

**POLYTECHNIQUE
MONTRÉAL**

UNIVERSITÉ
D'INGÉNIERIE



Prédiction des débits dans le réseau d'égouts en fonction des prévisions météorologiques et des précipitations enregistrées

Rédigé par : Chaimaa Lotfi

Supervisé par : Professeur Jonathan Jalbert, Polytechnique Montréal

Juillet-Août 2020

Sommaire

1	Introduction	3
1.1	Problème d'intérêt	3
1.2	But de projet	3
2	Les données.	4
2.1	Description de données.	4
2.2	Analyse exploratoire de données.	4
3	Modèles et Résultats.	7
3.1	Motivation :	7
3.2	Modèle utilisé :	7
3.3	Résultats	8
3.3.1	Résultats pour un seul débit	8
3.3.2	Résultats pour tous les débits, calcul d'efficacité et d'erreur relative :	12
4	Conclusion	14
5	Bibliographie	14
6	Remerciements	14

1 Introduction

1.1 Problème d'intérêt

Ce rapport résume une étude effectuée pour déterminer si on peut prédire des débits dans le réseau d'égouts en fonction des prévisions météorologiques et des précipitations enregistrées.

À Montréal, avant de rejeter l'eau traitée dans le milieu récepteur, il existe des réseaux destinés à acheminer l'eau sanitaire et pluviale vers la station d'épuration. Cependant, lorsque l'intensité des précipitations dépasse la capacité de transport du réseau ou la capacité de traitement de la station d'épuration, le mélange eau de pluie et eau sanitaire sont directement rejetés dans le milieu récepteur sans passer par la station d'épuration. C'est l'une des principales causes des maladies d'origine hydrique.

1.2 But de projet

Établir le lien entre les débits aux régulateurs et les précipitations enregistrées aux 51 stations pluviométriques réparties sur l'île de Montréal et les prévisions météorologiques.

2 Les données.

2.1 Description de données.

- Les précipitations enregistrées pour le mois Juillet : Ce sont les précipitations enregistrées dans les 51 stations de Montréal pour le mois juillet. Les mesures ont été récupérées toutes les 5 minutes pendant le mois juillet 2019. Donc, on obtient 8930 valeurs mesurées.
- Les débits aux régulateurs pour le mois Juillet : Ce sont les débits aux régulateurs de 39 sites de Montréal pour le mois juillet. Les mesures ont été prises toutes les 5 minutes pendant le mois juillet 2019. Donc, on obtient 8930 valeurs mesurées.

2.2 Analyse exploratoire de données.

Dans un premier temps, j'ai tracé les précipitations et les débits en fonction du temps sous forme d'un histogramme.

Pour faciliter le travail, j'ai sommé les précipitations des 51 stations pour le 5 et le 6 juillet, ainsi que les débits et les précipitations pour chaque heure (et non pas toutes les 5 minutes) et les ai tracé avec un pas de 3 heures.

- Pour les précipitations :

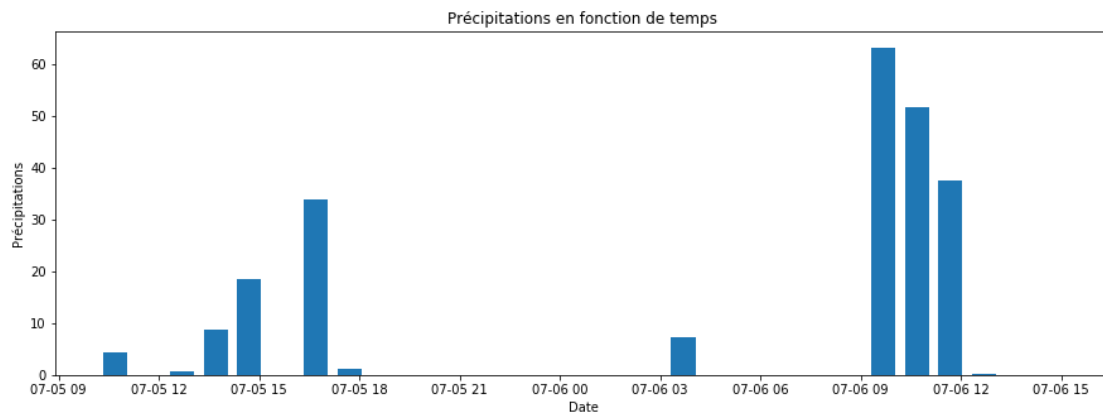


FIGURE 1 – Les précipitations en fonction de temps.

