Reply to reviewers comments for the UWO open dataset

ESSD-2024-47 | Data description paper
Submitted on 06 Feb 2024
The UWO dataset – long-term observations from a full-scale field laboratory to better understand urban hydrology at small spatio-temporal scales
Frank Blumensaat, Simon Bloem, Christian Ebi, Andy Disch, Christian Förster, Max Maurer, Mayra Rodriguez, and Jörg Rieckermann

Jörg Rieckermann
30.6.2025

Response to Reviewer #1 (Agnethe Nedergaard Pedersen)

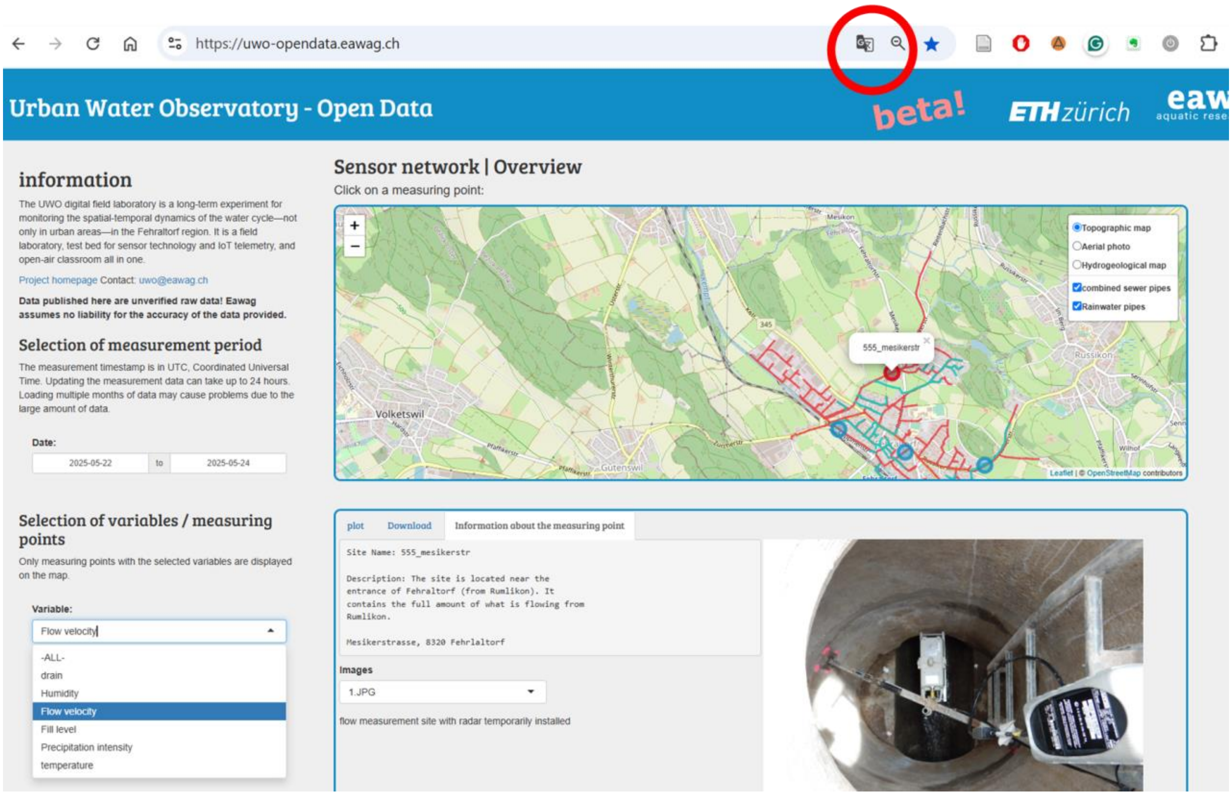
We thank the reviewer for the positive feedback on the quality and utility of the dataset and the accompanying manuscript. We appreciate the constructive suggestions and have addressed all comments (red colour) in the revised manuscript as outlined below. Final changes are formatted in blue colour.

RC1: 'Comment on essd-2024-47', Agnethe Nedergaard Pedersen, 09 Feb 2025 reply

Dear authors

Thank you for a very well written and descriptive paper of the UWO dataset. **The submission describes a three year long comprehensive measuring campaign of the urban drainage system in a Swiss minor city along with an accompanying SWMM model. The dataset is unique and well described in the manuscript with thorough supplementary material. The manuscript highlights well-thought potential application areas of the dataset.** Dataset is made accessible and with a link to a data viewer. **Translating this viewer would be of high value for the future users not able to read Swiss.** The data structure of the SQL databases are well structured and it is easy to understand data.

We appreciate the positive evaluation. Language should not be a barrier, because browsers like Google Chrome, Firefox, Edge and Safari offer translation features, either as built-in features or through extensions (red circle).



I have a few comments and corrections for the manuscript:

RC1.1: Maybe a question of taste, but **why do you use both footnotes and references**?

We agree this may be stylistically inconsistent. We have removed all footnotes in the main text and replaced them with inline citations or explanations.

RC1.2: Figure 2, top: It seems that the **top figure is missing.**
Thank you for spotting this. The missing figure has been reinstated in the revised manuscript.

RC1.3: Figure 2, buttom: It is not very clear which combination of the letters and numbers that are the naming of the manhole. **Suggestion to make bold text for names. The light colors are not clear when being printed.**

We updated the figure to use bold fonts for manhole IDs and improved the color contrast for better print readability.

RC1.4: Figure 3, left: It is not easy to read the small text. Please make it larger. The throttle says <75 L/s but in Figure 2 it says 80 L/s. Which one is correct? (Please also correct in text Line143). Are some of the text tagnames of the sensors?

Please excuse this error which carried over from a previous version of the manuscript. It is 80 L/s as documented in the Flow Data (signal: ***bf_f03_11e_russikerstr***)…

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	timestamp	average_v	battery_ch	battery_vo	device_te	distance	flow_rate	nos	pmr	rssi	surface_ve	water_level	
5117	12.12.2020 18:00	1.91454	83	13.1	6	530.1132	41.08670006	59	6.78	-85	2.109216	90.88628	
5118	12.12.2020 18:05	1.931091	83	13.1	6	526.1635	44.018738406000004	75	6.74	-85	2.11836	94.83851999999999	
5119	12.12.2020 18:10	1.909877	83	13.1	6	528.386040	42.09694761	69	6.72	-85	2.100072	92.61601999999999	
5120	12.12.2020 18:15	1.930573	83	13.1	6	526.4099	43.845165198000004	65	7.1	-85	2.11836	94.5896	
5121	12.12.2020 18:20	1.876471	83	13.1	6	531.1013	39.64985054	58	7.08	-85	2.0695919	89.89822	
5122	12.12.2020 18:25	1.860225	83	13.1	6	529.620479	40.22939528	70	6.78	-85	2.0482560	91.38158	
5123	12.12.2020 18:30	1.905396	83	13.1	6	529.1988	41.47554896	70	6.62	-85	2.097024	91.80068	
5124	12.12.2020 18:35	1.899331	83	13.1	6	529.445220	41.18595955	69	6.35	-85	2.090928	91.5543	
5125	12.12.2020 18:40	1.905701	83	13.1	6	527.718	42.43476932	74	6.6	-85	2.093976	93.28404	
5126	12.12.2020 18:45	1.972879	83	13.1	6	524.753840	45.91907861	57	6.19	-85	2.161032	96.24568000000001	
5127	12.12.2020 18:50	2.102907	83	13.1	6	518.3327	53.613951398999994	84	6.32	-85	2.2890479	102.6668	
5128	12.12.2020 18:55	2.2206203	83	13.1	7	507.9619	64.78970769	69	6.94	-85	2.395728	113.03761999999999	
5129	12.12.2020 19:00	2.122841	83	13.1	7	519.5697	53.20816914	74	6.22	-85	2.313432	101.4324	
5130	12.12.2020 19:05	2.141555	83	13.1	7	511.172459	60.01776936	87	6.42	-85	2.31648	109.8271	
5131	12.12.2020 19:10	2.221047	83	13.1	7	507.7155	64.99992357	82	6.41	-85	2.395728	113.28399999999999	
5132	12.12.2020 19:15	2.285726	83	13.1	7	503.024140	70.78199589	67	5.32	-85	2.456688	117.9779	
5133	12.12.2020 19:20	2.309652	83	13.1	7	498.825520	75.07810944	82	5.42	-85	2.474976	122.17399999999999	
5134	12.12.2020 19:25	2.327392	83	13.1	7	496.603	77.56562196	90	5.75	-85	2.490216	124.3965	
5135	12.12.2020 19:30	2.331933	83	13.1	7	493.6388	80.28398079	81	6.06	-85	2.490216	127.3607	
5136	12.12.2020 19:35	2.325472	83	13.1	7	494.134140	79.63402761	77	6.41	-85	2.48412	126.8679	
5137	12.12.2020 19:40	2.38378	83	13.1	7	493.3239	82.34979975	63	5.47	-85	2.54508	127.67563999999999	
5138	12.12.2020 19:45	2.342053	83	13.1	6	492.5822	81.55612755	80	6.17	-85	2.49936	128.4173	
5139	12.12.2020 19:50	2.366284	83	13.1	7	493.570259	81.52729542	88	6	-85	2.526792	127.42926000000001	
5140	12.12.2020 19:55	2.314621	83	13.1	7	493.748059	79.59421782	67	6.46	-83	2.4719279	127.25146000000001	
5141	12.12.2020 20:00	2.32852	83	13.1	7	493.997	79.85806020000001	69	6.25	-83	2.487168	127.0051	
5142	12.12.2020 20:05	2.314011	83	13.1	6	494.1748	79.20615123	79	6.24	-83	2.4719279	126.8273	
5143	12.12.2020 20:10	2.340803	83	13.1	7	493.4331	80.76920598	77	6.08	-83	2.49936	127.5664	
5144	12.12.2020 20:15	2.31011	83	13.1	7	492.940340	80.13628709999999	80	6.39	-83	2.465832	128.0617	
5145	12.12.2020 20:20	2.330836	83	13.1	6	492.445040	81.28465128	77	6.07	-83	2.487168	128.5545	
5146	12.12.2020 20:25	2.3698200	83	13.1	7	491.2131	83.73771666	82	5.25	-83	2.526792	129.7889	
5147	12.12.2020 20:30	2.331293	83	13.1	7	492.130079	81.57562236	74	6.26	-83	2.487168	128.8694	
5148	12.12.2020 20:35	2.343821	83	13.1	7	491.390940	82.66285233	83	5.91	-83	2.49936	129.6111	
5149	12.12.2020 20:40	2.377379	83	13.1	7	491.8837	83.4075036	85	6.13	-83	2.535936	129.11581999999999	
5150	12.12.2020 20:45	2.373142	83	13.1	6	490.8956	84.13543602	96	5.43	-83	2.52984	130.1039	
5151	12.12.2020 20:50	2.327727	83	13.1	7	492.6254	81.02093508	80	5.99	-83	2.48412	128.3767	
5152	12.12.2020 20:55	2.346686	83	13.1	7	491.390940	82.76367015	86	6.19	-83	2.502408	129.6111	
5153	12.12.2020 21:00	2.363572	83	13.1	6	491.568740	83.20138857	86	5.48	-83	2.520696	129.4308	
5154	12.12.2020 21:05	2.326782	83	13.1	7	491.322359	82.1210985	80	6.1	-83	2.481072	129.6797	
5155	12.12.2020 21:10	2.346411	83	13.1	7	491.568740	82.59774345	89	6.01	-83	2.502408	129.4308	

