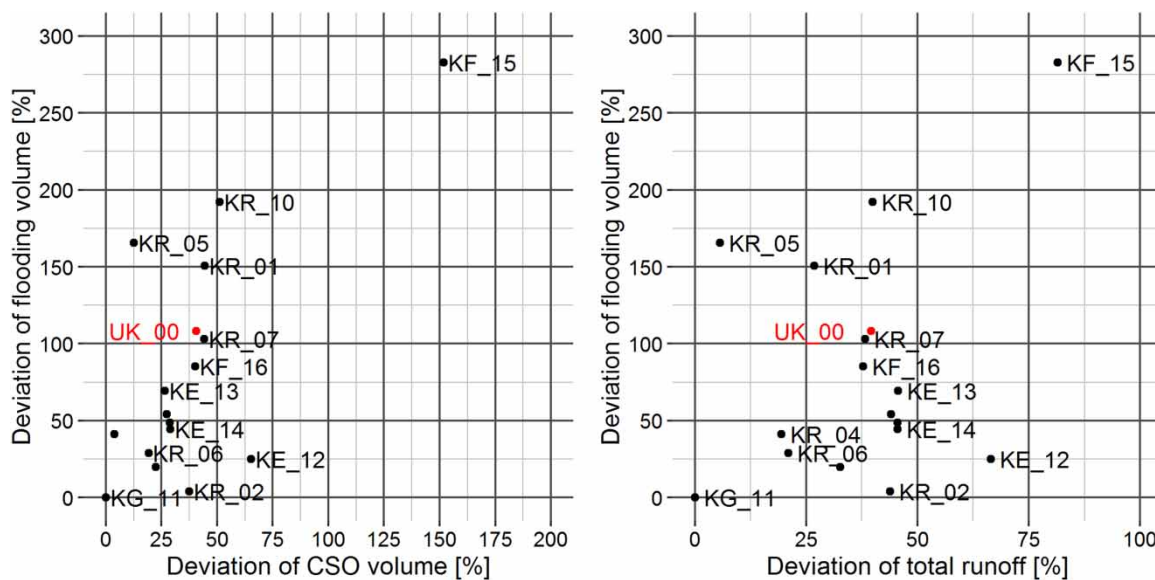


Figure 9 | Deviation of total runoff, flooding and CSO volume from the chosen scenario KG_11.

measurement data. The necessity of calibration speaks of the importance of measurement data being available in sufficient quantity and quality. It can be seen that spatial distributed rainfall measurement is advisable to minimize the uncertainties stemming from differences in rain intensities and distribution occurring in relatively small catchments, especially in mountainous regions. Additionally, the usage of a rain series is advantageous in comparison to the usage of single rainfall events, because it accurately represents antecedent moisture conditions.

The comparison of the model results using the different typical design storm events from all the surrounding data points (Austrian rainfall database for design storm events) showed substantial differences for the assessment of the sewers regarding urban flooding. The selection of the data point for the design storm event can have enormous consequences for the design of the network and, consequently, for the construction costs. This could be counteracted by either the usage of several data points in the planning process or the strict allocation of the nearest data points to the individual subcatchments. This emphasizes also the need of uncertainty analysis for hydrodynamic models in urban water management.

ACKNOWLEDGEMENTS

This work is part of the project 'DatMod – An efficient way from data to model' funded by the Austrian Federal Ministry

of Agriculture, Forestry, Environment and Water Management. We would like to acknowledge the input and good cooperation with the operating company Gemeindewerke Telfs.

REFERENCES

- Alegre, H. & Coelho, S. T. 2012 Infrastructure asset management of urban water systems. In: *Water Supply System Analysis – Selected Topics* (A. Ostfeld, ed.). InTech, Rijeka, Croatia, pp. 49–73.
- Barco, J., Wong, K. M. & Stenstrom, M. K. 2008 Automatic calibration of the U.S. EPA SWMM model for a large urban catchment. *Journal of Hydraulic Engineering* **134** (4), 466–474.
- Bertrand-Krajewski, J. L., Bardin, J.-P., Mourad, M. & Béranger, Y. 2003 Accounting for sensor calibration, concentration heterogeneity, measurement and sampling uncertainties in monitoring urban drainage systems. *Water Science & Technology* **47** (2), 95–102.
- Brombach, H., Weiss, G. & Fuchs, S. 2005 A new database on urban runoff pollution: comparison of separate and combined sewer systems. *Water Science & Technology* **51** (2), 119–128.
- Butler, D. & Davies, J. W. 2011 *Urban Drainage*, 3rd edn. Spon Press, London, UK.
- Deletic, A., Dotto, C. B. S., Fletcher, T. D., McCarthy, D. T., Bertrand-Krajewski, J. L., Rauch, W., Kleidorfer, M., Freni, G., Mannina, G. & Tait, S. 2009 Defining uncertainties in modelling of urban drainage systems. In: *8th International Conference on Urban Drainage Modelling*, Tokyo, Japan, 7–12 September, pp. 7–11.

