https://doi.org/10.71573/ny2ybb51

Fast urban drainage model generator using global open data sources

Thibaud Maruejouls[3], *Delmas Simon*[3], *Abdelghani Zaid*[3], *Philippe Ginestet*[2], *Marcello Serrao*[2 3]* https://orcid.org/0009-0001-9798-4223 & *Wolfgang Rauch*[1] https://orcid.org/0000-0002-6462-2832

[1]*University Innsbruck, Department of Urban Drainage Modelling, 6020 Innsbruck, Austria*
[2] *SUEZ International, Engineering & Construction – Innovation & Technical Office, 92040 Paris La Défense, France*
[3] *SUEZ, Le LyRE R&D Center, 33600 Pessac, France*
**Corresponding author email:* simon.delmas@suez.com

Abstract
Accurate flowrate data in a catchment area is essential for city stakeholders that needs to plan future developments for the next decades. Hydraulics models are powerful tools to rely on for such decisions. The lack of models around the world is mainly due to their development cost which is directly correlated to the data volume requirement. The rise of open-source datasets for geography, population, urban infrastructures provide an opportunity to deploy urban drainage on almost any point of the earth. This study aims at developing two automatic model generators, the first one using GIS dataset as input aims at being robust to poorly filled GIS while the second is solely relying on data from publicly accessible websites for auto configuration and calibration. The codes were applied to reproduce the urban drainage system of Dijon Metropole for which the simulated flowrates were compared to observed flowrates at Dijon water resource recovery facility inlet.

Key words:
urban drainage planification, hydraulics model, GIS gap filling, data requirement

Highlights
- Estimation of key design parameters using hydraulics model.
- Methodology applied on a real pilot site comparing simulations to observations.
- Automatic hydraulics model generation

Introduction
Being able to monitor and even anticipate urban drainage (UD) flows is gaining more important to tackle future crucial issues for cities such as: flooding risk assessment, UD impacts on river quality, UD and WWTP infrastructure design. Hydraulics models are very powerful tools that cities aim at developing to better manage its own system. However, modellers require extensive datasets (asset locations and depth, dimensions, material, rainfall, flows…) to have the proper knowledge of the system to build and maintain models making their cost extensive (Elga *et al*., 2015).

Public institutions have made important efforts in the recent years to propose open-source datasets for demography, geographical, and weather systems such as European Spatial Agency (ESA), National Aeronautics and Space Administration (NASA), United State Geological Survey (USGS), Copernicus program… These efforts have unlocked data availability required for UD models at various scales, reaching even global earth scale with high spatial and/or temporal resolution.

Automatic, or close-to, hydraulics model generators have recently been developed in several domains: drinking water (www.spatial-insights.co.uk, www.qatium.com), hydrology (www.pcswmm.com). Few investigations have already been investigated for urban drainage systems. Möderl *et al*. (2009) proposed a methodology for generating stochastic urban drainage simulations that can create virtual case studies. Schilling *et al*. (2022) proposed a QGIS (https://qgis.org/community/organisation/) add-on converting a GIS dataset into a SWMM configuration .INP file. This powerful tool allows to speed-up model generation. However, it needs a proper GIS quality to provide fast models, not being resilient enough to dataset gaps (missing pipes, diameters or elevations). Reyes *et al*. (2023) or Dobson et al (2025) developed software that generates a sewer infrastructure GIS based on assumptions relying on city infrastructures (roads, buildings, etc.). These tools need improvements to be widely used without too many additional efforts and to provide reliable or usable results.
To overcome data lack and hazardous assumptions, a stochastic approach is used to provide a probability flowrate range rather than a deterministic result.

This study presents the development of a tool aimed at generating UD models for almost any part of the world, regardless of whether an existing GIS is available. This tool relies on open-source external datasets for rainfall, topography, sewer infrastructure probability in urban areas, soil occupancy and population. The algorithm is resilient to low GIS data availability. The models generated were compared to observed data to assess their performance.

