https://doi.org/10.71573/vjpthb97

All models are wrong - more or less! Added value of model calibration for dry and wet weather in Munich

Holger Hoppe [1], Daniela Böckmann[1], Ehsan Rabiei[1], Dieter Sitzmann [2], Stefan Braunschmidt[3] & Tobias Knödlseder[4]*

[1]Dr. Pecher GmbH, Gelsenkirchen, Germany
[2] BWU, Scientific Consultant on Water and Environment, Lautertal/Coburg, Germany
[3] Dr.-Ing. Pecher und Partner Ingenieurgesellschaft mbH, Munich, Germany
[4] Münchner Stadtentwässerung, Munich, Germany
**Corresponding author email: holger.hoppe@pecher.de*

Abstract
Hydraulic models form the basis for far-reaching investment and operational decisions as well as emission- or immission-based considerations. Model calibration and validation are essential to ensure adequate model quality and to assess model uncertainties, particularly in large and complex networks. This study presents the methodology for the calibration of Munich's extensive hydrodynamic sewer network model, comprising 2,400 km of sewers and serving an area of 18,000 hectares. The calibration process is based on a comprehensive 18-month measurement campaign, which included flow, water level, and rainfall data collected from more than 500 temporary and permanent monitoring stations. The calibration process identified and corrected several model inaccuracies, including errors in flow paths, connection elevations, operational settings and assumptions on infiltration water discharge. The results demonstrate the critical role of measurement data in refining hydraulic models for large, highly meshed networks. In a future workflow, quality measurements will be included in the calibration process. The project impressively shows that (all) models are more or less wrong if not calibrated to measurement data. How wrong can only be answered within the framework of a model calibration by proofed measured data and thus whether the model is suitable for the respective application.

Highlights
- Calibration of a large hydrodynamic model using flow measurement- and radar-data.
- Detection and correction of model errors such as flow paths and operational mismatches.

Introduction
Urban drainage networks are complex systems, and their efficient design and operation are essential to prevent flooding and ensure reliable wastewater disposal. Besides the structure, assumptions, conceptual framework, etc. of hydraulic models their accurate calibration is inevitable for simulating a realistic hydraulic behaviour and for evaluating the performance of urban drainage systems. Although the value of calibrated models is undisputed in science (e.g. Despotović et al., 2002; Harremoës, 2003; Tscheikner-Gratl, 2016; Ashely et al., 2024) in everyday practice precipitation-runoff measurement campaigns for model calibration are still often neglected. This has not changed significantly in the last two decades (Schmitt et al., 2008). According to the updated DWA-A 118 guidelines in Germany (DWA, 2024), finally continuous rainfall-runoff-water level and flow measurements are (at least) “highly recommended” over an “extended period” for verifying input data and calibrating the hydraulic models used in overflow and flooding assessments and to assess model uncertainties. By calibrating models using measurement data, actual runoff conditions – both in dry and wet weather – can be more

accurately represented than using standard planning values and preset model parameters (Kleidorfer, 2010). Measurement data is also necessary to verify or correct assumptions about infiltration in dry weather flow, particularly in older (combined) systems where significant amounts of infiltration can lead to higher emission rates during wet weather. The Munich project impressively demonstrates how comprehensive measurement data of more than 500 flow and level monitoring sites can be utilized to improve model accuracy and optimize the performance of an exceptionally large and complex urban drainage network.

Methodology

Model Setup

The hydrodynamic model represents the extensive drainage network of the City of Munich. The network topology forms the basis for all subsequent calculations and is initially checked carefully for plausibility. During this process, the data is reviewed, supplemented and adjusted in coordination with the network operator (Münchner Stadtentwässerung, MSE) to address any missing or implausible information, to ensure a high level of accuracy. The finalized network data is then processed and transferred into the hydrodynamic sewer network model framework (++Systems®).

