Benchmarks monitoring algorithms for Non-Intrusive Water Monitoring

Andrea Cominola\textsuperscript{1}, Marie-Philine Becker\textsuperscript{2}, and Riccardo Taormina\textsuperscript{3}

\textsuperscript{1}Technische Universität Berlin - Einstein Center Digital Future, Chair of Smart Water Networks, Berlin, Germany (andrea.cominola@tu-berlin.de)
\textsuperscript{2}Technische Universität Berlin - Einstein Center Digital Future, Chair of Smart Water Networks, Berlin, Germany (marie-philine.becker@campus.tu-berlin.de)
\textsuperscript{3}Delft University of Technology, The Netherlands (r.taormina@tudelft.nl)

As several cities all over the world face the exacerbating challenges posed by climate change, population growth, and urbanization, it becomes clear how increased water security and more resilient urban water systems can be achieved by optimizing the use of water resources and minimize losses and inefficient usage. In the literature, there is growing evidence about the potential of demand management programs to complement supply-side interventions and foster more efficient water use behaviors. A new boost to demand management is offered by the ongoing digitalization of the water utility sector, which facilitates accurate measuring and estimation of urban water demands down to the scale of individual end-uses of residential water consumers (e.g., showering, watering). This high-resolution data can play a pivotal role in supporting demand-side management programs, fostering more efficient and sustainable water uses, and prompting the detection of anomalous behaviors (e.g., leakages, faulty meters). The problem of deriving individual end-use consumption traces from the composite signal recorded by single-point meters installed at the inlet of each household has been studied for nearly 30 years in the electricity field (Non-Intrusive Load Monitoring). Conversely, the similar disaggregation problem in the water sector - here called Non-Intrusive Water Monitoring (NIWM) - is still a very open research challenge. Most of the state-of-the-art end-use disaggregation algorithms still need an intrusive calibration or time-consuming expert-based manual processing. Moreover, the limited availability of large-scale open datasets with end-use ground truth data has so far greatly limited the development and benchmarking of NIWM methods.

In this work, we comparatively test the suitability of different machine learning algorithms to perform NIWM. First, we formulate the NIWM problem both as a regression problem, where water consumption traces are processed as continuous time-series, and a classification problem, where individual water use events are associated to one or more end use labels. Second, a number of algorithms based on the last trends in Artificial Intelligence and Machine Learning are tested both on synthetic and real-world data, including state-of-the-art tree-based and Deep Learning methods. Synthetic water end-use time series generated with the STReaM stochastic simulation model are considered for algorithm testing, along with labelled real-world data from the
Residential End Uses of Water, Version 2, database by the Water Research Foundation. Finally, the performance of the different NIWM algorithms is comparatively assessed with metrics that include (i) NIWM accuracy, (ii) computational cost, and (iii) amount of needed training data.