bf_f03_11e_russikerstr

Also documented in the SWMM file (*faf.inp*) for the CSO tank Russikon RKB_Morgental max flow =80 l/s

```
[CONDUITS]
;;Name           From Node        To Node          Length     Roughness  InOffset   OutOffset  InitFlow   MaxFlow
;;-------------- ---------------- ---------------- ---------- ---------- ---------- ---------- ---------- ----------
10-10a           10_              10a              45.054     0.011764706 0         0          0          0
10a-10b          10a              10b              45.037     0.011764706 0         0          0          0
10b-10c          10b              10c              44.970     0.011764706 0         0.18       0          0
10c-11           10c              11               23.105     0.011764706 0         0          0          0
11-11a           11               11a              58.919     0.011764706 0         0          0          0
11a-11b          11a              11b              58.950     0.011764706 0         0          0          0
11b-11c          11b              11c              46.934     0.011764706 0         0          0          0
11c-11d          11c              11d              46.949     0.011764706 0         0          0          0
2c-2d            2c               2d               46.356     0.011764706 0         0          0          80
2d-2e            2d               2e               36.293     0.011764706 0         2.31       0          0
2e-2f            2e               2f               98.954     0.011764706 0         0.02       0          0
2f-10            2f               10_              51.732     0.011764706 0         0          0          0
411-412          411              412              55.544     0.011764706 0         0          0          0
412-413          412              413              60.297     0.011764706 0         0          0          0
413-414          413              414              64.871     0.011764706 0         0          0          0
414-415          414              415              64.734     0.011764706 0         0          0          0
415-416          415              416              64.780     0.011764706 0         0          0          0
416-11b          416              11b              64.809     0.011764706 0         0          0          0
RKB_Morgenthal_throttle rus4           2c          200        0.011764706 0          0          80          0
2                rus3             rus4             30         0.011764706 0         0          0          0
```

RC1.5: Line 142. What does it mean that the nearby villages are “largely” connected to Fehreltorfs. Please specify.

We changed the text to:

ll.153ff:

> *The nearby villages of Russikon (3320 inhabitants) and Rumlikon (451 inhabitants) are also largely connected to Fehraltorf's wastewater treatment plant (WWTP) (HBT, 2016) (Figure 2). Only a very small neighbourhood of Russikon (Madetswil) connects to the catchment of WWTP Bläsimühle. The foul sewage contribution from this area is negligible; the contributing area is not included.*

RC1.6: Figure 4, left: The legends are hard to read. Are groundwater levels also a part of the dataset? Otherwise consider removing on figure, or indicate where to find data. I cannot see the WWTP on the figure. Is it hidden behind the other signatures?

We improved the legend, clearly labelled the WWTP and added a Reference to the GW data.

We deliberately excluded Groundwater level data to keep the complexity of the data set tractable. Apparently, the hydrogelological situation in the catchment is very complex. Nevrtheless, groundwater level data are available in this repository

https://opendata.eawag.ch/dataset/stein-variational-gradient-descent-fehraltorf/resource/faeaa307-3b42-47b2-adde-6fb59cceff96

Readme: "the hydraulic head data should be in the Python pickle file dictionary_data.p"

which we included in our list of references. We extended the Section 4.2 (now 6.1.2) to:

ll.520ff:

> Future work could use the spatially detailed GWI rates (Ramgraber, 2025) to better capture the sewer-groundwater interactions. This could build on recent work, which added a groundwater module to the SWMM model and simulated dynamic infiltration in Fehraltorf (Rodriguez Bennadji, 2022; Rodriguez et al., 2024).

RC1.7: Line 235ff: Is the maintenance information from the utility company available for the specific measuring period?

No, unfortunately not. No maintenance data from the utility's SCADA system are included, because the small utility (only four employees) does not systematically document sensor maintenance tasks.

RC1.8: Line 375ff: You mention the installation of a flow-limiting hardware. Can you clarify where the device is installed or give a reference to the supplementary information. Wouldn't it affect the measurements?

Thank you for pointing out this issue. We realised that this aspect needs additional and better explanation.

Despite the fact that the implementation of a new flow limiter is documented in the supporting information [*cf.* S3.5 Stormwater treatment facility RUB 59 (RUB ARA): "Until 2020, the inflow to the wastewater treatment plant (WWTP) was monitored using a level sensor and controlled to a calculated maximum inflow of 180 L s-1 through an electric-hydraulic slide gate at the in-sewer flow splitting structure US 58. In August 2020, a new flow control system from *Stebatec AG* (type *TF-PNA16*) was implemented in the main collector just upstream of the sewage treatment plant inlet."] we changed the description in *Section S3.5* to make this aspect clearer. In the manuscript, we provide a link to the Supplement (see line 398, page 15) and added further details there for a more precise description of the drainage situation.

> Initially (until July 2020), the inflow to the WWTP was controlled through an electric-hydraulic slide gate at the in-sewer flow splitting structure *US 58* located about 350 m upstream of the WWTP inlet in the main collector. The flow was monitored just upstream of the WWTP inlet using a level sensor. The level sensor signal was used to control the slide gate so that the WWTP inflow rate did not exceed 180 L s^{-1}. In August 2020, a flow control system (*Stebatec AG*, type TF-PNA16) was implemented in the main collector just upstream of the sewage treatment plant inlet (close to node 597a in the SWMM model). The operation mode of the RUB ARA remained unchanged. Also, the maximum WWTP inflow rate remained at 180 Ls-1, except that the impoundment point moved 'closer' to the WWTP. Thus, the upstream storage volume has increased by approximately 120 m^3.
>
> The installation of the new flow control device does affect the measurements: measured WWTP inflow (now directly derived from a flow measurement of the flow control system) had become more precise. Further, we expect that the additional in-sewer retention volume (in the flow limiter is operated accordingly) may affect the spilling activity at the overflow structure RUB 59 just upstream of the WWTP.

We are hesitant to include such detail in the main manuscript. Now, with the link to the SI in the chapter related to the SWMM model, we hope that this detail is now sufficiently addressed without lengthening the main text in the manuscript.

RC1.9: Line 393: Please clarify what "similar to previous work" means.

This refers to the experience gained from previous Open Data Publications in ESSD. In previous work (Špačková et al., 2021), we also provide a web-based dashboard as a pre-viewer to the data.

RC1.10: Line 444: You mention GWI ranges from 10-15 L/s. But where? At the treatment plant or at all sensor-points?

This refers to the estimated contribution from the Fehraltorf catchment. It is documented in the SI (S9.2) “The analysis allows quantifying the “inner” infiltration within the Fehraltorf catchment, excluding imported groundwater upstream of F02 and F03. This ranges from 10 to 15 Ls-1 depending on the season. Infiltration dynamics are illustrated at F00, using the WWTP inflow as an example.” We adjusted the text accordingly.

ll.512ff:

> To estimate the contribution of groundwater infiltration (GWI) rates from the Fehraltorf catchment, we analyzed long-term flow recordings, focusing on dry weather night-minimum flows and excluding rain-induced infiltration. Using the night-minimum flow, GWI rates were estimated as the difference between GWI at the catchment outlet (F00) and the inflow, i.e. two upstream contributions from Rumlikon and Russikon (F02 and F03).

RC1.11: Line 445ff: How do you know how many manholes the high groundwater table affected? Is it based on traceback assumptions? Please elaborate.

This is based on GIS analyses comparing manhole invert levels with interpolated groundwater levels from the thesis on GW modelling (see SI Section S9.2). As described above, GW data are not part of the dataset, but are openly accessible on ERIC (https://doi.org/10.25678/00035V)

We added the reference here

l.518ff:

> For example, in April 2018, using the data from Ramgraber (2025), a high groundwater table affected 256 out of 459 manhole inverts, while in October 2018, only 100 manholes were impacted (Figure 8).

RC1.12: Line 465ff: Great with an example of how redundant sensors can check data quality, but the text could be improved. I am not sure it is an incorrect level sensor processing, but erroneous sensor settings. Please clarify.

Thank you for the relevant hint. We agree with your comment: the current technical explanation as well as the formulation in the manuscript leaves room for improvement. We rewrote section 4.3 completely to better outline the principle and the value of sensor redundancy/ diversity when monitoring CSO activity. It now is section 6.1.3:

ll.528ff:

6.1.3 The value of redundant sensors in event-duration monitoring

Assessing CSO activity is crucial for quantifying pollution and optimizing sewer networks. Typically, level sensors are used to monitor water levels in the tank and next to overflow weir crests. This inside-tank level data is then used to derive overflow volumes or even spill duration. The latter process can be a very difficult topic, because it essentially means splitting a continuous process into “events”, which are not precisely defined (hysteresis). In addition, a single (level) sensor can drift or change over time due to issues such as temperature influence and the presence of spider webs. In our example, a second, capacitive sensor serves as a rather robust indicator for spill duration, i.e. the start and stop of a spill event. This becomes even more relevant as to date, more and more countries have implemented data-based compliance assessments of CSOs and make spill data, e.g. from overflow event duration monitoring, available to the public (EC, 2022; Rieckermann et al., 2021). However, ensuring data quality is a real challenge and research explores various monitoring techniques to make data-based compliance assessment more reliable.

In the UWO, we equipped all CSO tanks in the Fehraltorf system with multiple ultrasonic level and capacitive sensors to independently monitor overflow duration and to investigate how redundant signals would reduce the uncertainty of CSO event-duration monitoring (SI, Section S9). The results shown in Figure 9 (middle, right) demonstrate that comparing the capacitive sensor signal (bm dl332 | y-axis) with the level signal (bl dl311 | x-axis) greatly improves data confidence and allows for post-calibration. Initially, we severely underestimated the CSO activity at this particular location due to the tank level being recorded with an incorrectly configured level sensor. The opportunity to cross-compare with the ‘same’ spill duration, but derived from a simple yet reliable capacitive sensor signal (see Figure 9, middle), revealed the configuration error (i.e. an incorrect offset). Here, this resulted in 50 % less cumulative overflow duration in a total monitoring period of 1’077 days.

