

Digitalization meets reality – Concept and Experiences from long-term wireless data collection with 50+ sewer monitors

Frank Blumensaat^{1,2*}, Andreas Scheidegger², Christian Ebi², Simon Dicht², Fabian Schaltegger³,
Andreas Rüst³, Uwe Schmitt⁴ and Max Maurer^{1,2}

¹ Institute of Environmental Engineering, Chair of Urban Water Systems, ETH Zurich, Switzerland

² Swiss Federal Institute of Aquatic Science and Technology, Eawag, Switzerland

³ ZHAW Winterthur, Institute of Embedded Systems, Switzerland

⁴ Scientific IT Services, ETH Zurich, Weinbergstrasse 11, 8092 Zurich, Switzerland

Abstract: Blossoming visions foresee an ubiquitous availability of data enabled through emerging sensors, novel transmission technologies and revisited data analysis techniques. It is often expected that this likewise applies to urban water management (Kerkez *et al.*, 2016; Eggimann *et al.*, 2017). Contrary to the envisioned easiness to collect data, an increasing number of studies report about a declining effort to conduct experimental field research in catchment and urban hydrology, implying the risk of scientific advancement to be misdirected (Blume *et al.*, 2017; van Emmerik *et al.*, 2018). In this paper, we present results from a unique sensor network that collects, transmits, and organizes very high resolution status information from an urban drainage catchment. Our findings are based on a two-year operation at full-scale with more than 50 sensor nodes connected to a low-power wireless sensor network (LPWSN). In essence we i) illustrate the scalability of LPWSNs, ii) reveal telemetry limitations of the LoRaWAN standard (LoRa@Alliance 2015), iii) suggest a novel technique (LoRaMesh) to master this constraint, and iv) discuss challenges associated with managing the amount of data collected from diverse sources. We find: 1. the proposed transmission technique considerably improves telemetry performance of the entire network without introducing additional gateway stations (total packet loss 2.2 %). 2. Dynamic and flexible data management solutions come at a price. 3. It is not the pure amount of data that is most challenging but the diversity of different signals. 4. Full metadata integration in the data warehouse is laborious but clearly consolidates data interpretation in the long-term run.

Keywords: distributed monitoring; low-power wireless sensor networks; spatio-temporal in-sewer dynamics; LoRaWAN; data management

1. Introduction

Miniaturization of hardware components, increasing computational capacities and an omnipresent integration of various types of technology in our everyday life promise a data plethora in the near future. It is often expected that this likewise applies to urban water management (Kerkez *et al.*, 2016; Eggimann *et al.*, 2017). Despite such novel techniques that evolved through digital transition, experimental field work in hydrology remains hard. Our contribution aims at discussing this apparent contradiction from a conceptual and a practical point of view. With sharing our experiences from the Urban Water Observatory (UWO)¹ a long-term (two years) distributed data collection campaign, we intend to test the visions of digitalization against reality.

In the following we focus on two central aspects of UWO: i) the performance and improvements of the wireless underground transmission and ii) how the diverse data obtained can be efficiently structured and managed (see Figure 1). We previously reported on specific technical details of the UWO including sensor technology, network layout and principles of the low-power transmission technique LoRaWAN (see Blumensaat *et al.*, 2017).

¹ www.eawag.ch/uwo

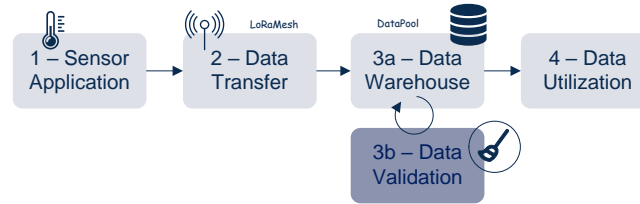


Figure 1: schematic overview data collection process as referred to in this paper.

2. MATERIAL AND MOTIVATION

Sensor network evolution: the 60 monitors deployed in the UWO can be partitioned into a) 9 backbone monitors, i.e. conventional rain gages, flow monitors operated on batteries and GSM data loggers, and b) a low-power wireless sensor network (LPWSN) consisting of two LoRaWAN gateways and 51 low-power sensor nodes that monitor in-sewer dynamics, e.g. water levels, temperature, conductivity. With the establishment of the first backbone monitors in Feb. 2016, and the ongoing LPWSN rollout started we steadily increased the number of UWO monitors since May 2016. To date we collect 37'000 data points per day. This corresponds to overall 123 monitoring signals representing the system status in a 5min resolution.

Data transmission: A telemetry performance analysis of the LoRaWAN network revealed an average data packet loss of approximately 10% through imperfect transmission (1 year of data; 23 of 34 *radio* nodes located underground). While this is an achievement, particularly for underground nodes, there is room for improvement. An in-depth examination shows that the Quality of Service (QoS) linearly decreases with increasing distance to the preferred gateway, and it depends on whether a radio node is positioned above or below ground. For underground nodes, a critical distance to the preferred gateway was found to be at approximately 500 m (see Fig. 2). Interestingly, varying weather conditions (rainfall) do not significantly influence the transmission.