The primary data for the model is provided by MSE as ESRI shapefiles, containing important details such as pipe diameters, slopes, pipe materials as well as building and shaft base areas. Munich's sewer network is an exceptionally large and complex system, with a total length of 2,400 km, 41,000 manholes and 55,000 conduits, serving a connected area of about 18,000 hectares.

The wastewater from the city is treated in two wastewater treatment plants (WWTPs): WWTP 1 (inflow capacity = 10 m^3/s) and WWTP 2 (inflow capacity = 6 m^3/s). The sewer network is divided into two more or less independent sub-systems located on the western and eastern sides of the river Isar. However, within these areas, the network is highly meshed, which complicates the delineation of sub-catchments. Superior operational control of the system is performed manually by the treatment plant staff rather than through automated processes, adding another layer of complexity to the model setup. Moreover, the dry weather discharge can be switched at an important network node to balance the load on the sewage treatment plants.

Rainfall-Runoff Measurement Campaign and quality control of the measurement data

To support the calibration of the sewer network model, an extensive measurement campaign was conducted over a period of 18 months. The campaign was strategically divided into two six-month intervals, with one focusing on the area west of the river Isar and the other on the eastern area. In addition, a selection of monitoring sites was maintained continuously throughout the entire 18-month period. These long-term measurements played an important role in quantifying the distribution of infiltration water. The campaign involved 210 temporary flow measurement devices (Figure 1), 48 temporary water level gauges, and 10 temporary precipitation gauges. These supplemented the 274 permanent measuring points already installed across the city’s sewer network.

The measurement data were continuously checked for data quality during the measurement campaign. This included establishing water level-discharge relationships, checking for outliers and conducting comparative measurements at selected flow monitoring sites. Monitoring sites with "poor measuring conditions" - especially low water levels that do not allow reliable measurements - were identified. Only checked, verified and documented measurement data provide a reliable basis for model calibration. Depending on the measurement conditions and method, even verified measurement data contain uncertainties that must be considered during the calibration process.

All relevant measurement data and operational information were transferred to a data management system (AquaZIS/++SYSTEMS®) which has an interface to the hydrodynamic sewer network simulation model (DYNA/++SYSTEMS®).