RC1.13: Figure 9, left: The triangles are really hard to read in print version.

Thank you for the comment. We did our best to provide the figure in high quality to meet the journal’s requirements. The triangles in the left plot show binary data at 5-minute resolution, which means that symbols overlap and appear as lines, no matter whether we use squares or even dots. We tried several alternatives but couldnot find a clearer way to show the data without losing its meaning.

RC1.14:Figure 9, middle and right: Please give appropriate titles to the figures.

Titles have been added to clarify what each subfigure represents. In addition, explanations given in the caption have been made clearer.

RC1.15:Line 488 and 489: Should K (-2 K and 5 K) be replaced with Celsius degree?

Yes, thank you. We replaced “K” with “°C” to avoid confusion. The temperature differences are now expressed as: “–2 °C and +5 °C.”

RC1.16:Table 2: You have an asterisk, what does it mean?

Nothing. We removed the Asterisk from year 2018, because it was a left-over from an earlier version of the manuscript and is not needed anymore.

Response to Reviewer #2 (Anonymous)

https://doi.org/10.5194/essd-2024-47-RC2

We thank the anonymous reviewer for the detailed and insightful comments. We have implemented revisions to the structure and clarity of the manuscript. Please find our point-by-point response below.

RC2: 'Comment on essd-2024-47', Anonymous Referee #2, 13 Mar 2025 reply

In this work, the authors presented the Urban Water Observatory dataset. The dataset includes the data of more than 120 sensors over three years in high temporal resolution, complemented by descriptions about the underlying sewer network. Besides the description of the dataset, data access, implemented quality check and some preliminary analyses are presented.

RC2.1: THE UWO dataset represents a comprehensive dataset highly needed in the field of urban drainage networks, and can have therefore a significant impact on future research.

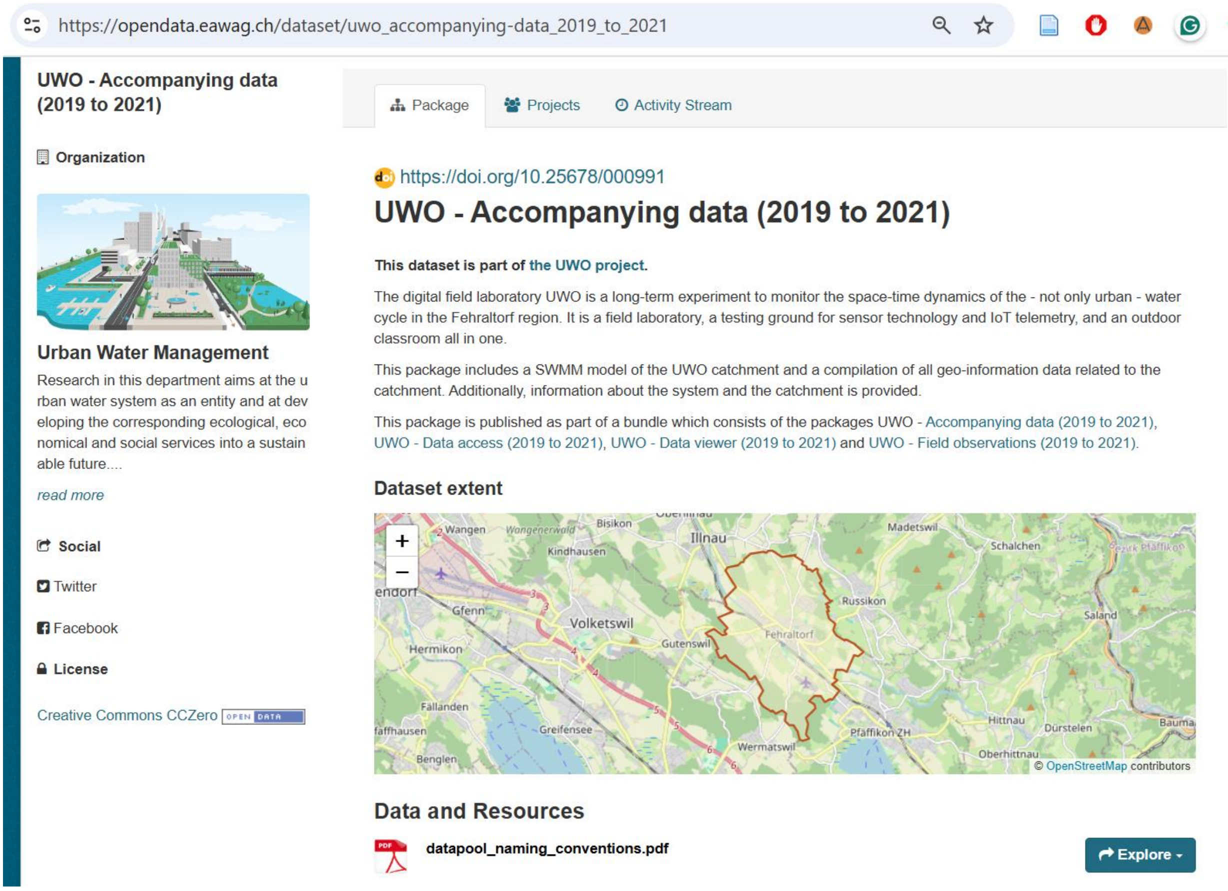
Thank you for this positive assessment of our work. It is in line with the comments we received from user groups who have started to work with our data (see below).

In the invitation to the review, it was mentioned that the provided dataset should be interpretable without looking at the manuscript. Therefore, I started to provide the feedback with the data set first and then continued with the manuscript. Please find detailed comments below, which should be addressed in a revision of this work.

Data base

RC2.2: It is specified in the journal instruction, that “data are distributed under a non-restrictive licence such as CC BY 4.0 or equivalent”. Currently, “No License Provided” is mentioned on the webpages. Therefore, please include a corresponding licence.

We have now clearly stated that all datasets are released under CC0, which is compatible with CC BY 4.0 as per journal policy (see Section 6, Data Availability).



6 Data availability

Data from the Urban Water Observatory are available at the Eawag Research Data Institutional Repository (ERIC/open). We provide the data under a CC0 license, granting users the freedom to use, modify, and distribute it without any restrictions, as the work is dedicated to the public domain. The geodata by swisstopo is supplied with conditions of use, which comply with

RC2.3:

1. In the provided links, there are two databases with different time lengths, namely the downloadable SQL databases with duration from 2019 to 2021 and the database which is accessible via the web platform with data up to now. What is the difference between these two databases besides the different time lengths, what was the motivation for having two different databases, and which database should be used in future? Additionally, it is unclear for me, why the downloadable SQL databases are only until 2021 and a time near data provision would clearly increase the impact of this work. To further increase the impact of this work, **I would also suggest to provide scripts for an access to the database accessible via the web platform, which allows the access to the newer data.**

Thank you. This is a valid point that we need to clarify. We would like to emphasize that the downloadable SQLite databases (packages A) represent a validated and 'frozen' subset of the data from 2019–2021, which has undergone manual and automated quality checks as described in Section 4.3.1. In contrast, the web-based data viewer also provides access to newer, quasi real-time data from ongoing monitoring operations. This real-time data is subject to automated flagging but may contain artefacts or unchecked values, as the implementation of robust, fully automated quality control remains a work in progress—a classic hen-and-egg problem. To minimize the risk of misinterpretation, we currently limit public downloads to validated data and intend to incrementally publish further curated subsets as quality assurance processes mature, in line with FAIR principles, as "living data" (ESSD LD) https://www.earth-system-science-data.net/living_data_process.html.

We changed the last paragraph of section 3.4 (now 4.4 Data access and tools) to:

ll.409-414:

> Public access is granted via ERIC/open, which currently offers only limited data exploration. Therefore, we provide an R-Shiny web dashboard for interactive preview and filtering by signal type, source, or location (Figure 6) (Eawag-SWW, 2025). The downloadable SQLite databases contain a validated and frozen subset (2019–2021) with broad sensor coverage, while the online viewer shows real-time data that may include artefacts, as automated quality control is still evolving. To prevent misuse of unchecked data, we restrict downloads to the validated subset and plan to release further curated data over time.

RC2.4:

2. All the information and files are provided for access to the databases. However, they are downloadable and described on separated webpages, requiring some efforts to understand how they are connected to each other. Therefore, I suggest placing them either all on one webpage or provide an overview of the individual webpages with key points on the content to improve the comprehensibility.

Thank you for the helpful suggestion. In fact, this was precisely the motivation for placing the overview in Figure 1—to clearly illustrate how the individual data packages are structured and connected. We also created a dedicated overview page for the Urban Water Observatory (https://opendata.eawag.ch/group/urban-water-observatory-uwo), which is linked as the first item on each data package page. While we understand the appeal of consolidating all content on a single page, Eawag's internal policy requires separate pages for each data package to ensure proper versioning, updates, and curation by the responsible teams. This modular setup has worked well in practice—several research groups have accessed and used the data without reported issues.

RC2.5:

3. When running the supplementary SWMM files I get the error of the missing rain file "r02_mm_utc0_1min_corr.dat". Besides, it seems that the coordinates for rus are wrong in faf_rus.inp.

