



HAL
open science

Parametrization of a wastewater hydraulic model under incomplete data constraint

Yassine Belghaddar, Carole Delenne, Nanée Chahinian, Abderrahmane Seriai, Ahlame Begdouri

► **To cite this version:**

Yassine Belghaddar, Carole Delenne, Nanée Chahinian, Abderrahmane Seriai, Ahlame Begdouri. Parametrization of a wastewater hydraulic model under incomplete data constraint. 2022. hal-03857301

HAL Id: hal-03857301

<https://hal.umontpellier.fr/hal-03857301>

Submitted on 17 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Parametrization of a wastewater hydraulic model under incomplete data constraint

Yassine Belghaddar^{1,2,3,4}, Carole Delenne^{1,4}, Nanée Chahinian¹, Abderrahmane Seriai³ and Ahlame Begdouri²

¹ HSM, Univ. Montpellier, CNRS, IRD, Montpellier, France.

² LSIA, Univ. Sidi Mohamed Ben Abdellah, Fez, Morocco.

³ Berger-Levrault, Pérols, France.

⁴ Inria Lemon, CRISAM – Inria Sophia Antipolis – Méditerranée, France.

yassine.belghaddar@berger-levrault.com

Abstract. Hydraulic simulation represents a powerful tool for studying wastewater networks. In order to achieve this target, hydraulic software require a set of parameters such as pipe slopes, roughness, diameters, etc. However, these pieces of information are rarely known for each and every pipe. Moreover, underground networks are frequently expanded, repaired and improved and these changes are not always reported in databases. The task of completing the required data represents the most time-consuming part of model implementation. In this context, we present algorithms that complete missing data required by hydraulic software. We automated this data insertion and transformation in SWMM© format to make it quicker and easier for the user. This automated solution was compared with manually estimated inputs. The simulation results show a coherent hydraulic behavior.

Keywords. Wastewater network, hydraulic simulation, missing data, parameter estimation, SWMM.

1. Introduction

Urban stormwater and wastewater networks play an environmental role in many ways. The former, draining rainwater, constitute a protection against floods whereas the latter are intended to route wastewater to treatment systems, when they exist, or directly into the receiving bodies. In the framework of flood and pollution risk management, the hydrodynamic modelling of such networks is essential. However, getting accurate and updated information on the underground wastewater networks is a cumbersome task, especially in cities undergoing urban expansion. Data related to wastewater networks can be divided into two categories: spatial data representing the geo-referenced position of the objects such as pipes or manholes, and attribute data associated to each object such as the diameter of a pipe.

The most common challenge when managing environmental systems is missing data [1-3]. For wastewater networks, incomplete databases have various consequences such as delays in repair and collateral damage on the networks thus reducing the quality of the services provided to the citizens and increasing expenses.

To help managers acquire precise information about the position of the buried objects, diverse solutions have been published. In [4], the authors propose a Bayesian mapping model to integrate information collected using Ground Penetrating Radar, surveyed manholes and available statutory records. A methodology has been developed in a previous project for an automatic mapping procedure in data scarce settings [5,6] using manhole covers detection in aerial images [7].

Even if the networks' maps obtained through these procedures are in good agreement with the actual networks in terms of topology, they cannot be used directly by hydraulic modelling software. Indeed, no or very little information is available on the main attributes of the network: pipe shapes and dimensions,

roughness, slopes, etc. To address the challenge of missing attributes related to wastewater networks, in [8], the authors use Missing Value Imputation techniques to estimate missing values of pipe diameters and the number of valves. Although the results were encouraging for some methods, this study was conducted on a small percentage of missing values (between 2% and 12%), which is not always the case in real life situations.

The present work aims at automatically completing the database associated to the mapped network, using the little data available and rules based on classical guidelines for the construction of such networks. This paper is organized as follows: Section 2 present the methodology and the materials used in this study. The results are described and discussed in Section 3. Section 4 concludes this paper.