- Deletic, A., Dotto, C. B. S., McCarthy, D. T., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T. D., Rauch, W., Bertrand-Krajewski, J. L. & Tait, S. 2012 [Assessing uncertainties in urban drainage models](#). *Physics and Chemistry of the Earth, Parts A/B/C* **42–44**, 3–10.
- Di Pierro, F., Djordjević, S., Kapelan, Z., Khu, S. T., Savic, D. & Walters, G. A. 2005 Automatic calibration of urban drainage model using a novel multi-objective genetic algorithm. *Water Science & Technology* **52** (5), 43–52.
- Djordjević, S., Prodanović, D. & Maksimović, Č. 1999 [An approach to simulation of dual drainage](#). *Water Science & Technology* **39** (9), 95–103.
- Dotto, C. B. S., Kleidorfer, M., Deletic, A., Rauch, W., McCarthy, D. T. & Fletcher, T. D. 2011 [Performance and sensitivity analysis of stormwater models using a Bayesian approach and long-term high resolution data](#). *Environmental Modelling & Software* **26** (10), 1225–1239.
- Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W. & Deletic, A. 2012 [Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling](#). *Water Research* **46** (8), 2545–2558.
- Dotto, C. B. S., Kleidorfer, M., Deletic, A., Rauch, W. & McCarthy, D. T. 2014 [Impacts of measured data uncertainty on urban stormwater models](#). *Journal of Hydrology* **508**, 28–42.
- DWA-A 118 2006 *Hydraulische Bemessung und Nachweis von Entwässerungssystemen [Hydraulic design of urban drainage systems]*. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V., Magdeburg, Germany, p. 33.
- Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirmbach, M., Vaes, G. & Vieux, B. 2004 [Towards a roadmap for use of radar rainfall data in urban drainage](#). *Journal of Hydrology* **299** (3–4), 186–202.
- Fencel, M., Rieckermann, J., Schleiss, M., Stránský, D. & Bareš, V. 2013 [Assessing the potential of using telecommunication microwave links in urban drainage modelling](#). *Water Science & Technology* **68** (8), 1810–1818.
- Freni, G., Mannina, G. & Viviani, G. 2009a [Assessment of data availability influence on integrated urban drainage modelling uncertainty](#). *Environmental Modelling & Software* **24** (10), 1171–1181.
- Freni, G., Mannina, G. & Viviani, G. 2009b [Uncertainty assessment of an integrated urban drainage model](#). *Journal of Hydrology* **373** (3–4), 392–404.
- Gamerith, V., Muschalla, D., Schröter, K., Könnemann, P. & Gruber, G. 2008 Application of multi-objective auto-calibration schemes in pollution load modelling based on high-resolution calibration data sets. In: *11th International Conference on Urban Drainage*, Edinburgh, Scotland, pp. 1–10.
- Gironás, J., Roesner, L. A., Rossman, L. A. & Davis, J. 2010 [A new applications manual for the storm water management model \(SWMM\)](#). *Environmental Modelling & Software* **25** (6), 813–814.
- Granata, F., Gargano, R. & De Marinis, G. 2016 [Support vector regression for rainfall-runoff modeling in urban drainage: a comparison with the EPA's storm water management model](#). *Water* **8** (3), 69.
- Hauduc, H., Neumann, M., Muschalla, D., Gamerith, V., Gillot, S. & Vanrolleghem, P. A. 2011 [Towards quantitative quality criteria to evaluate simulation results in wastewater treatment – A critical review](#). In: *8th Watermatex – IWA Symposium on Systems Analysis and Integrated Assessment*, San Sebastian, Spain, 20–22 June, pp. 36–46.
- James, W. 2005 *Rules for Responsible Modeling*, 4th edn. CHI, Guelph, ON, Canada.
- James, W. R. C., Wan, B. & James, W. 2002 Implementation in PCSWMM using genetic algorithms for auto calibration and design-optimization. In: *9th International Conference on Urban Drainage (9ICUD)* (E. W. Strecker & W. C. Huber, eds). Portland, OR, USA, pp. 1–10.
- James, W., Rossman, L. A. & James, W. R. C. 2010 *User's Guide to SWMM 5*, 13th edn. Water systems models. CHI, Guelph, ON, Canada.
- João, E. 2002 [How scale affects environmental impact assessment](#). *Environmental Impact Assessment Review* **22** (4), 289–310.
- Kleidorfer, M. & Rauch, W. 2011 [An application of Austrian legal requirements for CSO emissions](#). *Water Science & Technology* **64** (5), 1081–1088.
- Kleidorfer, M., Deletic, A., Fletcher, T. D. & Rauch, W. 2009a [Impact of input data uncertainties on urban stormwater model parameters](#). *Water Science & Technology* **60** (6), 1545–1554.
- Kleidorfer, M., Möderl, M., Fach, S. & Rauch, W. 2009b [Optimization of measurement campaigns for calibration of a conceptual sewer model](#). *Water Science & Technology* **59** (8), 1523–1530.
- Kleidorfer, M., Tschiesche, U., Tscheikner-Gratl, F., Sitzenfrie, R., Kretschmer, F., Muschalla, D., Ertl, T. & Rauch, W. 2014 [Von den Daten zum Modell – Anforderungen an hydraulische Entwässerungsmodelle in kleinen und mittleren Gemeinden \[Data to model – requirements for hydraulic urban drainage models for small and medium sized operators\]](#). In: *Kanalmanagement 2014. ÖWAV-Informationsveranstaltung – Wien 2014, TU Wien, 24. April 2014* (T. Ertl, ed.). Wiener Mitteilungen 231, Inst. für Wassergüte, Ressourcenmanagement und Abfallwirtschaft, TU-Wien, Wien, Austria.
- Korving, H. & Clemens, F. 2005 Impact of dimension uncertainty and model calibration on sewer system assessment. *Water Science & Technology* **52** (5), 35–42.
- Lindblom, E., Ahlman, S. & Mikkelsen, P. S. 2011 [Uncertainty-based calibration and prediction with a stormwater surface accumulation-washoff model based on coverage of sampled Zn, Cu, Pb and Cd field data](#). *Water Research* **45** (13), 3823–3835.
- Mikovits, C., Jasper-Tönnies, A., Einfalt, T., Huttenlau, M., Rauch, W. & Kleidorfer, M. 2015a [Klimawandel, Stadtentwicklung und urbane Wasserinfrastrukturplanung – Risiken und Möglichkeiten \[Climate change, urban development and the design of urban water infrastructure – risks and opportunities\]](#). *Österr Wasser- und Abfallwirtschaft* **67** (5–6), 214–221.
- Mikovits, C., Rauch, W. & Kleidorfer, M. 2015b [A dynamic urban development model designed for purposes in the field of urban water management](#). *Journal of Hydroinformatics* **17** (3), 390–403.
- Muthusamy, M., Patania, A., Schellart, A. & Tait, S. 2015 Analysis of sub-kilometre variability of rainfall in the context of urban