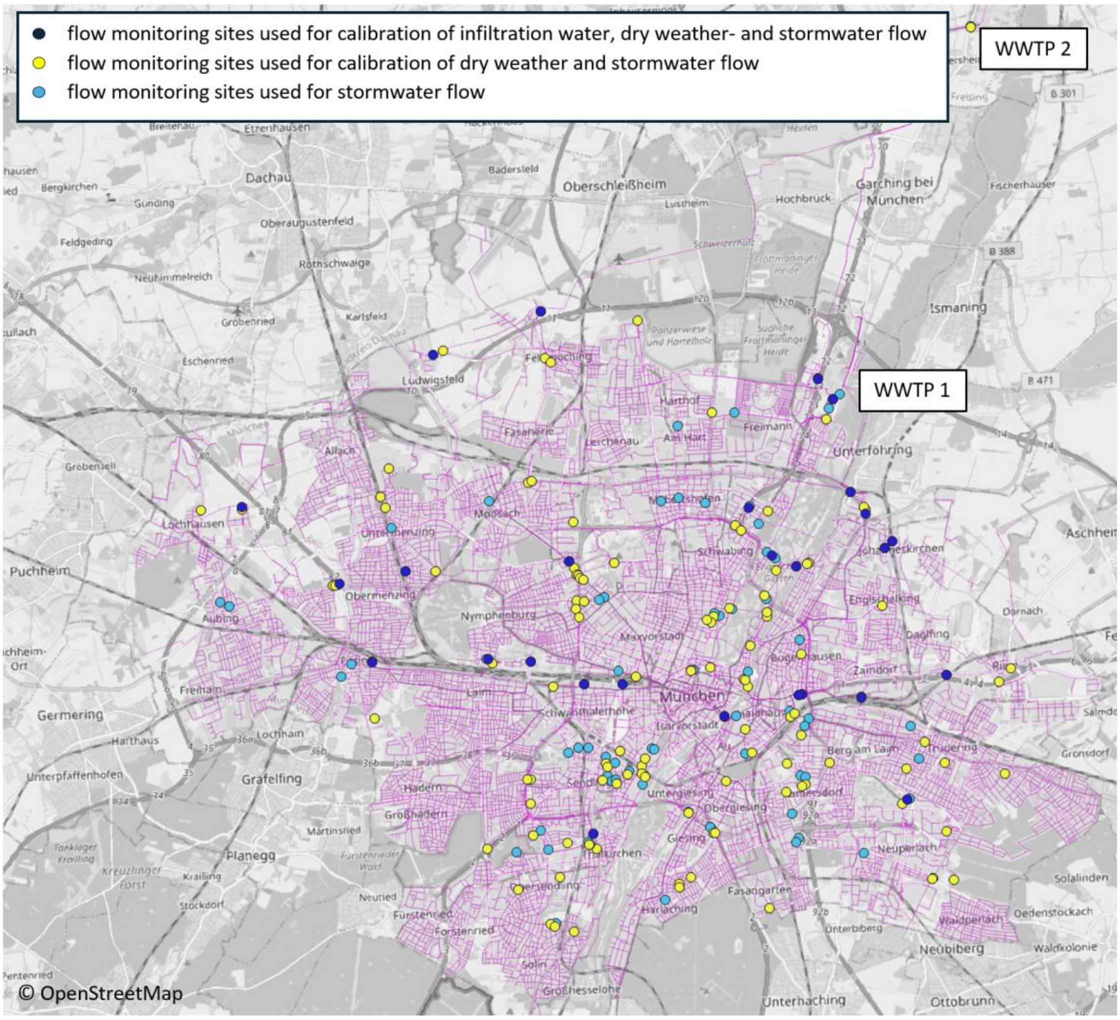

Figure 1. Sitemap of temporary flow measurement devices. (Map data from OpenStreetMap https://www.openstreetmap.org/copyright)

Model Calibration Procedure

The goal of model calibration in this project is to determine the “best” parameter set for all individual calibration events. Based on this, an "average" parameter set will be established for subsequent calculations (e.g. rehabilitation planning). The calibration of the sewer network model is divided into two phases: dry- and storm-weather calibration (Figure 2).

The first step of the dry-weather calibration involves distributing infiltration water across the city using data from the long-term flow measurements. This step ensures that infiltration is accurately represented in all sub-catchments of the model. Subsequently, all measurement data and corresponding model results at locations with dry-weather flow are compared. This iterative process of adjustment and validation refines the model to accurately replicate dry-weather conditions. The dry weather calculation was performed for the entire network using the infiltration flow for the selected reference date. To demonstrate the “correct modelling” of dry weather discharge (infiltration and wastewater) the simulated and measured dry weather daily hydrographs were compared in a second calibration step (total volume, nighttime minimum flow and pattern of daily hydrograph). These quality criteria (deviation criteria) are:

- Deviation dV_{dw} of the total dry weather (d_w) discharge volume using daily simulated and measured dry weather flow volume ($V_{dw,sim}$; $V_{dw,meas}$):

$$dV_{dw} = \frac{V_{dw,sim} - V_{dw,meas}}{V_{dw,meas}}$$

- Comparison of the nighttime minimum flow rate (dQ_{night}) using 2-hour simulated and measured averages of flow $Q_{min,sim}$ and $Q_{min,meas}$:

$$dQ_{night} = \frac{Q_{min,sim} - Q_{min,meas}}{Q_{min,meas}}$$

- Daily hydrograph anomalies assessment using the Nash-Sutcliffe Index (NS_{dw}).

$$NS_{dw} = 1 - \frac{\sum_{i=1}^{N}\left(Q_{i,dw,meas} - Q_{i,dw,sim}\right)^2}{\sum_{i=1}^{N}\left(Q_{i,dw,meas} - \overline{Q_{dw,meas}}\right)^2}$$

The threshold values for the deviation dV_{dw} (dry weather discharge volume) in the categories "very good" to "sufficient" according to Table 1 should be observed. If this was not the case, causal investigations were conducted (re-examination of drainage conditions and quality of the flow measurements, connections and branches in the sewer network model). The quality of model fitting is evaluated by an expert assessment in relation to the significance of each monitoring site and its impact on the model results.