2. Materials and methods

We assume that the map of a wastewater network is available in the form of a polyline shapefile which may be obtained either from the local stakeholder or from the previously developed mapping process. In this work, we use the wastewater network of Prades-le-Lez, a small town located in Southern France, produced in [5]. We already know that no or little data is available in the attribute table associated to the underground network. Moreover, other information may be available from national or open access geographical databases. In France, they correspond to commercial products developed by the French Geographical Institute IGN: Bd-Topo[®] for roads and buildings and RGE-alti 1m[®] for elevation.

The minimum information required to run a hydraulic software are: inlet positions, depth and associated input discharges, pipe geometries, slopes and roughness. Inlet positions are assumed to be known from the map. In a first approximation, pipe roughness can be assumed uniform given the most probable material used in the city. Pipe geometry can also be assumed circular with little impact on the results leaving one parameter to be estimated: their diameters. Finally, if the absolute depth of inlet is not mandatory, the corresponding pipes slope is a parameter of great importance, as classical hydraulic models cannot compute gravity fed flow on counter-slopes, unless pumping stations are installed.

We thus present the methodology used to automatically assign diameter and slope values to each pipe of the network and to assess input discharges.

2.1. Diameter estimation

Using minimal and maximal value bounds chosen by the user, the algorithm allocates the pipe diameters according to Strahler's order [9], thus ensuring the general increase of the diameters from upstream to downstream. When a minimum of 20% of the attributes are available, one may use a semi-supervised learning method to predict the missing values as described in [10].

2.2. Slope estimation

The developed algorithm first allocates to pipes nodes the ground elevation value minus the minimum burying depth of pipes, i.e. 0.8m in France. Starting from the outfall, it associates an upstream and downstream point, and thus a direction to each pipe, assuming gravity fed flow. However, even if ground elevation has proven to be well related to manhole depths, the precision of this simple estimation is not sufficient and requires modifications to ensure the conveyance of flow in the correct direction. The algorithm proposed here can be summed up as follows:

- Separate the network into "branches" where all the points can have one upstream and one downstream point at most;
- For each branch starting from the general outfall:
 - Check the coherence of the elevations and correct if needed to ensure a general decrease from upstream to downstream.
 - If several corrections are to be carried out on a same point, choose the one with minimum elevation.

Figure 1 describes the main steps of the slope estimation.