Thank you very much for the really careful review. We apologize for this oversight and have corrected this pointer to wrong rain input file ("r02_mm_utc0_1min_corr.dat") in the new versions of the SWMM input files for Rumlikon and Russikon: rum.inp; rus.inp

The corresponding lines in the SWMM input file reads as follows:

Old: [RAINGAGES]

;;Name Format Interval SCF Source

;;-------------- --------- ------ ------ ----------

rain_gauge VOLUME 00:01 1.0 FILE "r02_mm_utc0_1min_corr.dat" r02 MM

New: [RAINGAGES]

;;Name Format Interval SCF Source

;;-------------- --------- ------ ------ ----------

rain_gauge VOLUME 00:01 1.0 FILE "r02_mm_utc0_1min_feb16_apr20.dat" r02 MM

Also the coordinates have been transformed from the historical Swiss LV03 to the concurrent LV95 projection. Below the example of **rus.inp** file original

[illegible]	[illegible]	[illegible]
549	699244.230	250048.540
54a	699186.140	250026.430
550_SW	699367.140	250040.840
550a_SW	699373.770	250041.610
560_SW	699338.130	250009.440
570	699276.210	250002.190
570a	699230.090	250014.820
590b	699371.750	250176.400
590c	699358.870	250144.630
555	699121.480	249997.650
Bacheinl_RÜB_Rum	699327.941	251359.730
RÜBRumlikon	699538.256	251450.663

[VERTICES]

;;Link	X-Coord	Y-Coord
8	699570.373	251434.140
69	699872.083	251548.565
69	699872.083	251548.565
501_SW-403	699758.450	250051.780
501_SW-403	699741.621	250042.006
528_SW-529a_SW	699528.550	250007.030
529_SW-L328	699481.135	250023.548
529_SW-L328	699460.369	249996.391
529a_SW-529_SW	699494.623	250032.328
529a_SW-529_SW	699494.620	250032.330
529b_SW-529_SW	699503.271	250052.471
529b_SW-529_SW	699487.088	250031.334
538_SW-L349	699441.182	250073.505

revised

[illegible]	[illegible]	[illegible]
548	1699309.960	2250061.040
549	1699244.230	2250048.540
54a	1699186.140	2250026.430
550_SW	1699367.140	2250040.840
550a_SW	1699373.770	2250041.610
560_SW	1699338.130	2250009.440
570	1699276.210	2250002.190
570a	1699230.090	2250014.820
590b	1699371.750	2250176.400
590c	1699358.870	2250144.630
555	1699121.480	2249997.650
Bacheinl_RÜB_Rum	1699327.941	2251359.730
RÜBRumlikon	1699538.256	2251450.663

[VERTICES]

;;Link	X-Coord	Y-Coord
8	1699570.373	2251434.140
69	1699872.083	2251548.565
69	1699872.083	2251548.565
501_SW-403	1699758.450	2250051.780
501_SW-403	1699741.621	2250042.006
528_SW-529a_SW	1699528.550	2250007.030
529_SW-L328	1699481.135	2250023.548
529_SW-L328	1699460.369	2249996.391
529a_SW-529_SW	1699494.623	2250032.328
529a_SW-529_SW	1699494.620	2250032.330
529b_SW-529_SW	1699503.271	2250052.471
529b_SW-529_SW	1699487.088	2250031.334
538_SW-L349	1699441.182	2250073.505
538a_SW-538_SW	1699473.620	2250113.315

RC2.6:

4. **Additionally, what does A1 to A4 in the file "data_uwo_sqlite_content_overview.csv" mean? It is described in the manuscript, but it should be also described at the platform.**

Thank you once more for excellent review and for shedding light on this aspect!

"We added this information to the data repository and revised the Readme.md documentation accordingly."

> 'data_uwo_sqlite_content_overview.csv': Information about which package (A1-A4) a source belongs to.
>
> A1: contains all precipitation and further meteorological variables,
>
> A2: all hydraulic measurements,
>
> A3: all temperature readings and
>
> A4: the complete wireless sensor network and its nodes.

RC2.7:

5. I tested the provided Python script for extracting data from the SQL databases, please also include an information, if there are any limitation regarding Python version and if yes, which Python version is needed for running the scripts. Further suggestions to improve the usability are to include some exemplary information in the main function as comment, e.g., "# data_uwo_2019.sql in the folder uwo_data_slice" for add_argument filename and "#data_uwo_sqlite_content_overview.csv in the folder uwo_data_slice" for add_argument contentlist.

We updated the README to state compatibility with Python ≥3.8. Example commands and inline comments have been added. Below screenshots from the file.

```
# coding: utf-8
# Requires Python 3.8 or higher


import pendulum
import argparse
import pathlib
```

```
def example_query_1(db_file: str) -> pd.DataFrame:
    """List all datapoints recorded between 'start_date' and 'end_date' at location '11e_russikerstr'."""

    end_date = pendulum.datetime(year=2021, month=9, day=30)
    start_date = end_date.subtract(days=1)

    location = "11e_russikerstr"

    example_query = f"""

    SELECT
        signal.timestamp,
        value,
        unit,
        variable.name,
        source_type.name,
        source.name
    FROM signal
        INNER JOIN site ON signal.site_id = site.site_id
        INNER JOIN variable ON signal.variable_id = variable.variable_id
```

```
if __name__ == "__main__":

    parser = argparse.ArgumentParser()
    parser.add_argument(
        "-sd", "--sourcedirectory",
        default="./uwo_data_slices",
        help="Directory containing the data files. Example: ./uwo_data_slices")
    parser.add_argument(
        "-fn",
        "--filename",
        default="dp.sqlite",
        help="SQLite database filename in the folder 'uwo_data_slices'. Example: data_UWO_2019-01_2020-01.sqlite"
        )
    parser.add_argument(
        "-cl",
        "--contentlist",
        default="./package_information.csv",
        help="CSV file listing the sources per package (A1 to A4) in the folder 'uwo_data_slices'. Example: data_uwo_sqlite_content_overview.csv"
        )

    args = parser.parse_args()
```

~~~~~

Manuscript

In general, the manuscript is well written and understandable. However, the structure and content of the individual sections as well as their differences are not always comprehensible:

RC2.8:
~~~~~

1. This is for example obvious for Section 2 Material and Section 3 Methods, which both include parts of introduced data pipeline in Figure 3 (e.g., for the sensors and the data collection it is referred to section 2, while data warehouse with data quality checks are section 3). Also, the content is described unevenly in these two sections (e.g., much more description for the utilised sensors, while the description of the data warehouse is really short and referred to the supplementary for more information). Therefore, I recommend to combine these two sections into one section (e.g., Materials and Methods), to shorten the text and to focus on the most important information required for the readers (for detailed information, it can be referred to the Supplementary as already done by the authors) to improve the comprehensibility.

Thank you for your helpful suggestion. We agree that the original structure blurred the boundary between sensor deployment and data processing. Rather than merging Sections 2 and 3, **we have restructured the manuscript to better distinguish between the real-world system** (sensor network and observations) **and the "digital system"** (data pipeline, storage, and quality control). The new structure is inspired by similar papers on ESSD, such as the CAMELS-DE (https://doi.org/10.5194/essd-16-5625-2024), CAMELS-CH (https://essd.copernicus.org/articles/15/5755/2023/) and the *Bellinge* dataset (sewer oriented) (https://essd.copernicus.org/articles/13/4779/2021/) and includes:

- Dedicated sections on the **Catchment** and **Sensor Data,** focusing on physical infrastructure and measurements.
- A separate section on the digital system **Data Pipeline and Quality Assurance** aligned with Figure 5, covering data collection, QA/QC, and access tools.

Here is a short overview on the new structure:

1. **Introduction**

2. **Description of the Catchment**
   - Description of the study area
   - The urban wastewater system and the River Luppmen/Kempt

3. **Sensor Data**
   - Overview of sensors and installation
   - Measurement types and locations
   - Sensor maintenance and metadata

4. **Data Pipeline and Quality Assurance**
   - Data collection and transmission (Figure 5)
   - Data storage and structure (Datapool)
   - Quality control (range/gradient checks, consistency)
   - Data access and tools

5. **Simulation Model**
   - SWMM implementation
   - Model calibration

6. **Research Opportunities**
   - Applications of the datasets and model
   - Future research directions

7. Conclusions

This structure improves clarity for both infrastructure-focused and data-focused readers, while addressing your comment about uneven detail. We also included a new section of future use of the data, also requested in comment (RC2.10), and improved the cross-referencing to the Supplementary Information.

RC2.9:

2. Since the structure of this manuscript was not presented in the introduction, I would have expected that the manuscript follows the standard structure for articles, including a section on results and discussions. However, in the current version, the results and discussion section is missing, which makes the results of this work less comprehensible and unfortunately also reduces the impact of this work from my point of view. From the Introduction, I would have expected the presentation of the UWO dataset including some first analysis as the main result of this work, while from the descriptions in the Method section, it would be the analysis of the performance of the implemented processes for data quality checks. Besides, results and discussion can be found throughout the manuscript (e.g., evaluation of LoRaWAN ranges in method). Therefore, I highly recommend to include a section “Results and Discussion” focusing on the provided UWO dataset to better present this great dataset, and extending it with already performed analysis throughout the manuscript. By doing so, this should also make the conclusions drawn in section 7 better comprehensible.

Including a dedicated “Results and Discussion” section is an interesting suggestions. We fully agree that many ESSD articles do present results in the form of derived or preprocessed data products, especially in the domain of remote sensing where a novel method to process satellite data, yields new insights (XREFs). However, this manuscript presents a “pure” observational dataset, as described in the ESSD-Editorial (https://essd.copernicus.org/articles/10/2275/2018/) and thus also follows the typical structure of **observational infrastructure and open dataset** papers, such as the CAMELS datasets or the Bellinge dataset (see RC2.8), where the **main result is the dataset itself**, not a new processing method or analytical insight.

In our case, the contribution lies in:

- the **design and implementation** of a long-term, high-resolution urban hydrology observatory;
- the **provision and documentation** of curated raw and quality-flagged data;
- and the **enabling of future research**, which we illustrate through use cases and example opportunities in Section 4 and the Supplement (cf. S9 - Research opportunities).

We acknowledge that this may not have been sufficiently clear in the introduction and conclusions. To address this, we have:

- Clarified upfront that the **dataset and infrastructure** are the primary outcomes of the work.

ll.104-108:

> The remainder of the article is structured as follows (Figure 1): we first describe the catchment of Fehraltorf and then the sensor data. We then describe the methods used to collect, process and explore the data and the accompanying data, including a hydraulic rainfall-runoff model implemented in SWMM. Finally, we highlight five exemplary research opportunities . We emphasize that the main novelty of this work is the dataset itself, not the described methods for data curation and cleaning or the provided examples, which motivate future research on the UWO dataset.

- Modified the **"Research Opportunities"** section into “examples” and “future research” to explain the role of the presented applications as illustrative, yet not conclusive, analyses.

- Improved the linkage between these examples and the conclusions, to highlight how the dataset supports broader reuse.

We hope this clarification better reflects the manuscript's alignment with the ESSD scope and intention.

RC2.10:

3. The aim of section 4 is to present some possible research opportunities. However, the descriptions are more about presenting previous research performed by the authors, while there is a lack of concrete research opportunities to the scientific community noticeable. For me, this analysis could be part of the Results and Discussions section discussed in the previous comment. Additionally, I highly recommend to include also concrete research questions/task into the manuscript.

Thank you for this valuable suggestion. We agree that the goal of section "Research Opportunities" is to support future use of the UWO dataset by the wider scientific community. The section currently draws on example analyses from our own team to demonstrate, as explicitly requested by ESSD Editors, "the validity and applicability of [our] datasets and data products."[1] — but we understand that this may have read more as retrospective results than forward-looking opportunities.

To address this we revised the original Section 4 (now 6) and structured it into "Applications of the dataset" and "Future research".

- The "applications" section emphasizes that we provide illustrative applications of the data as use cases intended to inspire reuse.
- In the "future research", we provide explicit research questions to highlight and guide potential directions for future work (e.g., anomaly detection, sensor performance evaluation, hydrological modeling). We added future research opportunities in a dedicated subsection, rather than moving them into a traditional "Results and Discussion" section, in line with ESSD conventions for dataset-focused papers.

