



HAL
open science

Comparaison des performances d'un modèle de réseau de neurone et d'un modèle distribué: application au débit déversé par un déversoir d'orage.

Frédéric Gogien, Magali Dechesne, Rémi Martinerie, Gislain Lipeme Kouyi

► To cite this version:

Frédéric Gogien, Magali Dechesne, Rémi Martinerie, Gislain Lipeme Kouyi. Comparaison des performances d'un modèle de réseau de neurone et d'un modèle distribué: application au débit déversé par un déversoir d'orage.. Novatech 2023, Jul 2023, Lyon, France. hal-04176901

HAL Id: hal-04176901

<https://hal.science/hal-04176901v1>

Submitted on 3 Aug 2023

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.

Comparaison des performances d'un modèle de réseau de neurone et d'un modèle distribué : application au débit déversé par un déversoir d'orage.

Performance comparison between a neural network model and a distributed model: application to a combined sewer overflow case study.

F. Gogien^{1,2}, M. Dechesne³, R. Martinerie⁴, G. Lipeme Kouyi¹

¹ Univ Lyon, INSA Lyon, DEEP, EA7429, 69621 Villeurbanne, France

² Veolia EAU- Région Centre-Est, 2-4 avenue des Canuts, 69120 Vaulx-en-Velin, France - frederic.gogien@veolia.com

³ Veolia Recherche & Innovation, 765 rue Henri Becquerel, 34965 Montpellier, France

⁴ Valence Romans Agglo, Direction de l'Assainissement, des Eaux Pluviales et des Rivières, 1 place Jacques Brel, 26000 Valence, France

RÉSUMÉ

La construction d'un modèle hydraulique est une opération complexe et souvent coûteuse pour les maîtres d'ouvrage de réseaux d'assainissement. Outre les modèles conceptuels distribués, les modèles fondés sur l'intelligence artificielle et en particulier les réseaux de neurones peuvent constituer une alternative technique et économique intéressante en fonction de l'objectif poursuivi. Cette communication compare les performances de ces deux types de modèles dans le cadre de simulations pluie/débit réalisées sur un déversoir d'orage de la ville de Valence. Les performances du réseau de neurones semblent un peu en retrait par rapport à celle du modèle distribué. Des perspectives sont évoquées en fin de communication pour améliorer les performances de ce type de modèle.

ABSTRACT

Building a hydraulic model is a complex and often costly operation for project owners of sewer networks. In addition to distributed conceptual models, models based on artificial intelligence and in particular neural networks can be considered as an interesting technical and economic alternative depending on the objective pursued. This communication compares the performance of these two types of models in the context of rain/flow simulations carried out on a combined sewer overflow in the city of Valence. The performance of the neural network seems a little off compared to that of the distributed model. Prospects are mentioned at the end of the communication to improve the performance of this type of model.

MOTS CLÉS

Modélisation pluie débit, déversoir d'orage, réseau de neurones, modèle distribué

1 INTRODUCTION

La modélisation est un outil auquel les maîtres d'ouvrage de réseaux d'assainissement ont régulièrement recours dans le cadre d'études diagnostic ou de schémas directeurs. Le développement d'un modèle hydrologique et hydraulique de réseau est généralement mis en œuvre à l'aide de modèles conceptuels distribués. La construction d'un tel modèle est un processus nécessitant une bonne connaissance de la structure du réseau et des bassins versants urbains qui y sont rattachés. Il est notamment indispensable de disposer de l'ensemble des informations d'altimétrie du réseau. A l'opposé du modèle distribué, les modèles pluie-débit basés sur l'intelligence artificielle ne nécessitent pas de disposer d'une connaissance fine de la structure du système étudié. Ce type de modèle, plus facile à construire, peut dans certains cas présenter une alternative à la construction d'un modèle distribué complet. Cette communication s'attache à comparer les performances d'un modèle distribué avec celles d'un modèle de réseau de neurones de type NARX pour la modélisation pluie-débit déversé au niveau d'un déversoir d'orage du système d'assainissement de Valence (Drôme).

2 LE SITE D'ETUDE

Le site étudié est le réseau d'assainissement de la ville de Valence. Il s'agit d'un réseau unitaire de 700 km de longueur dont l'exutoire principal est une station d'épuration de capacité 172 000 EH. Le réseau est équipé de déversoirs d'orage dont le principal est le déversoir Gambetta qui contribue à environ 45% du volume annuel déversé par le système d'assainissement, avec en moyenne 450 000 m³ déversés par an. C'est ce déversoir et la relation pluie-débit à son niveau qui sont étudiés dans cette communication.