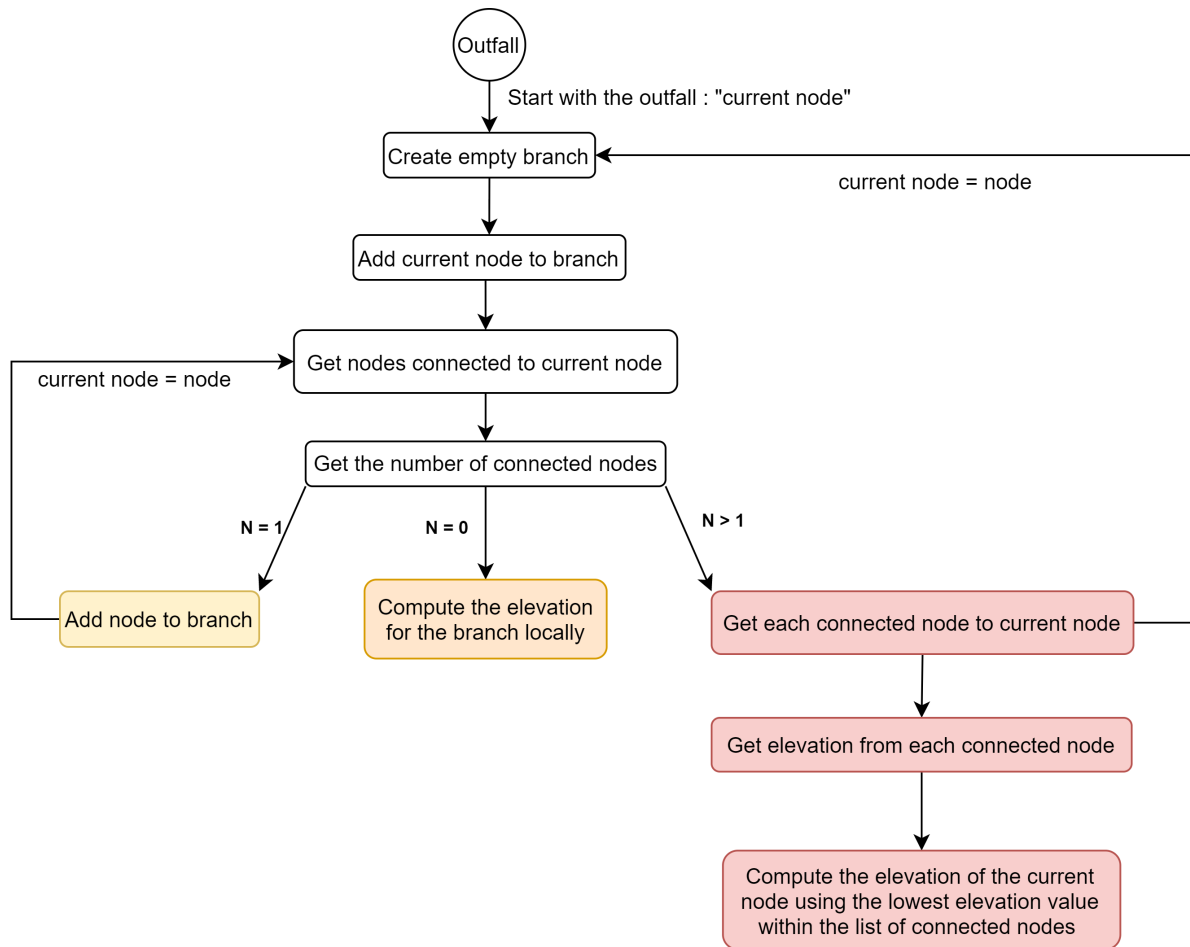


Figure 1: Slope estimation steps.

2.3. Input discharges

At each node, input discharge is estimated from the following pieces of information: the average number of inhabitants in the surrounding houses and their mean daily sewage discharge assessed from drinking water consumption. We assume that each construction is more likely to be connected to the closest pipe in terms of Euclidean distance. Figure 2 presents an illustration of the process based on the IGN building database (BD-Topo©). Each Polygon is linked to the closest node and an input discharge is estimated for each node.

SWMM© [11] is a Windows-based desktop program, used for planning, analysis, and design related to stormwater runoff, combined and sanitary sewers, and other drainage systems. We use SWMM© to run the hydraulic simulation. The proposed process should result in the automatic creation of all the attributes required by this software. This will ensure considerable gain in preparation time for the user. Tests were carried on the wastewater network of Prades-le-Lez. The database contains approximately 800 pipes and almost no associated characteristics, namely no elevation.

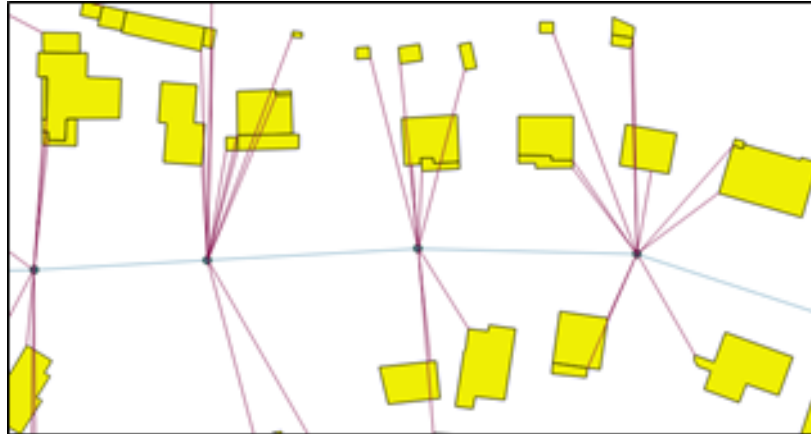


Figure 2. Linking each building to the closest network node.

3. Results and discussion

The procedure succeeded in completing the database and running hydraulic simulations without any errors. As no validation dataset is presently available for a better hydraulic validation, the attributes' values automatically estimated by the above-mentioned algorithms are compared with those manually estimated by a hydraulic engineer. For the elevation estimation the correlation is almost equal to 1 (Figure 3). Moreover, the simulation run on the entire dataset shows coherent hydraulic behavior (Figure 4).