Methodology

Two algorithms were developed to generate stochastic SWMM (Rossman, 2004) models. The first algorithm utilizes an existing GIS dataset, while the second relies solely on open-source data with limited information about the actual system. These models were implemented within the Dijon Metropole area, and their performance was evaluated by comparing the simulated outlet UD system flowrates with observed flowrates recorded throughout the years 2021 to 2025.

Dijon Metropole case study description

The UD system of Dijon is the pilot site where the generated model flowrate results will be compared to observed flow rate at the ODIVEA WWTP inlet. The population of Dijon is 207,906 inhabitants (2020), the UD system is a mix of combined sewers (272 km), wastewater network (181 km) and stormwater network (119 km). One WRRF (400,000 PE, Q_{max} 129,600 m^3/d) receives the effluent from Dijon Metropole and six other towns. Dijon has an average annual precipitation of 743 mm.

Open-source data

Various open-source datasets were scanned and analysed for: 1) rainfall intensities: ERA5 from Copernicus programme, 2) topography: NASADEM HGT v001 from NASA,; land cover: ESA Worldcover from Copernicus, European Spatial agency, 3) city infrastructures (Open Street Map); GIS (ODIVEA from Dijon Metropole); population density (Worldpop).

Model generation

The sewer model comprises hydrologic modelling and infrastructure network development.
The method employs the D8 algorithm (O'Callaghan and Mark, 1984) to delineate watersheds within defined urban boundaries. The hydrologic component integrates watershed characteristics from DEM and landcover, with watershed outlets defined at lowest elevation points that are proximally located to the WWTPs.
The network is obtained using GIS data, from manhole elevations, roof and pipe diameters and length. Missing data are estimated through interpolation using information from adjacent nodes, traversing the network graph sequentially from one neighbouring node to the next. The network is also

compressed, while preserving the geometric characteristics of each branch to improve simulation times.
The model integration assigns the nearest manhole to each subcatchment's natural outlet to serve as the outlet for receiving dry weather and infiltration flows. The model is then encoded into SWMM-compatible INP files.

Results and discussion

Figure 2 compares the SWMM subcatchments delineation manually made by an engineer through dedicated study with the subcatchments delineation generated by the algorithm. The number of catchments is initially 244 while the generated ones are 30. The resolution of catchments could be modified to be smaller, but they were found enough detailed to model runoff with an overall average size of 63 ha versus 12ha for the initial model.

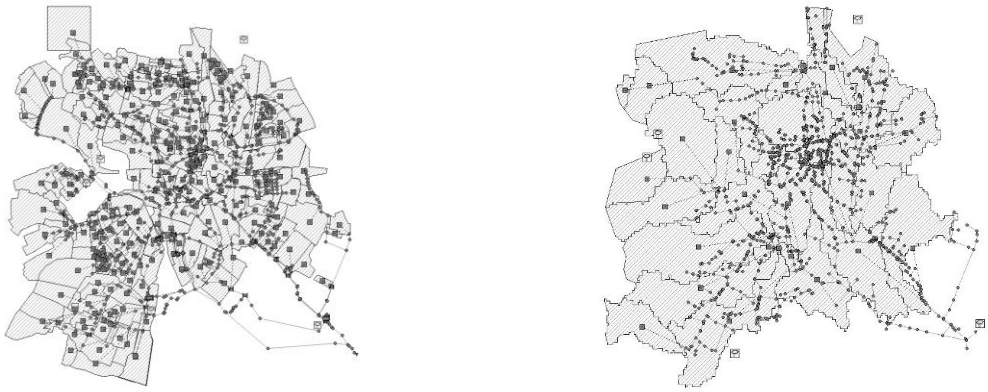

Figure 1. Original manually made model subcatchments delineation on the left compared to the automatically generated subcatchments based on topography on the right.

The results shown in Figure 2 look promising, particularly with regard to the reproduction of the dry weather flow and infiltration. However, the model tends to overestimate peak flows during rainfall events. This phenomenon is influenced by two parameters during model generation, namely the percentage of imperviousness of built-up areas and the position of CSO in the network.