Table 1. Quality criteria for the assessment of calibration parameters

quality criteria	very good	good	sufficient	further analysis recommended
dV	< 0.03	0.03 – 0.10	0.10 – 0.15	> 0.15
dQ	< 0.10	0.10 – 0.15	0.15 – 0.20	> 0.20
NS	> 0.75	0.75 – 0.50	0.50 – 0.25	< 0.25

For storm-weather calibration, six representative rainfall events were selected for both the western and eastern parts of the network. To ensure robust calibration, events with different rainfall intensities were included. Since such events often exhibit pronounced spatial variability, radar data was procured, processed and adjusted for the selected time periods, using ground station measurements. The calibrated radar data is subsequently used in the hydrodynamic model to simulate and validate rainfall-runoff behaviour under varying conditions. A particular challenge of the model calibration in Munich is that the operating state of the sewer systems varies from calibration event to calibration event and even changes during an event due to manmade external control interventions. The documented operating states were implemented in the model to represent this event-dependent control. At the beginning of the calibration process, the connected and runoff-effective area $A_{u,prelim}$ is "determined" as a preliminary value from the model input data . The following uncertainties remain, among others:

- Actual degree of connection of the drainage areas to the sewer network
- Runoff coefficients of the sub-areas

Model calibration with measured rainfall-runoff events requires a correction of the determined runoff effective area $A_{eff,prelim}$ if measured ($V_{r,meas}$) and simulated rainwater runoff volume ($V_{r,sim}$) differ significantly. The correction factor x_k is the ratio:

$$x_k = \frac{V_{r,meas}}{V_{r,sim}}$$

The final "runoff effective area" $A_{eff,model}$ is then:

$$A_{eff,model} = x_k \times A_{eff,prelim}$$

The correction factor x_k is not global for the entire catchment but is individually evaluated for each sub-catchment. It should preferably be applied to the paved runoff effective areas. The precipitation events used for calibration should have a precipitation level between 10 and 20 mm. At > 10 mm, the influence of initial hydrological losses is too great, and at > 20 mm, unpaved areas can contribute significantly to runoff. For x_k values << 0.8, the original values from the area-related baseline data should be checked. For example, incorrect area allocations may be present. Values x_k >> 1.2 indicate significant unpaved areas that affect runoff or incorrectly recorded paved areas. In addition to adjusting the connected and runoff-effective areas, the course of the runoff and water level hydrographs can be influenced by additional model parameters. These parameters (especially the runoff concentration) are model-dependent and may only be changed within the limits of physically reasonable quantities.
The balancing period (event period) must be defined carefully as it influences the quality criteria like the NS-coefficient and volume error (Hoppe, 2006). The beginning of the balancing period is set at the beginning of the precipitation event (or at the time when the runoff hydrograph exceeds the "normal" dry weather water level). The end of the balancing period occurs when the runoff has decreased almost to the level at the beginning of the precipitation event (dry weather water level). The balancing period will be defined individually for each of the 90 relevant measuring points and each of the twelve events.