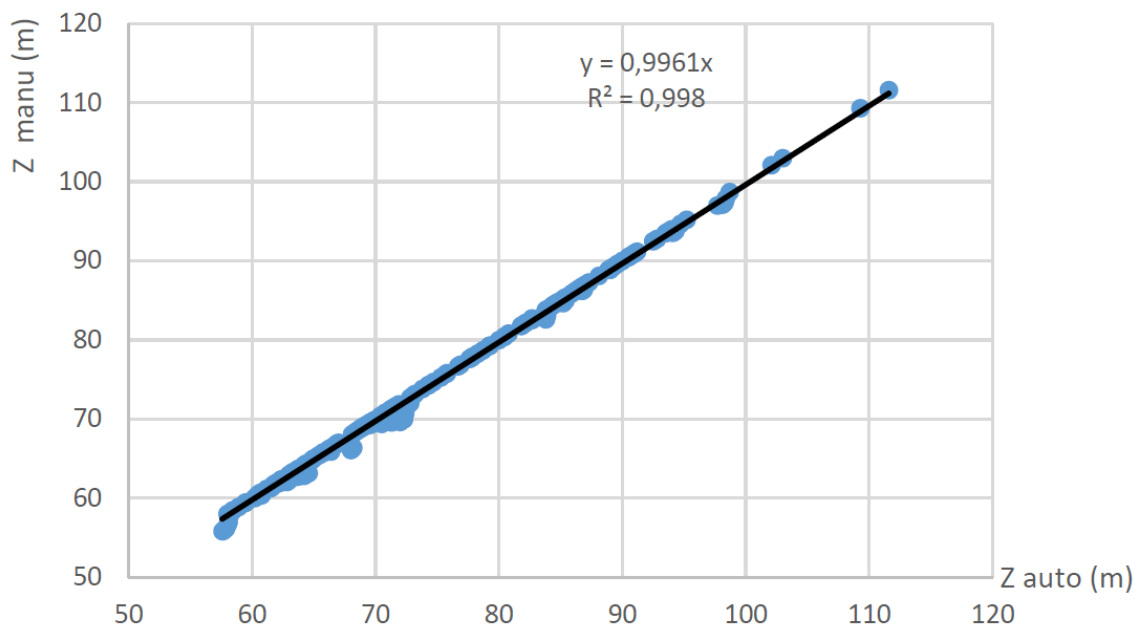


Figure 3. Automatic elevations (x-axis) and manual elevations (y-axis).

Figure 5 shows an example of the hydraulic simulation using SWMM©, where the various colors indicate water flow in each pipe and manhole.

It should be reminded that the network used in this study is the result of the map created in [5], where mapping errors may occur and that our propositions to compute the missing values are carried out under the worst conditions i.e. where no attributes values are provided in advance. By running the hydraulic simulation, we were able to detect some mapping errors during this process, precisely on the slope estimation step.

We noticed that assigning a slope value based on the elevation and the structure of the network may produce abnormal values such as a too big difference of altitude between two nodes of the same pipe. Such error impacts directly the entire network, thus the simulation directly. We put forward a tracing algorithm to identify the pipes where the errors are initiated. After comparing the reconstructed map with the map provided by the operator, we noticed that all the pipes where the errors are identified are not part

of the ground truth network and are due to detection or mapping errors. Figure 6 shows an example of such a situation, where the difference between the node identified by “657” and “738” is of almost 15 meters. The pipes from the reconstructed network (in blue) having the endpoints (“657”, “738”) and (“738”, “688”) do not exist in the actual network (in red).