3 CONSTRUCTION DES MODELES, CALAGE ET VALIDATION

Le jeu de données utilisé pour mettre au point les deux types de modèles est le même, à savoir l'année 2019 pour le calage des deux modèles et l'année 2018 pour la validation. Avec une pluviométrie moyenne de 870 mm/an à Valence, on pourra considérer 2018 comme une année humide (914 mm) et 2019 comme une année sèche (729 mm) Le choix de l'année 2019 pour le calage et de l'année 2018 pour la validation a été guidé par le nombre de déversements observés sur le site en retenant l'année ayant totalisé le nombre de déversements le plus élevé pour l'étape de calage (2018 : 45 déversements ; 2019 : 56 déversements).

3.1 Le modèle distribué

Le modèle distribué a été construit à l'aide du logiciel InfoWorks-CS™. Le réseau décrit couvre un linéaire de 50 km. La surface totale est discrétisée en 154 bassins versants élémentaires. Le modèle hydrologique retenu pour la surface imperméable est de type réservoir linéaire et celui retenu pour les surfaces perméables de type refus à l'infiltration. Le calage en temps de pluie a été réalisé à partir de données de quatre points de mesure. Les critères retenus pour évaluer le calage et la validation sont pour chaque épisode le volume, le débit de pointe et le critère de Nash-Sutcliffe. Les performances atteintes pour chacun des critères retenus ont permis de conclure à un calage et une validation du modèle (Bertrand-Krajewski et al., 2000 ; Aguilar et al., 2016).

3.2 Le modèle réseau de neurones

La grande majorité des études impliquant l'utilisation de réseaux de neurones en hydrologie s'attache à des horizons de prédiction courts. L'objectif ici est de construire un modèle dont l'usage pourrait être équivalent à celui d'un modèle distribué, c'est-à-dire capable de réaliser des simulations longues durées sans recalage intermédiaire à l'aide de mesures, et ce avec une qualité de prédiction équivalente à celle d'un modèle distribué. Une étude bibliographique a conduit à sélectionner un type de réseau de neurones bien adapté à la simulation de séries temporelles, le réseau de type NARX (Figure 1) pour « Nonlinear AutoRegressive with eXogenous input » (Zhang et al., 2018 ; El-Ghazouli et al., 2019). Les ajustements à réaliser pour obtenir les meilleures performances concernent le nombre de neurones ainsi que la profondeur de l'historique de données à prendre en compte en entrée pour simuler la sortie à chaque pas de temps.

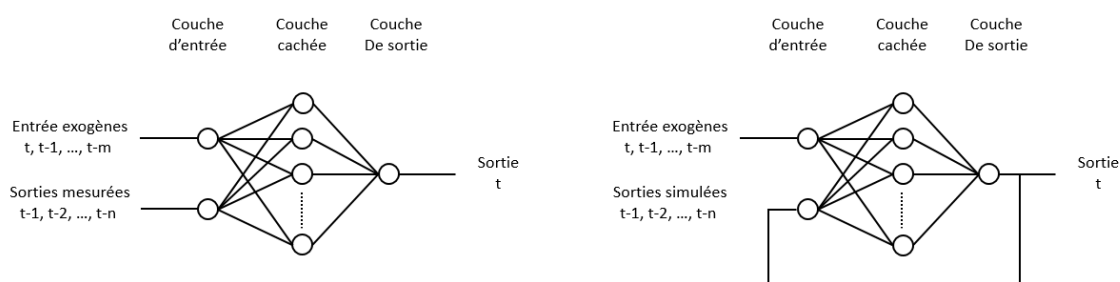


Figure 1 : Architecture d'un réseau de neurones de type NARX à une couche cachée, avec à gauche le fonctionnement en boucle ouverte lors de la phase d'apprentissage, et à droite le fonctionnement en boucle fermée en validation et simulation. Les indices $t, t-1, \dots, t-n$ correspondent aux numéros des pas de temps pour lesquels les valeurs sont prises en compte en entrée du réseau.

Concernant le réseau de neurones, les étapes d'apprentissage, de test et de validation ont été réalisées sur les données de l'année 2019. Les données 2018 ont quant à elles été utilisées en prédiction, qui sera considérée ici comme une validation de façon à disposer d'une configuration équivalente à celle du modèle distribué et ainsi rendre la comparaison possible.

4 RESULTATS ET DISCUSSION

4.1 Comparaison des performances des deux modèles

Les résultats mettent en évidence sur la période de validation (2018) de très bonnes performances du modèle distribué et des performances très acceptables quoiqu'un peu en retrait pour le NARX (voir Figure 2).