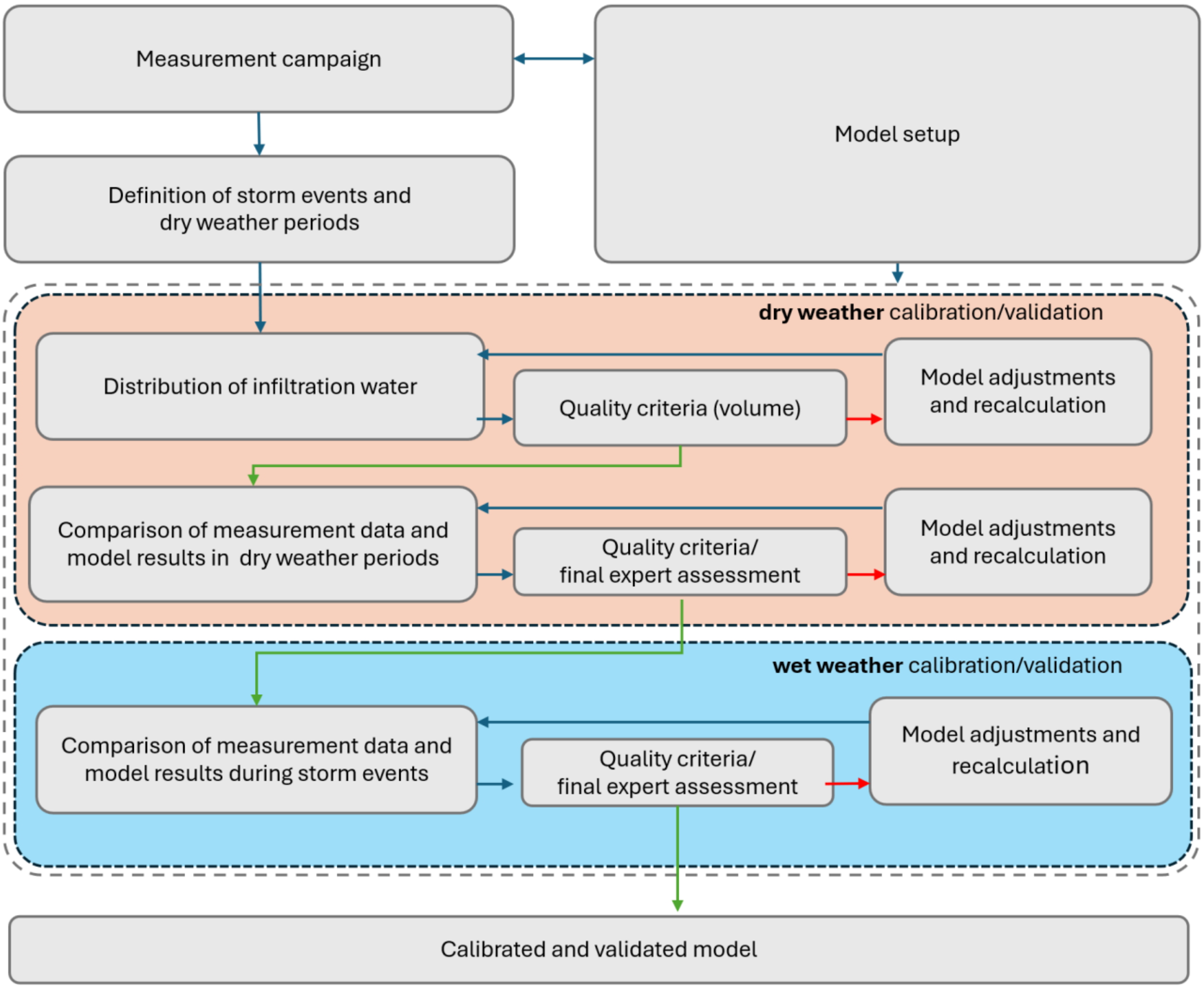

Figure 2. Flowchart for model calibration (dry weather calibration until autumn 2025, wet weather calibration until spring 2026)

When evaluating the results for individual rainfall events and deviations from the “target values” of the quality criteria (Table 1), the following points must be considered:

- Precipitation
- Runoff flow depths and flow velocities during the event (robustness of measurements)
- Ratio of stormwater to dry weather flow volume
- Duration of balance period