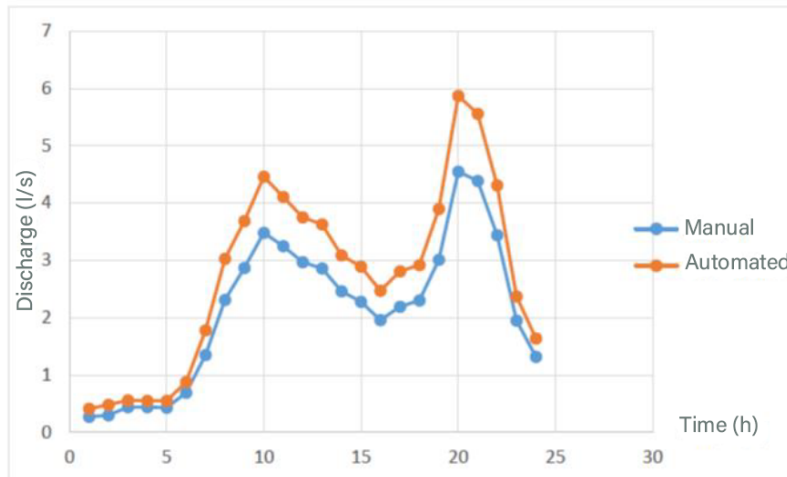


Figure 4. Comparison between outputs hydrographs from automatic and manual inputs.

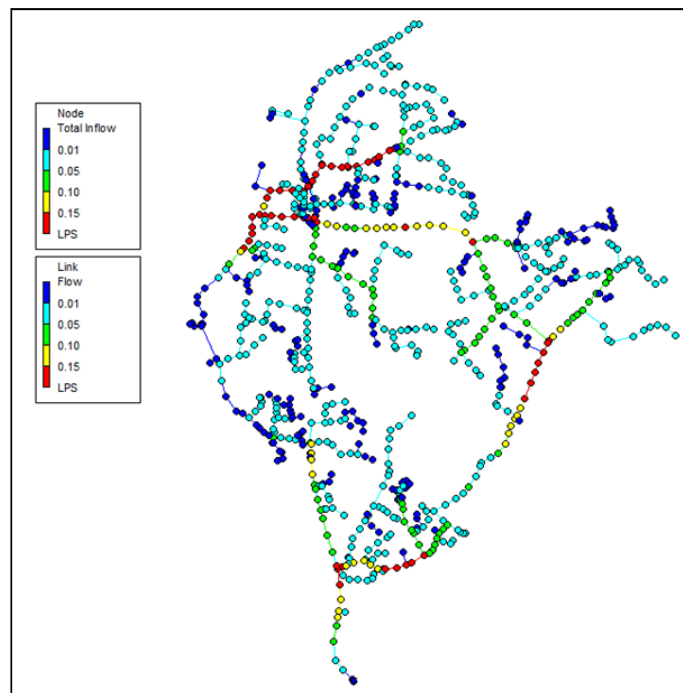


Figure 5: Simulation example using the SWMM© software with the proposed estimation process.

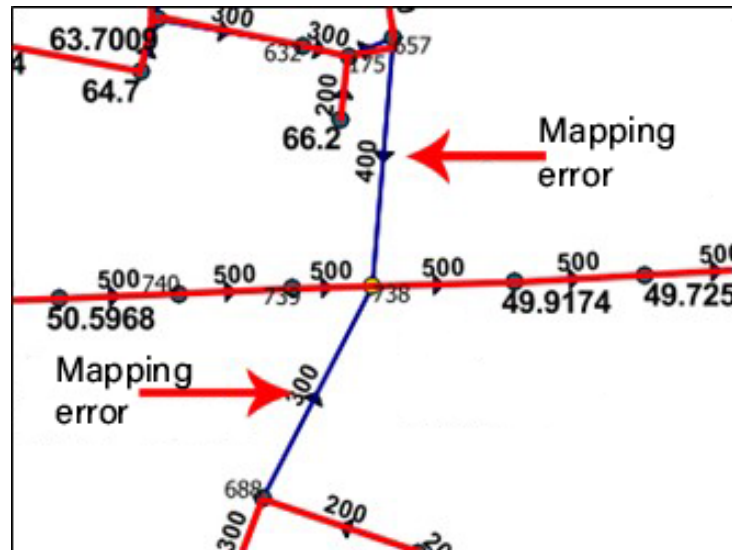


Figure 6: Identification of mapping errors.