We hope this revision makes the section more useful and engaging for potential dataset users, and we appreciate your suggestion to suggest concrete future research opportunities.

ll. 597ff:

6.2 Future research directions

[1] https://essd.copernicus.org/articles/10/2275/2018/ 3.5 Extensive validations
Authors will need to demonstrate, first to reviewers and later to a wide range of users, the validity and applicability of their datasets and data products. Exact mechanisms and options for validation will vary substantially among and across data products. Because ESSD serves to ensure the suitability of published data for future research, each ESSD paper should demonstrate skill and utility of the submitted data product [...].

In the future, our drainage systems should be more adaptive and resilient and we should have a good understanding of their performance. Future research should therefore focus on optimizing sensor placement for both understanding and control, developing predictive models that account for uncertainty, and adopting adaptive monitoring strategies. Advancing FAIR data infrastructure will further support broad, intuitive access and reuse of high-resolution observational data, e.g. for machine learning and data-driven modelling.

6.2.1 Optimizing sensor placement and sampling for system understanding and control

Monitoring networks in urban drainage systems are important to understand the behaviour of the system and to enable real-time control, but they are also challenging to maintain, and manage. Future research should therefore focus on methods for sensor placement that maximize the value of the data collected. Key research questions include: “How many locations should we monitor with which measurement type to capture critical flow dynamics?” “Which monitoring points provide little (added) value?”. Addressing these questions requires a systematic assessment of the information contribution of each sensor location, possibly using information theory or graph-based methods (Crowley et al., 2025; Villez et al., 2016).

Also, future research should investigate adaptive spatio-temporal sampling strategies. Sewer observation using uniform high-frequency monitoring is rarely sustainable, as it consumes power and continuously changing batteries puts a practical limit to the number of sensors a utility can maintain. Future monitoring strategies should therefore be adaptive, adjusting logging intervals based on system needs and external triggers, e.g. weather radar. Key questions include “When should sensors switch to high-frequency mode?” and “How can targeted sampling schemes for sewer capacity or groundwater infiltration or routine monitoring look like?” The UWO dataset is dense, well-documented and, with the accompanying SWMM model, provides an excellent foundation for sensor placement.

6.2.2. Adapting Urban Drainage to Climate Change with Grey-Green Infrastructure

Urban drainage systems must be adapted to climate change using smart combinations of grey infrastructure and Blue-Green Infrastructure (BGI), yet the interaction between these systems remains poorly understood. Future research should therefore focus on quantifying the contribution of BGI to overall system performance, especially under extreme weather conditions and in the context of climate resilience. Key research questions include: “How effectively does BGI reduce peak runoff under extreme rainfall?” and “Can BGI also support thermal energy recovery or mitigate urban heat stress?” To address these questions, detailed hydraulic and thermal simulations are needed that account for both conventional and nature-based infrastructure. Combining the available rainfall, flow, and temperature measurements with a model such as SWMM-HEAT —offers a unique opportunity to simulate such future climate scenarios and assess the performance of integrated grey-green solutions.

6.2.3 Developing predictive models and Real-Time Control strategies

Real-time control (RTC) of urban drainage systems depends on the ability to predict system states reliably under a wide range of rainfall and runoff conditions. Future research could use the UWO data to develop predictive models that support effective control strategies. Central research questions include: “Which types of rainfall events are most critical for triggering control actions?”, “How can we use monitoring data, which are never 100% accurate, to quantify the performance of RTC systems?”, “Can observed flow patterns be used to predict system states and guide real-time decisions?”. One very interesting direction, which is not widely discussed, is “Can we use the SCADA system to identify important characteristics of the full-scale system?” where controlled interventions in the full-scale system can actively test sensor performance and system behaviour. For example, operators can trigger actuators like pumps or gates and directly observe the sensor response to identify delays, outliers, or faults (Sant’Anna et al., 2024).

6.2.4 Accounting for uncertainty and improving statistical inference

Reliable model predictions depend on properly accounting for uncertainty especially given the fact that urban drainage models contain structural bias (Del Giudice et al., 2013, 2015). Also, flow and rainfall measurements are also often systematically wrong, e.g. through offset, calibration shifts or drift (Del Giudice et al., 2016). Future research could therefore use the provided data packages to explore what methods are promising to describe prediction uncertainty in a statistically correct way, e.g. conformal predictions (Vovk et al., 2025) and how autocorrelation, multiple data sources and input errors from rain gauges can be accounted for (Auer et al., 2024; Sun and Yu, 2022).

6.2.5 Enabling smart data infrastructures

Despite the increasing volumes of monitoring data, urban drainage datasets remain underused due to insufficient metadata, inconsistent formats, and limited user access. To unlock their full potential, future research should invest in FAIR data infrastructure that enables intuitive, domain-aware interaction with complex environmental data. One promising avenue is the development of interfaces that translate natural language queries into structured database queries (text-to-SQL), allowing users to retrieve relevant data without deep technical knowledge of the schema (Allemang and Sequeda, 2024). Ideally, one could directly ask the database “Which pipe is at capacity?”, “What was the annuity of the thunderstorm last Friday?” and “Which city district has the highest infiltration rate?” To support such functionality, data models must be semantically rich and ideally rely on ontologies and standardized knowledge graphs. A starting point could be the Dutch Urban Drainage ontology (GWSW) (The Dutch Urban Drainage Ontology (GWSW), 2025), which defines classes and relationships within urban drainage systems and is already integrated with national tools for asset management and modelling. As it is mostly focused on asset representation rather than on sensor time series, further semantic developments are needed to support such queries.

RC2.11:

4. The conclusions drawn by the authors are not always comprehensible, as they are discussed for the first time in the manuscript (e.g., learning effects, …). Maybe this is already solved by the recommended inclusion of a Results and Discussion section, otherwise the authors can also include a subsection with a critical discussion about their learnings.

Although reviewer #1 did not suggest modifications to our conclusions, we agree with the reviewer #2 that it is not good practice to bring up entirely new information, e.g. on the ontologies. We addressed this by including the subsection on “Future research”, which makes it possible to reduce the original “Conclusion and Outlook” to a “Conclusion” section.

The key conclusions are:

- Open data is still the exception—structured data sharing remains rare.
- IoT-based technology greatly helps, but it does not replace traditional monitoring- a balanced combination of few reliable sensors plus many low-cost ones give both accuracy and spatial insight.
- More flow sensors require more rain gauges to interpret runoff patterns.

- Real-time data only adds value if validated—automated checks help, but manual checks are still needed.
- The UWO data is relevant, which we demonstrate by several applications, such as understanding infiltration, CSO, sewer heat exchange and also for the telecommunications domain.
- There are a lot of promising future research possibilities, especially regarding ML development, uncertainty quantification, testing BGI impacts and smart databases which can make sewer data truly usable.

The complete changes are:

ll.668ff:

- Urban drainage monitoring is evolving, but despite progress in data acquisition and processing, widespread open data sharing is still the exception rather than the rule.
- IoT-prone transmission technologies and advances in sensor development offer the opportunity to revolutionize sewer process monitoring by enabling real-time, spatially differentiated information. However, they do not replace traditional methods and are not yet a plug-and-play solution for ubiquitous sensing. Power supply of many sensors is still a major bottlenecks and limits scalability, especially for small utilities with few resources.
- In terms of sensors, we conclude that a dual-sensor strategy pays off: using a few high-quality sensors for key measurements (e.g., rainfall, pipe flow) ensures trustworthy reference data, while low-cost sensors, such as ultrasonic water level sensors support dense deployment and gathering spatial information, which then provides a deeper understanding how the urban drainage system functions. It is important that sensor deployment must be supported by adequate rainfall and climate monitoring, especially if spatially resolved runoff patterns are to be interpreted meaningfully. Simply put: More flow sensors generally necessitate more rain gauges.
- To fully benefit from monitoring data, it is crucial to ensure that streaming the sensor data in real time is sufficiently reliable. Collecting field observations without timely validation does not make sense, and regular *manual* checks for consistency and homogeneity remain essential to ensure data accuracy. We found that *automated* quality checks, such as range and gradient tests, are adequate to detect anomalies and become particularly important when handling more than 30 to 40 sensor signals. Differentiating whether anomalies are caused by sensor malfunction or actual system behaviour requires more advanced analysis which still has to be standardized.
- For important urban drainage processes, such as sewer infiltration, overflow behaviour and thermal energy exchange, the suggested applications of the UWO dataset demonstrate that high-resolution, long-term monitoring is essential to understand the relevant dynamics. In addition, the provided telemetry data offer a real-world benchmark for assessing IoT performance from underground environments, enabling further research on optimizing network design, power use, and data reliability in smart city applications.

- Future research should use the UWO dataset to develop and benchmark ML-based anomaly detection and data recovery methods using the long time series. It also seems very promising to investigate in how far BGI can help us to make our urban drainage systems more resilient and adaptable and especially to test modern methods of uncertainty quantification, such as conformal predictions. Future work should also focus on smart databases by combining semantic web technologies with data-driven modelling, e.g. Large Language Models. Urban drainage ontologies and standardized descriptions of monitoring data will probably improve integration, interoperability and even reasoning. Combined with natural language interfaces, these systems could allow users to query sewer data more intuitively, e.g. “Which pipe is at capacity?”, “What was the annuity of the heavy storm last Friday?”

RC2.12:

5. For a better reference to the individual subfigures, I recommend including subheadings such as (a), (b) ,… instead of top/bottom.

Very good suggestion. Adjusted as recommended.

RC2.13:

6. Figure and table captions are quite long and should be shorten. Instead, the figure and table contain should be described in more details in text.

Here, we disagree slightly. Informative captions which give the main message of the figure make it easier for the reader to screen the paper whether the content is relevant. However, we shortened the captions as much as possible.

Figure 2a): Overview of the Fehraltorf catchment, including urban drainage network, geography and land use. b): Simplified flow scheme of the combined sewer system of Fehraltorf. The main flow path through the network is from top right to the left (bold line). The relevant characteristics of sub-catchments can be found in the supporting information. Indicated flow times are calculated assuming a constant flow velocity in the sewer of 1 ms-1.

Figure 3a): CAD model of the overflow structure “RÜB Morgental” including inflow chamber, detention tank, overflow to receiving water, and installed sensors, b): Cross-section at the sensor bl_plsRKBM1201_rubmorg_inflow (not drawn to scale). Detailed information of all other detention basins and flow control structures is provided in the SI.