Using the final correction value x_k and other model parameters for the catchment areas, a final calculation run is performed for each event, the results of which form the basis for further assessment of the uncertainty of the model. The “quality criteria” for the assessment are like the dry weather simulation and calculated for each measuring point and each event:

- Volume errors of the total discharge hydrographs
- Comparison of extreme hydrograph values (flow or water level)
- Nash-Sutcliffe index of the hydrograph values (flow or water level)

Case study - peculiarities of the sewer network

The highly meshed structure of the sewer network presents significant challenges in defining accurate flow paths. Selecting suitable measurement locations is crucial to ensure that the collected data can be effectively utilized during the calibration process. In a complex and interconnected system like in Munich, identifying optimal monitoring points is particularly challenging but essential for capturing representative flow patterns and hydraulic behaviour. During the analysis of infiltration water, the measurement data proved invaluable for uncovering model inaccuracies. These included errors in assumed flow paths which were corrected using the data.Despite a carefully planned selection of measurement locations, the direct catchments of each measurement point in such a highly meshed sewer network can only be determined by a combined evaluation of dry-weather model runs and measurement data in a time-consuming but worthwhile iterative process. The prior definition of direct catchment areas is necessary before individual infiltration rates can be allocated to these areas. Anyway, the catchment area boundaries can shift slightly again and again because the water distribution at the meshes of the sewer network is not fixed but changes as well. This iterative approach allows for a more accurate identification of flow paths and the quantification of infiltration in different parts of the network, ensuring a more precise calibration of the model.

First results and discussion

The results of the model calibration process, based on the comprehensive measurement campaign in Munich’s sewer network, highlight the effectiveness of using measured data to refine hydraulic models for large, complex urban sewer systems.

Dry-weather- calibration I: Infiltration water

Already the dry-weather calibration which first focused on the distribution of infiltration water contributions across the city (Figure 3), revealed several key insights. Through the iterative process of comparing the results of model runs with measurement data, direct catchments for measurement points were identified, a task that proved challenging due to the highly meshed structure of the network. Once the direct catchments were defined, it became possible to determine the spatial distribution of infiltration water. This step was critical in refining the model and improving its accuracy, as it allowed for the identification of areas with significant infiltration. The ability to detect model inaccuracies that would otherwise go unnoticed, such as incorrect flow paths and operational assumptions (e.g. valve status), demonstrates the critical role of measurement data in improving model performance.

First results for the dry-weather calibration II – wastewater/flow and daily variation

The second step of the dry-weather calibration provided additional insights into the distribution of infiltration water due to an increased number of monitoring sites considered. However, a detailed comparison of simulated and measured flow curves offered even deeper understanding of the models' accuracy. In particular flow patterns from industrial inflows can be identified in the hydrographs which may indicate incorrect connections or inaccurate daily or weekly industrial flow patterns within the model (Figure 4a). Comparing hydrographs and quality criteria before and after each step of the calibration highlights the model improvements achieved (Figure 4b).