4. Conclusions and perspectives

In this study we proposed a process to estimate the missing data needed to run hydraulic simulations using a different approach for each required parameter. We based our algorithms on information that are often available: the ground elevation, the average number of inhabitants per house and the mean daily discharge. In addition to being quicker and easier to use, the results accomplished by the automated solution indicate a coherent hydraulic behavior when comparing to manually estimated inputs.

The estimated data are not accurate and do not systematically reflect the reality on the ground. Consequently, the simulation results are inevitably uncertain and imprecise, and they can lead to wrong conclusion despite running without errors. In addition, estimation approaches are used when official recorded data are unavailable, and they are not intended to substitute for good quality field data. Thus, currently we are working on improving our process with the aim to account for the accuracy of the data that the user may define in the input map. Indeed, when the values of several pipe characteristics are known with certainty, they should be maintained by the algorithm and propagated to the surrounding pipes. For example, if a pipe's diameter is known, the algorithm should be able to modify other diameters to ensure the general increase of pipe conveyance from upstream to downstream. Furthermore, since information about the networks can be collected from different sources, using various techniques such as Radars and images, a data fusion framework, which includes all the available data, could help the managers surpass the problem of missing data and thus improve the accuracy of the simulations.

Acknowledgments

This work was carried out within the framework of the CIFRE-France/Morocco Program and the project Cart'Eaux funded by the European Regional Development Fund (ERDF) and the Languedoc Roussillon Region.

We thank Mustapha Derras for the general supervision of the research group in Berger-Levrault and administrative support.

References

- [1] Kofinas D.T., Spyropoulou, A., Lapidou, C.S. A methodology for synthetic household water consumption data generation. *Environ. Model. Softw.* 2018, 100, 48–66.
- [2] Schneider T. Analysis of incomplete climate data: Estimation of Mean Values and covariance matrices and imputation of Missing values. *J. Clim.* 2001, 14, 853–871.
- [3] Junninen H.; Niska, H.; Tuppurainen, K.; Ruuskanen, J.; Kolehmainen, M. Methods for imputation of missing values in air quality data sets. *Atmos. Environ.* 2004, 38, 2895–2907.
- [4] Bilal M., Khan W., Muggleton J., Rustighi E., Jenks H., Pennock S. R., Atkins P.R & Cohn A. (2018).

Inferring the most probable maps of underground utilities using Bayesian mapping model. *Journal of Applied Geophysics*, 150, 52-66.

- [5] Chahinian N., Delenne C., Commandré B., Derras M., Deruelle L., Bailly J.-S., (2019) Automatic mapping of urban wastewater networks based on manhole cover locations, *Computers, Environment and Urban Systems* 78 (2019) 101370.
- [6] Commandre B., Chahinian N., Bailly J.-S., Chaumont M., Subsol G., Rodriguez F., Derras M., Deruelle L. and Delenne C. (2017). Automatic reconstruction of urban wastewater and stormwater networks based on uncertain manhole cover locations. In ICUD, 14th IWA/IAHR International Conference on Urban Drainage, pages 2345–2352, Prague, Czech Republic, 10-15 September.
- [7] Pasquet J., Desert T., Bartoli O., Chaumont M., Delenne C., Subsol G., Derras M. et Chahinian N. (2016). Detection of manhole covers in high-resolution aerial images of urban areas by combining two methods. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 9(5):1802 – 1807.
- [8] Kabir G., Tesfamariam S., Hemsing J., & Sadiq R. (2020). Handling incomplete and missing data in water network database using imputation methods. *Sustainable and Resilient Infrastructure*, 5(6), 365-377.
- [9] Strahler, A. (1957). *Quantitative Analysis of Watershed Geomorphology*. American Geophysical Union, 38, 913- 920.
- [10] Belghaddar, Y., Chahinian, N., Seriai, A., Begdouri, A., Abdou, R., & Delenne, C. (2021). Graph Convolutional Networks: Application to Database Completion of Wastewater Networks. *Water*, 13(12), 1681.
- [11] Environmental Protection Agency, Storm Water Management Model (EPA SWMM), <https://www.epa.gov/>.