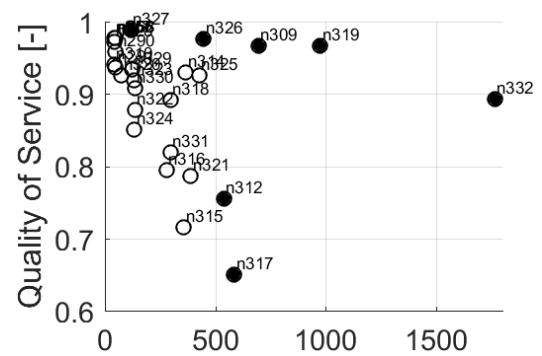


Figure 2: Quality of Service [-] (defined as fraction of packets reaching gateway) vs. distance from gateway [m] for original LoRaWAN transmission (Data: Jan – Dec 2017). Filled circles indicate above-ground radio nodes; empty circles sensor nodes positioned below ground.

Data Management: With an increasing amount and diversity of observations, the data management concept gains importance. Diversity in type of sensors, (often undocumented) file formats, units, time zones (daylight saving) increase the efforts to systematically unify the data and systematically link meta-data to it. Only by linking the observations with metadata the data can be utilized including for open data publication. While certain tasks cannot be automated (e.g. writing scripts to convert between different file formats), a good data management system ensures that such efforts are not duplicated by every analysts. Our justified hypothesis is that a systematic and consistent data management requires high efforts in the start-up period, but it pays off in the long-term as it increases the quality of the data, and it makes data usage much more efficient.

3. METHODS, RESULTS, DISCUSSION

Data transmission: To address the LoRaWAN telemetry limitation we developed a new transmission protocol 'LoRaMesh' that allows forming individual sub-networks by introducing intermediate repeater nodes (RN). This way, we modify the original LPWAN architecture from an exclusively star-type (LoRaWAN) towards a tree-type network topology, enabling a multi-hop transmission and thus achieving a more resilient routing. Still, both protocols are based on the same physical layer (LoRa®); the deployed radio node hardware is identical.

Verified through in a repeated field test (operation period: 45 days, 11 test locations) we provide a proof-of-concept for a LoRa-based LPWSN extended by the *LoRaMesh* protocol. Figure 3 illustrates the enhanced telemetry performance where we achieve a reduced packet loss of, in average 2.2 % compared to the standard LoRaWAN protocol (10 %).

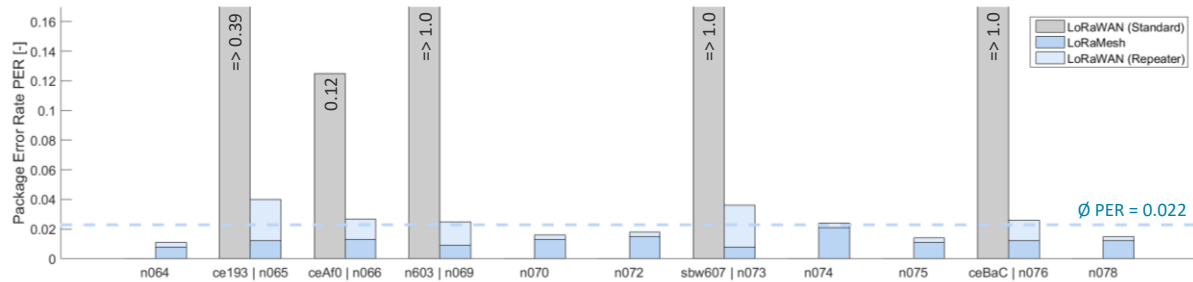


Figure 3: comparison of the total Packet Error Rate (PER = 1 - QoS) for selected underground monitoring locations (node IDs at x-axis) for i) the standard LoRaWAN network (grey bars) and ii) the, with *LoRaMesh* extended network. An PER of ≥ 1 indicates that there was no radio coverage at all before introducing the extended routing.

Data Warehouse application: Existing data management solutions are generally designed with a very specific work- and dataflow in mind – which is most unlikely to correspond to exactly *your* needs. So even while the required database concepts are very common and well-established, every application needs a careful technical specification, a design concept and an individual configuration. For the UWO we aimed at a solution that i) can handle diverse data of potentially not yet known sources and formats, ii) stores and links data and metadata simultaneously, iii) ensures consistency of the data whenever possible, iv) makes data usage at the front-end easy, and v) does not require in-depth IT skills to add novel data sources. Note, the flexibility of point i) limits the potential for automatization and makes point v) inevitably harder.