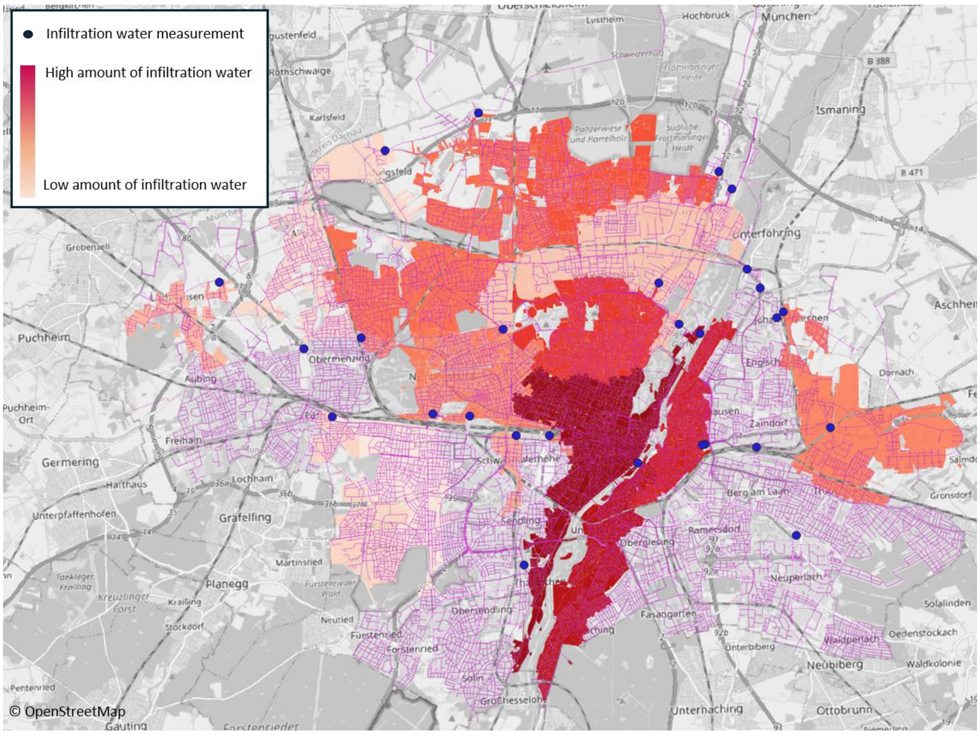

Figure 3. Data and model-based distribution of infiltration water in the urban area of Munich - one of many results of the model calibration process (Map data from OpenStreetMap https://www.openstreetmap.org/copyright)

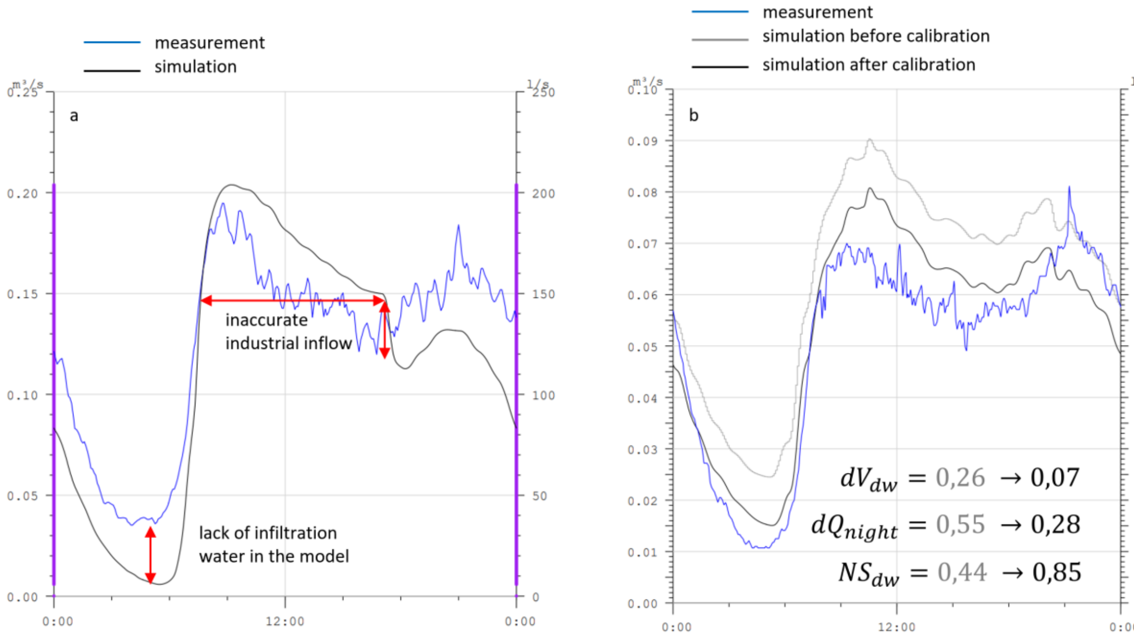

Figure 4. Identification of model inaccuracies by comparing simulated and measured dry-weather hydrographs (example a) and evaluation of calibration effects by hydrographs comparison before and after a calibration step (example b).