Figure 4a): Georeferenced locations of sensors, b): Example of a heat map representing the completeness of data in the monitoring period from 01-01-2019 to 31-12-2021. A1 = precipitation data; A2 = hydraulic data; A3 = temperature data. A4= LPWAN sensor data is not shown. The color saturation indicates the degree of data completeness (weekly granularity). Dark blue indicates periods with 100 % data completeness; white indicates periods with no data. Light blue indicates a reduced number of data points, either through sensor maintenance, sensor outage or incomplete data transmission. We provide a dynamic plot in package C, which can be used to interactively explore the data availability and view details. See SI Section 6 for details. Groundwater levels are not part of this dataset, but available at ERIC (Ramgraber, 2025).

Table 1: Overview of installed sensors including relevant characteristics on sensor type, temporal granularity of data, and the type of data transmission.

Figure 5: The individual components of the data pipeline for collecting, storing, and using UWO data from the sensor to the data service. The numbers refer to the corresponding Sections.

Figure 7, left: Observations with potential anomalies. Highlighted sections show the benefit of multiple sensors in detecting anomalies. While the high variability on the left is reflected by all sensors, on the right only a single signal appears abnormal. Right: F1 scores for different methods for pre-processing (A-D). An F1 score of 0.5-0.8 is considered medium quality, 0.8-0.9 good and above 0.9 excellent. For real-world data (A-C), the performance increases with increasing levels of pre-processing. For real-world data, ARIMA performs best and the Autoencoder never reaches the performance on synthetic data (D).

Figure 8, left: In total 256 sewer manholes in the Fehraltorf network are potentially affected by groundwater (April 2018). The darker the colour, the more are manholes submerged in GW. Right: in total 100 sewer manholes were affected by groundwater in October 2018.

Figure 9a): Continuous tank level measurements (dots) and binary data (triangles). The latter are derived from the capacitive sensor signal, reflecting an overflow activity during a period of two days for which two independent overflow events were recorded, b) and c): Event-specific overflow durations derived from one capacitive sensor (bm_dl332_rub_morg) and one ultrasonic level sensor (bl_dl311_rubmorg_inflow). The overflow duration derived from the erroneously calibrated level sensor (dl311) with an incorrect offset in b) is corrected to a correct offset in c).

RC2.14:

7. As described in the manuscript composition, footnotes should be avoided in the text.

Agree. This has been addressed in RC1.1.

RC2.15:

8. L17: Would be ‘barriers’ a better word instead of ‘deterrents.

Yes, this would indeed be much better. We replaced it.

RC2.16:

9. L32: From the descriptions, it is unclear for me how the development of ontologies and knowledge graphs can help to extend the application of sewer observation data. Besides, knowledge graph is only mentioned in the abstract.

Yes, we agree it’s not ideal to introduce new ideas in the conclusions. That said, the current success of large language models relies on massive, well-structured data collections—only possible if data are semantically labeled. Ontologies provide shared concepts and relations, allowing unique identifiers for things like sewer flow data. Knowledge graphs help structure metadata (e.g., diameter, slope), enabling automated checks using simple physical models. Since LLM integration with ontologies has progressed during the review phase, we think it is appropriate to mention this as a future research direction.

ll.655ff:

6.2.5 Enabling smart data infrastructures

Despite the increasing volumes of monitoring data, urban drainage datasets remain underused due to insufficient metadata, inconsistent formats, and limited user access. To unlock their full potential, future research should invest in FAIR data infrastructure that enables intuitive, domain-aware interaction with complex environmental data. One promising avenue is the development of interfaces that translate natural language queries into structured database queries (text-to-SQL), allowing users to retrieve relevant data without deep technical knowledge of the schema (XREF). Ideally, one could directly ask the database “Which pipe is at capacity?”, “What was the annuity of the thunder storm last Friday?” and “Which city district has the highest infiltration rate?” To support such functionality, data models must be semantically rich and ideally rely on ontologies and standardized knowledge graphs. A starting point could be the Dutch GWSW ontology (XREF), which defines classes and relationships within urban drainage systems and is already integrated with national tools for asset management and modelling. As it is mostly focused on asset representation rather than on sensor time series, further semantic developments are needed to support such queries.

Further reading in this direction, which we did not integrate into our article to keep it tracktable, is

- A popular very recent study suggests that LLMs struggle with formal logical reasoning and depend heavily on pattern matching. https://arxiv.org/abs/2410.05229
- Well-designed ontologies (vocabularies) make environmental and geospatial data FAIR and machine-reasonable, improving reuse and automated interpretation. https://www.semantic-web-journal.net/system/files/swj653.pdf
- There is evidence that investing knowledge graphs, namely the ontology, provides higher accuracy for LLM powered question answering systems.https://arxiv.org/abs/2405.11706

RC2.17:

10. L38 - 52: In my opinion, the lists of challenges require more detailed explanations to make them comprehensible without detailed knowledge of drainage networks (e.g., from the literature review it is not clear, why there is a need for better understanding the rainfall-runoff processes).

We understand this, and added some concise explanations to clarify key points without increasing the length too much. More detailed information is available in the referenced literature.

ll. 45-56

> "However, reliably collecting minute-by-minute rainfall data over many years can be challenging, because rain gauges are prone to clogging, especially from leaves, insects, or vandalism. Moreover, urban drainage catchments are often small and heterogeneous, so a single rain gauge cannot reliably capture localized, convective rainfall events—multiple, well-placed gauges are usually needed. Monitoring stormwater runoff, wastewater flows, and pollution processes at a minute scale is equally resource-consuming, as it requires high-frequency sensors that must operate reliably in harsh environments such as sewer pipes. Specialized equipment, training, and software (Dürrenmatt et al., 2013; Mourad and Bertrand-Krajewski, 2002), as well as considerable investments, are required to collect and manage the data, and to maintain the sensors (Blumensaat et al., 2019; Hoppe et al., 2016). Arguably, sensor maintenance to ensure good data quality is one of the biggest challenges in urban monitoring, as sensors can drift, foul, or lose power without notice (Mourad and Bertrand-Krajewski, 2002; Nedergaard Pedersen et al., 2021). Data quality is often dubious due to the low data literacy of the sewer workforce and lack of incentives to use data for evidence-based management of urban drainage systems (Manny et al., 2021). In addition, the lack of standards and meta-data makes it difficult to work with existing or historical data, especially when trying to reuse them across different systems, tools, or studies."

RC2.18:

11. L43: Please add information which high-resolution data is needed, e.g., temporal, spatial, or both.

Thank you for the comment. The type of high-resolution data needed depends on the application. For rainfall-runoff processes, temporal resolution is generally more important to capture rapid dynamics during storm events. In contrast, spatial resolution is more critical for understanding infiltration patterns. For combined sewer overflows (CSOs), both temporal and spatial resolution are relevant.

We added a sentence and reference that explains this for the rainfall-runoff process:

ll.43ff:

> “Understanding urban hydrological processes require high-resolution data on both, the input - such as rainfall - and the output - such as wastewater flows and pollution. For rainfall-runoff, a high temporal resolution seems to be more important than spatial resolution, though both interact closely (Ochoa-Rodriguez et al., 2015). However, reliably collecting minute-by-minute rainfall data over many years can be challenging (Bianchi et al., 2013),”

RC2.19:

12. L50: I would suggest to rephrase “often dubious due to the low data literacy of the sewer workforce” with something such as “is influenced by errors in the data knowledge of employees”.

Yes. We replaced it with:

ll.53ff:

> "Data quality is often affected by limited familiarity of operational staff with data handling and interpretation as well as a lack of incentives to use data for evidence-based management."

RC2.20:

13. L61: The descriptions how the following factor has “fueled” the demand for open data sets can be improved, e.g., for me the advancements in lower-power electronics have opened new ways in data collection rather than increasing the need for open data sets.

Yes, that is a good point. We re-phrase to describe the “push” for open datasets, which describes it much better.

ll.65-74:

> In recent times, several developments have contributed to a growing push for open datasets in urban hydrology. First, advancements in low-power electronics, data transmission (Ebi et al., 2019) and sensor application (Boebel et al., 2023; Mathis et al., 2022) have significantly reduced the effort of data collection. Second, efforts for standardized meta-data and exchange formats facilitate data sharing (Bustamante et al., 2021; Taylor et al., 2013). Third, scientific data collection efforts, such as the data set “Catchment Attributes and Meteorology for Large-sample Studies (CAMELS)” (Addor et al., 2017; Newman et al., 2015), have truly revolutionized the field of hydrology through the use of advanced data-driven models (Kratzert et al., 2018). Fourth, the public demands greater transparency of urban infrastructure performance (Benyon, 2013; Giakoumis and Voulvoulis, 2023) and regulatory bodies, such as those in the UK (Environment Act, 2021) and in the EU (EC, 2022), are demanding more monitoring. At the same time industry initiatives, such as STREAM (Stream - Portal, 2025), further emphasize the importance of collecting and sharing water company data.

RC2.21:

14. Please specify, what the main novelty of this work is compared to the mention projects CAMEL and STREAM, as it seems, they are having the same aim?

This is a good question. CAMELS focuses on river hydrology on the national level and is not suited for urban or in-sewer applications. STREAM is industry-driven and aims at transparency, but only provides aggregated summaries (e.g. number of overflows per year), not full time series or raw data needed for scientific analysis and modelling. So the ideas are similar (CAMELS- scientific focus, STREAM- urban drainage focus), but do not provide both, which is the main novelty of the UWO dataset.

RC2.22:

15. L87: For me, the description about the provided data packages belong to the methods section rather than in the Introduction. Instead, I would expect an outlook, how the manuscript is structured.

Agree. Implemented with RC2.9.

RC2.23:

16. Fig. 1 und Fig. 5 are quite redundant and can be combined into one.

Thank you for the suggestion. We respectfully disagree. Fig. 1 outlines the structure of the manuscript and the overall data concept, while Fig. 5 focuses on the actual datasets, their spatial and temporal coverage and availability. Since they serve different purposes—one conceptual, one data-driven—we think that both are really needed. Also no similar concerns were raised by other reviewers.

RC2.24:

17. Fig2: It seems that the top is missing. I also suggest to include the meaning of triangles (instead of the Figure caption) and circles in the legend.

    Yes, thank you. That mistake occurred during final formatting. Next time we'll use Latex again…

RC2.25:

18. L119: Reference is made to Fig 2., but none of this information is shown there.

    Fixed with RC2.24.

RC2.26:

19. L121 - 126: Instead of the historical weather values, the authors can include the measured values from the dataset. Besides, what does a high variability and frequency mean?