- runoff modelling. In: *Proceedings of the 10th International Conference on Urban Drainage Modelling* (T. Maere, S. Tik, S. Duchesne & P. A. Vanrolleghem, eds). pp. 201–205.
- Nash, J. E. & Sutcliffe, J. V. 1970 *River flow forecasting through conceptual models part I – a discussion of principles*. *Journal of Hydrology* **10** (3), 282–290.
- Notaro, V., Liuzzo, L., Freni, G. & La Loggia, G. 2015 *Uncertainty analysis in the evaluation of extreme rainfall trends and its implications on urban drainage system design*. *Water* **7** (12), 6931–6945.
- ÖWAV-Regelblatt 19 2007 Richtlinien für die Bemessung von Mischwasserentlastungen [Guidelines for the design of combined sewer overflows]. Österreichischer Wasser- und Abfallwirtschaftsverband, p. 21.
- ÖWAV-Regelblatt 11 2009 Richtlinien für die abwassertechnische Berechnung und Dimensionierung von Abwasserkanälen [Guidelines for the calculation, dimensioning and design of sewers]. Österreichischer Wasser- und Abfallwirtschaftsverband, p. 96.
- Quirnbach, M. & Schultz, G. A. 2002 Comparison of rain gauge and radar data as input to an urban rainfall-runoff model. *Water Science & Technology* **45** (2), 27–33.
- Rauch, W. & Harremoës, P. 1999 *On the potential of genetic algorithms in urban drainage modeling*. *Urban Water* **1** (1), 79–89.
- Statistik Austria. 2015 Gemeindegroßenklassen mit Einwohnerzahl. [Municipalities classified by inhabitants]. Available at: http://www.statistik.at/web_de/klassifikationen/regionale_gliederungen/gemeinden/index.html.
- Wang, K.-H. & Altunkaynak, A. 2012 *Comparative case study of rainfall-runoff modeling between SWMM and fuzzy logic approach*. *Journal of Hydrological Engineering* **17** (2), 283–291.
- Weilguni, V. 2009 Bemessungsniederschläge in Österreich [Design rainfall events in Austria]. In: *Hochwässer, Bemessung, Risikoanalyse und Vorhersage; ÖWAV-Seminar, Bundesamtsgebäude Wien, 26 May 2009* (G. Blöschl, ed.). Wiener Mitteilungen 216, Inst. für Wasserbau und Ingenieurhydrologie, Techn. Univ. Wien, Wien, Austria, pp. 71–84.

First received 3 March 2016; accepted in revised form 9 August 2016. Available online 3 September 2016