Conclusions and future work

The calibration of Munich's large and complex drainage network model demonstrated the importance of integrating extensive measurement data with iterative model refinement processes. By leveraging measurement data, the methodology successfully identified and corrected several model inaccuracies, such as errors in connection elevations, walled up pipes and incorrect valve states. These corrections significantly improved the accuracy of the model, enabling a more realistic representation of flow paths, sub-catchment related infiltration and hydraulic behaviour under various conditions. Despite the availability of comprehensive data and rigorous model verification processes, measurements still revealed errors in the model that were subsequently corrected. This highlights the critical role of real-world measurement data in uncovering issues that might not be evident through theoretical or simulation-based evaluations alone. Regarding infiltration water the sub-catchment specific distribution leads to a more realistic overflow behaviour of stormwater overflow tanks in contrast to a uniform distribution of infiltration water.

The calibration process also revealed challenges that warrant further investigation. The highly meshed structure of the network required an iterative process to define direct catchments of each monitoring site, which was both time-intensive and complex. Overcoming these challenges could improve the efficiency of model calibration processes.

In the next steps, the stormwater calibration will be carried out and the value of the measurement campaign and the model calibration will be evaluated by comparing the model results of the uncalibrated "initial model" with the calibrated model and considering the uncertainty. The results represent an important basis for decisions on the allocation of future investments in the areas of basic data/data quality, required monitoring programs and planning and implementation of measures within the framework of integrated planning of wastewater systems. The compilation of errors and uncertainties in the basic data shows the value of model calibration and should be considered in future planning.

The project impressively shows that (all) models yield more or less false results if not calibrated by means of measurements. How wrong can only be answered within the framework of a model calibration by proofed measured data and thus whether the model is suitable for the respective application.

References

Ashley R., Smith B., Shaffer P. and Caffoor I. (2024). ICE Handbook of Urban Drainage Practice. Emerald Publishing, ISBN 978-0-7277-4178-3, pp. 295.

Despotović J., Petrović J. and Jacimović N. (2002). Measurement, calibration of rainfall-runoff models and assessment of the return period of flooding events at urban catchment Kumodraz in Belgrade. Water Science and Technology, Vol. 45,(2): 127-133.

DWA (2024). Bewertung der hydraulischen Leistungsfähigkeit von Entwässerungssystemen (Assessment of the hydraulic performance of drainage systems) 2024; corrected version 2025. ISBN 978-3-96862-659-8.

Harremoës P. (2003). The role of uncertainty in application of integrated urban water modelling – Keynote Lecture. Proceedings of the IMUG Conference 23.- 25.05.2003, Tilburg, Netherlands.

Hoppe, H. (2006). Unsicherheiten der Grundlagendaten im Rahmen integrierter Planungen urbaner Abwasserentsorgungs-systeme (Uncertainties of the basic data within the framework of integrated planning of urban wastewater disposal systems). ISBN 978-3-9810255-1-4.

Kleidorfer M. (2010). Uncertain calibration of urban drainage models: A scientific approach to solve practical problems. Forum Umwelttechnik und Wasserbau, Vol 7. Innsbruck University Press.

Schmitt T.G., Becker M., Flores Ch., Pfeiffer E., Sitzmann D. and Uhl M. (2008). Quality Assurance in Sewer Flow Simulations through Model Calibration. Korrespondenz Abwasser, Abfall 2008, Vol. 55, No. 12. DOI: 10.3242/kae2008.12.002.

Tscheikner-Gratl F., Zeisl P., Kinzel C., Leimgruber J., Ertl T., Rauch W. and Kleidorfer M. (2016). Lost in calibration: why people still do not calibrate their models, and why they still should. Water Science and Technology, Vol. 74(10): 2337-2348. doi: 10.2166/wst.2016.395.