Good comment. We now include measured annual rainfall from our dataset (FAF) and the nearby MeteoSwss station (KLO):

2019: FAF: 1096 mm, KLO: 1073 mm

2020: FAF: 1059 mm, KLO: 833 mm

2021: FAF: 1339 mm, KLO: 1083 mm

which confirms that our measurements are within the typical range of long-term annual rainfall. The phrase "high variability and frequency" refers to short, intense rainfall events, which are common in convective storms. We added a reference to Section S1.6, where we briefly describe the new extreme value statistics from MeteoSwiss (not available at submission). For the nearby station KLO, tt estimates the intensity for a 10-year event is 119 mm/h (95 % CI: 107–141 mm/h) for a 5min duration. For a 20-year return period, it rises to 134 mm/h (CI: 118–168 mm/h). Starting in May 2025, MeteoSwiss is gradually making its data accessible as Open Government Data.

ll.135ff:

> Our measurements show annual rainfall totals between 1096–1339 mm, compared to 833–1083 mm from the MeteoSwiss station KLO. The climate region has a high variability in weather patterns and the frequency of storms and extreme weather events. The intensity for a 5-minute rainfall with a 20-year return period is estimated at 134 mm/h (95% CI: 118–168 mm/h) according to new extreme value statistics from MeteoSwiss (see Section S1.6).

RC2.27:

20. L142 + Fig. 3: Is the information about RUB Morgental needed in the manuscript?

Yes, we firmly believe that it is needed for completeness. In our view, it also nicely illustrates the level of detail we provide and the effort that went into curating the dataset and metadata. We improved the readability of the annotations (see RC1.4).

RC2.28:

21. Figure 4: The legend is hardly readable, and how are sensor and sensor nodes distinguished? Additionally, I recommend to include names of rivers and special structures, as spatial information in combination with Fig 2. Instead, the right part of the Figure would belong the results and not to Materials.

We increased the size of the legend and included the names of rivers and special structures. As Figure 4a shows the spatial coverage, we believe that it is reasonable to show the temporal coverage in 4b. We deleted sensor nodes in the caption to avoid confusion.

In the UWO dataset, the difference between a sensor and a sensor node is:

- Sensor: The actual device that measures a physical parameter (e.g. water level, temperature, flow, rainfall).
- Sensor node: A complete unit that includes the sensor plus supporting components—typically a logger, power supply (e.g. battery), and communication module (e.g. LoRaWAN, cellular).

This is described in ll.189-190.

RC2.29:

22. L201: What does in contrast to Bellinge dataset mean?

“In contrast to Bellinge dataset” means that the UWO dataset differs in its focus and design priorities. In contrast to the Bellinge dataset, which is a rich and comprehensive utility dataset, even including CCTV inspections and two hydrodynamic models, the UWO dataset was purpose-built as a scientific field lab. This design makes UWO better suited for understanding fine-scale urban hydrological processes, especially where detailed and reliable rainfall-runoff observations are essential—such as anomaly detection, infiltration studies, and sensor performance evaluation.

We extended the description to:

ll.219-221:

> “In contrast to the Bellinge rainfall data, which come from a national network (Nedergaard Pedersen et al., 2021) and are likely quality-controlled and processed using standardized tools, the UWO rain gauges were installed specifically for local monitoring, providing more targeted and detailed rainfall-runoff data although their measurements have not been quality controlled a priori.”

RC2.30:

23. L263: Is it right, that operating LPWANs has become straight forward? In contrast, why did the author reduce the number of sensors, developed the LoRa-based mesh technology and recommend a monitoring backbone?

Thank you. We interpret and address this comment as follows: *why didn’t we (the authors) implement even more sensors if installing LPWAN sensor networks is so straightforward?*

As a matter of fact the ever growing IoT community not only develops new approaches but also collects more experience in existing, yet new techniques. As actors gain more insights and experience, technological solutions become more robust and more comfortable to be used. Our team had the chance to witness and shape (hopefully a bit) this evolution in an early stage. Nowadays large-scale rollouts in the context of Smart City projects exist [e.g. Amsterdam, Zurich, Basle, Berlin, Hamburg, etc.], underlining that the LPWAN techniques made their way from a niche technology to an established, flexible and cost-effective monitoring approach especially in urban areas. Still, implementation and maintenance of a what-so-ever number of sensors is effort and consumes resources. Since also our resources and access to skilled technicians were limited, we had to reduce the number of sensors following the constraints given by the available budget. In addition to that, we were always keen on minimising the numbers of sensors that have limited value, because high maintenance efforts are difficult to justify if the sensor ultimately contributes little information. We started to develop the LoRa-based mesh technology in 2016, almost a decade ago, to overcome limitations we identified with the standard LoRaWAN technique when monitoring underground infrastructure, which is a rather special application case. Systematically we reduced those sensors in the standard LoRaWAN which had poor to mediocre performance. We replaced them with mesh sensors as long as we had our own prototypes available.

Overall, it can be unambiguously stated that operating LPWANs has become straightforward, but more advanced sensors (e.g. flow monitors as part of the so-called *backbone*) are still required to provide rather accurate reference measurements. The challenge to optimise the monitoring layout beyond how we have done this is formulated as a research opportunity in Section 6.2.1.

RC2.31:

24. L330: Could the authors give more details about the quality check – what was the motivation to focus on range and a gradient check, what was used for the valid range (e.g., system boundaries, measurement range, calibration range) and maximum or minimum gradient to distinguish between measurement fluctuations with ultrasonic sensors and rainfall event, …?

Sewer flow and level data follow rather complex dynamics. Despite the fact we did research more advanced data validation techniques, we identified that simple data validation rules (range, gradient checks) were sufficient and robust (!) enough to identify questionable data. A further differentiation of "questionable" data often requires expert knowledge, and - so far and to the best of our knowledge - no reliable more advanced data validation routine for in-sewer process data is available. Hence we formulated this task as research opportunity in Section 6.1.1 and S 9.1.

To address your comment we added the following sentence:

ll.352ff:

> Upper and lower limits for range checks are individual for each sensor as they are defined through i) the device specification and configuration as well as ii) constraints given through actual installation in the field.

… and the following at line 336:

> The values implemented in gradient tests vary depending on the observed system dynamics, i.e. they were manually adjusted to obtain meaningful results.

RC2.32:

25. L350: Please add details about the types of regular check, how often were they performed, how is data consistency and homogeneity defined, and what conditions have to be fulfilled to classify the data as wrong?

The procedure we refer to in Section 3.2.2 (now 4.3.2) aims at ensuring data consistency and homogeneity by regular semi-automated data validation.

More concretely that is accomplished by i) a regular (every morning), visual inspection of sensor data time series focusing at new sensor data that had arrived overnight. With the growing number of sensor data acquired, this process was/is supported by ii) an automated flagging routine based on range and gradient checks. Methodological details are given in Section 4.3.1.

Eventually, the daily (workdays only) data inspection routine, carried out by a person in charge, enabled us to identify gross errors in the recent data in quasi real-time, and to timely address these errors by on-demand sensor inspections in the field. While basic automated data validation techniques (step ii) help to filter out "questionable data", expert knowledge is used to further differentiate phenomena (sensor failures *or* system-immanent process dynamics) in "questionable data" and identify errors and the need for sensor maintenance (step i).

We added the following section after the first paragraph in Section 4.3.2

ll.373-395:

> The continuous verification of consistency and homogeneity helps to ensure that the UWO dataset reflects real-world processes adequately across time and space. In this context we define consistency and homogeneity as follows:
>
> **Data consistency** refers to the sufficient availability of plausible data over time. Specifically, data should not contain large gaps, inexplicable jumps or unrealistic values that conflict with the known behaviour or physical constraints of an urban drainage system, e.g. manhole water levels beyond terrain level.
>
> **Data homogeneity** refers to the uniformity of data characteristics over the observation period. For example, a sensor's signal should consistently represent the same physical quantity under the same conditions. Changes in installation, sensor type, or surrounding hydraulics that could affect the signal interpretation were tracked to ensure the dataset remains temporally coherent.
>
> More concretely, we accomplish this by i) a visual inspection of sensor data time series every morning, reviewing new sensor data that has arrived overnight in the context with historic data. With the growing number of UWO sensors, i.e. data collected, this process was supported by ii) an automated flagging routine based on range and gradient checks (see Section 4.3.1). This daily inspection routine enabled us to identify anomalies in quasi real-time (e.g. due to sensor malfunction, clogging, configuration changes, or physical alterations to the sewer system), and to timely address these phenomena by on-demand sensor maintenance in the field. While basic automated data validation routine (step ii) help to filter out "questionable data", expert knowledge is used in i) to further differentiate observed phenomena (is it a sensor failure *or* just typical, normal system behavior?) in "questionable" data and identify errors and the need for sensor maintenance. We differentiate between "doubtful" and "undoubtedly" data points. The so as "doubtful" identified data were kept in the dataset but annotated, enabling users to exclude or inspect them as needed. Clemens et al., 2021 propose more classes, but we found that this was difficult to implement in our case.

RC2.33:

26. L354 – 357: This paragraph should belong to the Future Research section as it is discussing future research opportunities.

Yes, this is a good point. We integrated this into the Future Research section.

RC2.34:

27. L363: Could you please describe which data (period, sensors, …) was used for the calibration.

This information has been provided in *Appendix S7 - Hydrodynamic sewer models*. Basically, since 2016, the SWMM model has been thoroughly revised continuously: dry weather flows, storage volumes, pump curves, and control logic were verified, while pipe and manhole geometries largely followed the official municipal cadastre from the original MIKE Urban model.

The final calibration of the provided SWMM model was based on measured inflow to the WWTP and flow observations at four additional locations during several non-extreme rain events (March–May 2016). Model performance was evaluated using cumulative volume bias and Nash-Sutcliffe efficiency (NSE), which is described in *Section S7.3* and *Figure S40*.

ll.457ff:

> The final calibration of the provided SWMM model was based on measured inflow to the WWTP and flow observations at four additional locations during several non-extreme rain events (March–May 2016). Details are described in Section S7.3 and Figure S40.

RC2.35:

28. L375: Is it correct, that the calibrated base model has a different underlying hydraulic structure then the provided dataset? If yes, could you please provide a description about the modifications and the expected impact (e.g., which sensors should correspond with the simulations and for which sensors there is a high difference expected). This would be a really important information for future work integrating hydrodynamic model and measurement data.

The given calibrated base model has a very similar structure as the provided data set. The only major limitation is the fact that the STEBATEC pneumatic flow limiter (is not implemented in the Model. We addressed this by including more details in the supporting information S3.5 and explicitly mentioning it in the main manuscript. See answer to RC1.8.

ll.467-471:

The SWMM model structure files as well as input files, flow patterns are available in package B. As with any other drainage system, the sewer network in Fehraltorf is subject to changes. For instance, with the commissioning of new flow-limiting hardware in August 2020, modifications were made to the network that are not yet considered in the currently provided model structure (cf. SI Section S3.5). Further information on the rainfall-runoff model, as well as on the accompanying data can be found in the SI.