The *Datapool*², our data management solution is based on separation of concerns: a field scientist (cf. *data provider* - Figure 4) adding a new sensor is responsible for i) providing a file with meta data and ii) a conversion script (scripting language R, Python, Julia or Matlab) to transform raw files from the sensor into a standardized text format. As soon a new raw file arrives from a sensor the *Datapool* applies automatically the corresponding script and imports the data into a relational database (here: PostgreSQL). Finally, *data users* can either directly SQL query the (meta)data, use a handy R/Python/Matlab package with a basic set of query functions, or use a web interface³. To make this rather simple workflow robust, *Datapool* must essentially perform consistency checks, provide error logs, be designed to guarantee scalability, and manage read/write-permission of the different users.

² <https://datapool.readthedocs.io/en/latest/>

³ www.uwo-opendata.eawag.ch

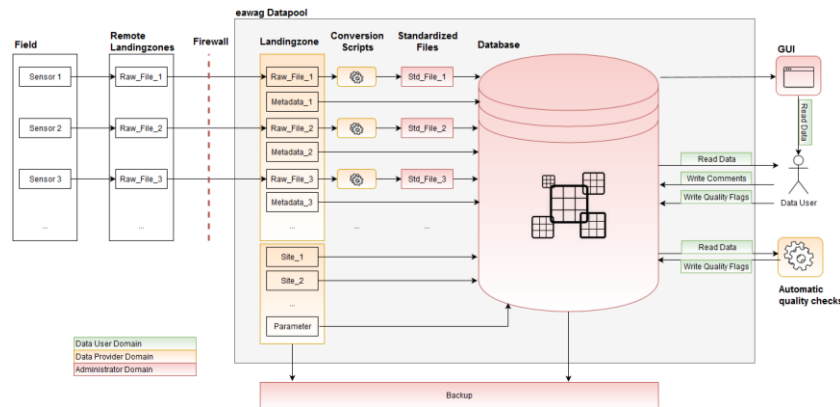


Figure 4: Dataflow of the data warehouse application *DataPool*.

4. CONCLUSION

The recent technical developments make the “data driven urban water management” envisioned by Eggimann *et al.* (2017) more feasible than ever. Still, many of these novel techniques are in its infancy or need adaptations for the application in sewers, requiring a considerable engineering and scientific efforts before a wide spread adaption becomes possible. By providing concepts and sharing real-world experiences we hope to contribute to a new era of data collection in drainage networks with various useful implications, also for the modelling community.

ACKNOWLEDGEMENTS

The authors are thankful for support from the municipality Fehraltorf (CH), the LoRaWAN hardware developers from *Decentlab* (Dübendorf, CH), the backend experts from *Ioriot.io* (Thalwil, CH), and we particularly thank Philipp Bachmann (Civilian Servant at Eawag) for his tremendous contribution to the final hardware prototyping phase.

REFERENCES

- Blume, T., van Meerveld, I., and Weiler, M. (2017). "The role of experimental work in hydrological sciences – insights from a community survey." *Hydrological Sciences Journal*, 62(3), 334-337.
- Blumensaat, F., Ebi, C., Dicht, S., Rieckermann, J., and Maurer, M. (2017). "Highly Distributed Long-Term Monitoring of in-Sewer Dynamics Using Low-Power Radio Technology." 12th IWA Specialized Conference on Instrumentation, Control and Automation, P. A. Vanrolleghem, ed., IWA, Québec City, Québec, Canada.
- Camhy, D., Gamerith, V., Steffelbauer, D., Muschalla, D., and Gruber, G. "Scientific Data Management with Open Source Tools – An Urban Drainage Example." *9th International Conference on Urban Drainage Modelling*, Belgrade.
- Eggimann, S., Mutzner, L., Wani, O., Schneider, M. Y., Spuhler, D., Moy de Vitry, M., Beutler, P., and Maurer, M. (2017). "The Potential of Knowing More: A Review of Data-Driven Urban Water Management." *Environmental Science & Technology*, 51(5), 2538-2553.
- Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., Bedig, A., Kertesz, R., Braun, T., Cadwalader, O., Poresky, A., and Pak, C. (2016). "Smarter Stormwater Systems." *Environmental Science & Technology*, 50(14), 7267-7273.
- LoRa@Alliance. (2015). "LoRaWAN™ What is it? - A technical overview of LoRa® and LoRaWAN™." LoRa® Alliance.
- Sonnenberg, H., Rustler, M., Riechel, M., Caradot, N., Rouault, P., and Matzinger, A. (2013). "Best data handling practices in water-related research." *Water Practice and Technology*, 8(3-4).
- Tarboton, D., Horsburgh, J., and R. Maidment, D. (2008). *CUAHSI Community Observations Data Model (ODM) Version 1.1 Design Specifications*.
- van Emmerik, T., Popp, A., Solcerova, A., Müller, H., and Hut, R. (2018). *Reporting negative results to stimulate experimental hydrology*.