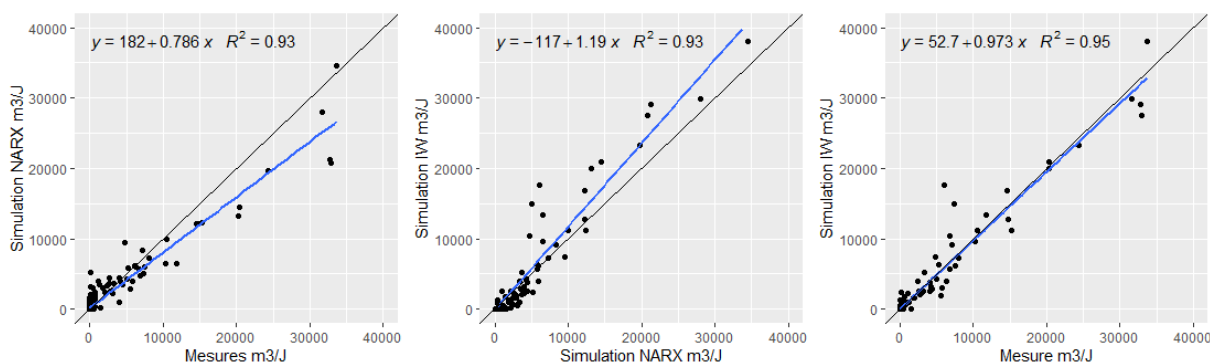


Figure 2 : Comparaison des volumes journalier déversés pour l'ensemble des épisodes de déversement sur l'année de validation (2018), avec à gauche le NARX confronté aux mesures, à droite le modèle distribué confronté aux mesures et au milieu le modèle distribué confronté au NARX.

Le principal inconvénient du réseau de neurones est sa tendance à sous-estimer dans le cas présent, les volumes déversés pour les épisodes de fort déversement (supérieurs à 10 000 m³). La corrélation avec les mesures de l'année 2018 reste cependant correcte.

4.2 Discussion

La question se pose de la quantité de données utilisée pour mener cette étude, notamment pour assurer un bon apprentissage du réseau de neurones. Un apprentissage sur un jeu de données plus important devrait pouvoir améliorer les performances.

Si les performances du réseau de neurones semblent un peu en dessous de celles du modèle distribué, elles

restent néanmoins acceptables. La question se pose donc dans le cadre d'un usage opérationnel d'utiliser ce type de modèle simple lorsque l'objectif poursuivi est compatible, et en gardant à l'esprit que ce type d'approche ne permet pas de simuler des scénarios d'aménagement ou tout autre fonctionnement impliquant un changement structurel du système étudié. En revanche, elle pourrait être utilisée pour simuler les déversements sur des déversoirs non instrumentés ou en cas de défaillance sur un système de mesure. L'approche pourrait aussi être utilisée pour quantifier l'impact d'un changement du régime pluviométrique sur le site étudié. Pour ce dernier cas d'usage, si les types de pluie considérés sont très différents de ceux utilisés lors de la phase d'apprentissage, il conviendra de considérer les résultats avec précaution, mais cette remarque est valable de la même manière si on considère un modèle distribué.

5 CONCLUSION

Deux approches distinctes de modélisation ont été menées sur un même système d'assainissement avec un même jeu de données : une approche « classique » basée sur un modèle conceptuel distribué et une approche basée sur un réseau de neurones. Si la précision du réseau de neurones semble un peu en dessous de celle du modèle distribué, l'utilisation d'un jeu de données plus conséquent pourrait améliorer les performances. Ce type d'approche relativement simple peut constituer une alternative techniquement et économiquement viable pour certains usages tels que pallier une absence ou un défaut de capteur, ou encore déterminer l'impact d'une variation d'un régime pluviométrique sur un site.

BIBLIOGRAPHIE

- Aguilar, M. F., McDonald, W. M., & Dymond, R. L. (2016). *Benchmarking laboratory observation uncertainty for in-pipe storm sewer discharge measurements*. *Journal of Hydrology*, 534, 73-86.
- Bertrand-Krajewski, J.-L., Laplace, D., Joannis, C., Chebbo, G. (2000). *Mesures en hydrologie urbaine et assainissement*. Tec&Doc, Lavoisier. ISBN : 2-7430-0380-4
- El Ghazouli, K.; El Khatabi, J.; Shahrour, I. & Soulhi I. (2019). *Comparison of M5 Model Tree and Nonlinear Autoregressive with exogenous inputs (NARX) Neural Network for urban stormwater discharge modelling*. MATEC Web of Conferences. 295. 02002. 10.1051/mateconf/201929502002.
- ZHANG, Duo, LINDHOLM, Geir, et RATNAWEERA, Harsha. *Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring*. *Journal of Hydrology*, 2018, vol. 556, p. 409-418.