RC2.36:

29. L379: Section 3.4 shows some overlapping with section 6 Data availability. To avoid these overlaps, I recommend to combine these two sections in a joint section.

Yes, this makes sense. This has been merged into the new structure and is now considered in Section 4.4 “Data access and tools”

RC2.37:

30. Figure 9: The total overflow duration is the same in the middle and right chart, although there should be a difference according to L467. Based on the results, why is the capacitive sensor signal not directly used for estimating the total overflow duration, and what is the need of the level sensor?

Thank you for pointing this out. We agree that the formulation in L466 ff is a little misleading. To be clearer we adopted the text in the manuscript as follows:

ll.540-551:

“In the UWO, we equipped all CSO tanks in the Fehraltorf system with multiple ultrasonic level and capacitive sensors to independently monitor overflow duration to investigate how redundant signals would reduce the uncertainty of CSO event-duration monitoring (SI Section S9.3).

The results shown in Figure 9 b) and c) illustrate that the CSO duration derived from an erroneously calibrated level sensor (Fig. 9b - dl311, x-axis) can significantly differ from the CSO duration derived from a correctly configured level sensor (Fig. 9c - dl311, x-axis). Only by comparing the level-sensor prone information with the CSO duration derived from the capacitive sensor signal (reference signal), the systematic deviation due to the wrongly configured level sensor becomes obvious. In the given example, with the incorrect level sensor configuration the CSO activity is severely underestimated - about 50 % less cumulative overflow duration is documented in the overall monitoring period of 1’077 days. Generally, the redundant information from a robust capacitive sensor (only dry/wet status; consecutive wet intervals are counted to give an overflow duration) helps to verify data from typically implemented level sensors. Once successfully verified, level data can then be used to quantify not only overflow duration but also overflow volume, i.e. spill rates. Capacitive sensors alone provide reliable information on overflow duration and overflow frequency.

RC2.38:

31. L485: Please include a definition of headspace and bulk liquid temperatures and a description how condensation and evaporation affect the sewer heat transfer processes for a better understanding, as it remains unclear from the descriptions.

We added the definitions. l.567f

Headspace temperature refers to the air above the wastewater surface in the sewer pipe, bulk liquid temperature to the wastewater itself.

RC2.39:

32. L487: In Figure 10, only the temperature is shown and it remains unclear, how the author determines if there is a real evaporation or condensation. From my understanding, the humidity is an important factor affecting these processes. Additionally, it is unclear, what medium evaporates or condensate.

This is correct, we forgot to include important information in the description of this section, which is only available in the Supporting Information of Figueroa et al.

Basically, the latent heat transfer analysis in the Supplementary Information was intended as a **didactical demonstration**, rather than a direct representation of dynamic field conditions.

To clarify:

a) As shown in Figure 2 (Figueroa SI, see below), we measured relative humidity (RH) in the sewer headspace at four locations in the Urban Water Observatory (UWO). The observations show that RH typically ranges from 83% to 98%, but only rarely reaches full saturation (100%)—even during humid summer conditions. This suggests that net latent heat fluxes (evaporation or condensation) are indeed possible under real operating conditions.

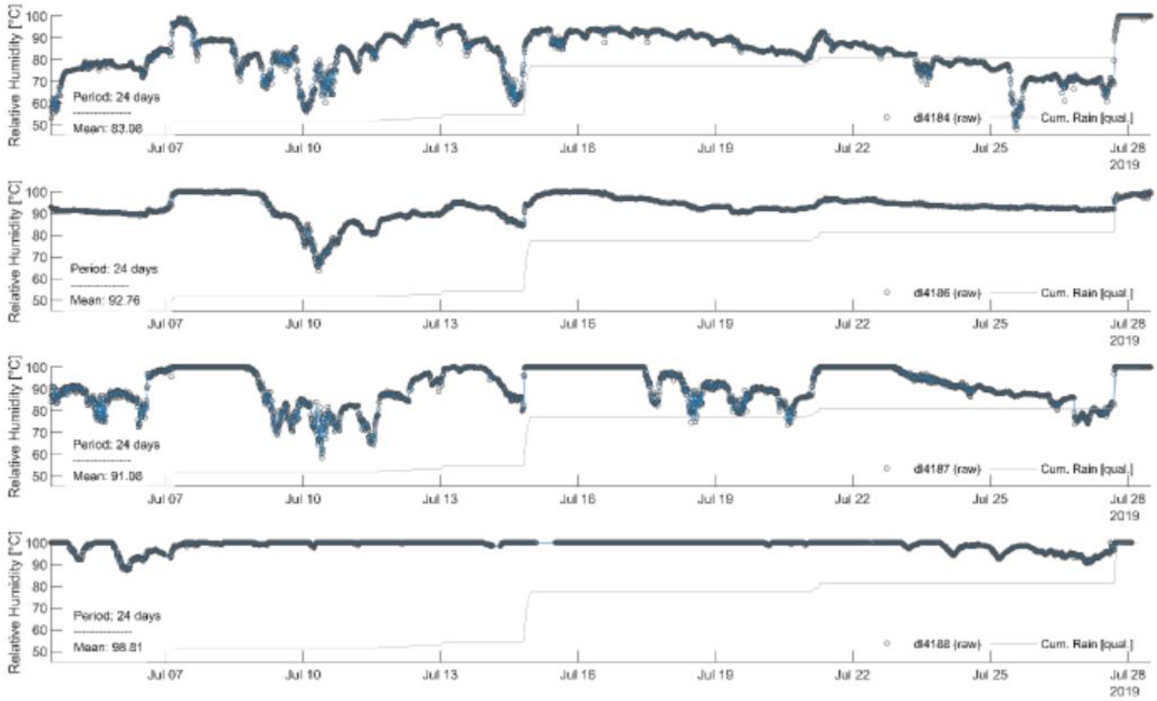


Figure 2: Relative humidity observed at four locations in the in-sewer headspace in Fehraltorf, Switzerland. Grey lines indicate the cumulative rain depth as qualitative information. Mean relative humidity during the period of July 2019 ranges from 83 % to 98 %. During dry weather periods, values drop down to 70 % at some locations (dl4184, dl4187). In turn, rain events, e.g. occurring at 15 July 2019, lead to an increase up to 100 %.

To analyse the effect of these processes, we made a simplifying assumption:

We treated the relative humidity as stationary with a fixed value of 91.4%, which corresponds to the mean RH across the observed locations in the above Figure.

Using this fixed RH, we classified field observations of sewer headspace and bulk liquid temperatures into periods of:

- Evaporation (when T_headspace > T_liquid, shown in red), and
- Condensation (T_headspace > T_liquid, shown in blue)
  as visualized in the corresponding figure.

We then selected two representative periods for detailed exploration (indicated by green and yellow boxes in the plot). These were used to simulate two contrasting thermodynamic scenarios, where we examined how bulk liquid temperature responds to variations in:

- Relative velocity between air and water surface ($rvel$),

- Relative humidity ($rhum$),
- and Bulk liquid depth.

As an example, here are the results of Figure 5

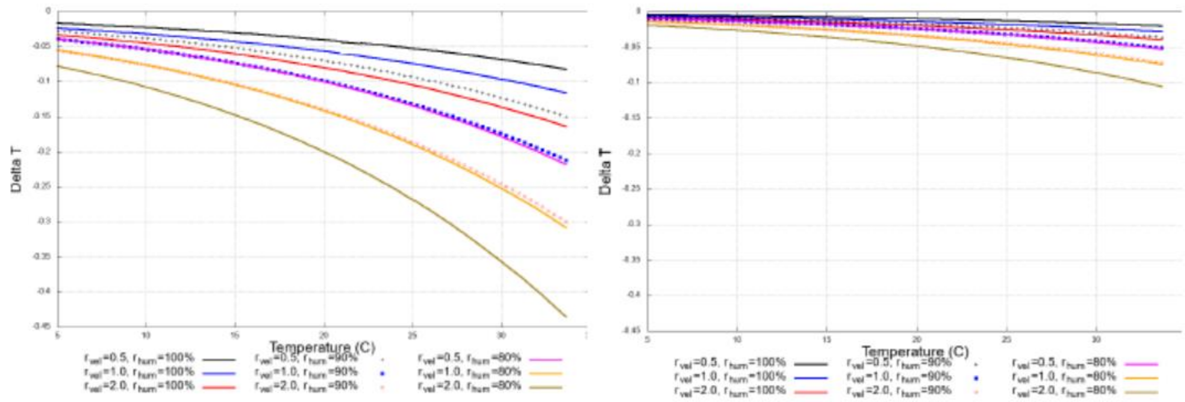


Figure 5 Change of bulk liquid temperature ΔT. Left: Bulk liquid depth = 5% of the pipe diameter. Right: Bulk liquid depth = 20% of the pipe diameter.

To avoid this confusion, we suggest the following changes:

ll.565ff:

> Using a subset of the provided temperature time series, we investigated the difference between headspace and bulk liquid temperatures to identify conditions for A) condensation and B) evaporation (SI of Figueroa et al. (2021)). Headspace temperature refers to the air above the wastewater surface in the sewer pipe, bulk liquid temperature to the wastewater itself. To do this, we analyse the provided temperature data in three locations dl933, dl935 and dl931 (Figure 10), assuming a fixed relative humidity of 91.4%, which is the mean value observed across several sewer locations (Figueroa et al. (2021) Figure 2, SI). This simplification enabled a didactical classification of temperature observations into evaporation (red) and condensation (blue) periods, based solely on the temperature gradient between the headspace and the wastewater.
>
> We found that condensation takes place in April when the typical temperature difference between the sewer headspace and the bulk liquid temperature is -2 °C. In contrast, with warmer ambient air and headspace temperatures, evaporation occurs. Maximum differences can amount to 5 °C, at location dl931. Based on these results, Figueroa et al. (2021) concluded that latent heat transfer should not be neglected, especially in areas with shallow flows and high velocity differences between wastewater and headspace, i.e. in steep catchments and peripheral regions. Even in scenarios with high relative humidity values, latent heat processes play a crucial role and should be considered. This makes the provided temperature

data an ideal source for further developing heat exchange models for sewer networks.

RC2.40:

33. Figure 10: The y-labels have the unit [°C], while the temperature differences in the text are °K.

Thank you, we fixed this (see RC2.39)

RC2.41:

34. L518: It should be section 5 instead of section 6, which is the same for the following sections.

Thank you, we fixed this.

RC2.42:

35. Author contributions: Andreas Scheidegger and Uwe Schmitt are mentioned here, but they are not listed in L5.

Thank you for spotting this. This was a leftover from an earlier version of the paper. Given their more technical support role, it is more appropriate that they are mentioned in the acknowledgements.