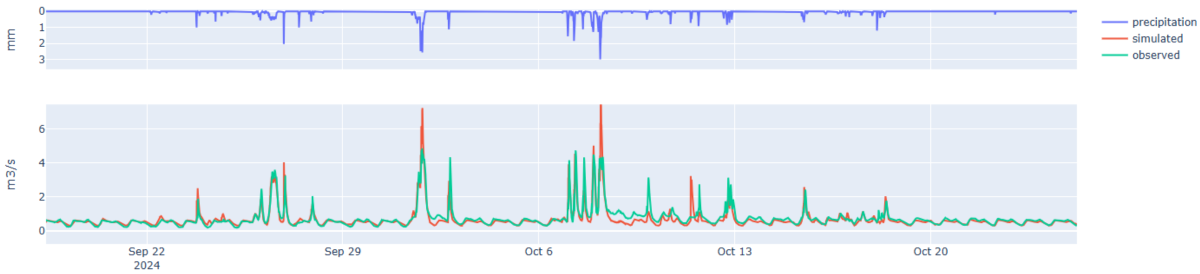

Figure 2. Comparison between observed and simulated WWTP inflow.

Conclusions and perspectives

A new method aiming at generating urban drainage models in SWMM at global earth scale using poor a priori knowledge has been developed and presented. It relies on large open-data sources for weather, hydrology, and hydraulics that were selected for their global earth coverage and (temporal/spatial) resolution fitting with flowrate modelling purposes. Results show model performances good enough to be used for several use cases such as long-term scenario simulations or yearly average urban drainage outlet flowrate prediction.

Several challenges emerged during the study, including parameter definition and generally poor GIS quality that requires robust algorithms to overcome data gaps. The authors are working on overcoming this limitation by providing a stochastic approach to cover certain parameter ranges for the poorly known one. In parallel, the authors are developing a second algorithm that generates models without GIS, using only open-source data. These results will be presented during the conference as they are currently being generated.

Acknowledgement

We thank the operators of participating wastewater treatment plants for allowing us to use the data.

References

Elga S., Bronders J. and B. Okke (2015) Hydrological modelling of urbanized catchments: A review and future directions. *J. of Hydro*., 529, 62-81. http://dx.doi.org/10.1016/j.jhydrol.2015.06.028

Möderl M., Butler D. and W. Rauch (2009) A stochastic approach for automatic generation of urban drainage systems. *Wat. Sci. Tech*., 59(6), 1137-1143. DOI:10.2166/wst.2009.097

Schilling J. and Tränckner J. (2022). Generate_SWMM_inp: An Open-Source QGIS Plugin to Import and Export Model Input Files for SWMM. *Water*, 14, 2262. https://doi.org/10.3390/w14142262

Reyes-Silva J.D., Novoa D., Helm B. and P. Krebs (2023) An Evaluation Framework for Urban Pluvial Flooding Based on Open-Access Data. *Water*, 15, 46. https://doi.org/10.3390/w15010046

Dobson, B., Jovanovic, T., Alonso-Álvarez, D., & Chegini, T. (2025). SWMManywhere: A workflow for generation and sensitivity analysis of synthetic urban drainage models, anywhere. Environmental Modelling & Software, 186, 106358.Rossman L.A. and M.A. Simon (2022) Storm water management model – User's manual version 5.2. National Risk Management Research Laboratory - U.S. ENVIRONMENTAL PROTECTION AGENCY

O'Callaghan, J.F. and Mark, D.M. (1984) The Extraction of Drainage Networks from Digital Elevation Data. Computer Vision, Graphics and Image Processing, 28, 328-344.http://dx.doi.org/10.1016/S0734-189X(84)80011-0

Tarjan, R.E. (1972). Depth-First Search and Linear Graph Algorithms. *SIAM J. Comput*., 1, 146-160.

Rossman, L. A., Dickinson, R. E., Schade, T., Chan, C. C., Burgess, E., Sullivan, D., & Lai, F. H. (2004). SWMM 5-the next generation of EPA's storm water management model. *Journal of Water Management Modeling*.