- Pour les débits aux sites :

J'ai tracé seulement les deux premiers débits aux sites BCFI0010 et BSFI0010 en fonction de temps sous forme d'un histogramme de pas de 3 heures.

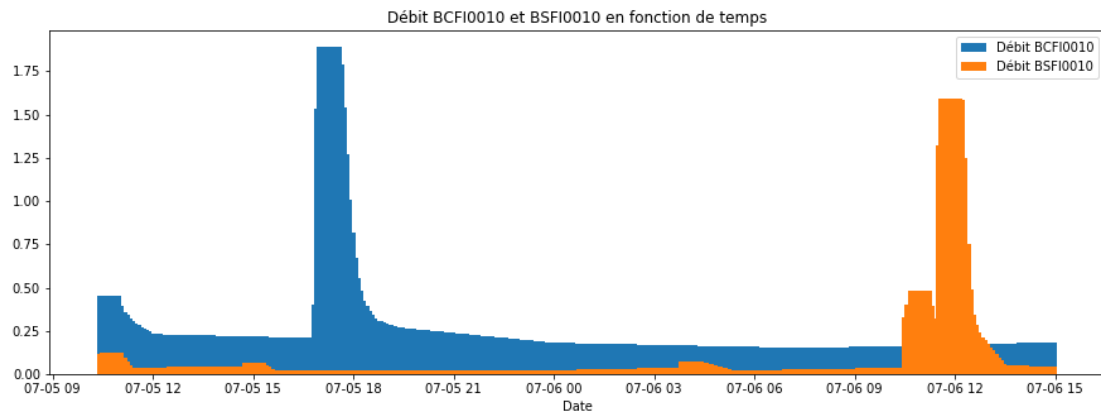


FIGURE 2 – Débits BCFI0010 et BSFI0010 en fonction de temps.

Finalement, j'ai tracé les débits aux sites en fonction de précipitations.

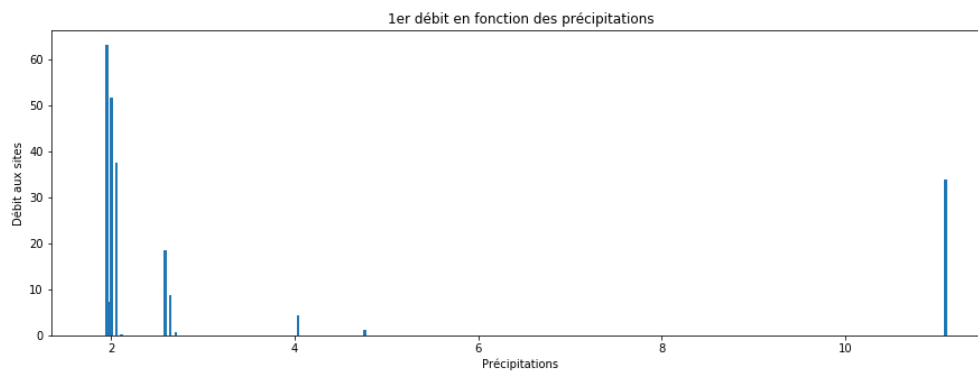


FIGURE 3 – Débits aux sites en fonction de précipitation

Conclusion : j'ai bien obtenu qu'entre 9h et 12h nous avons un taux de précipitation élevé, ce qui correspond à l'événement qu'on veut observer (entre 10h20 et 15h15). L'événement dans la ligne 45 montré dans la figure 6.

pluvieux	Date début	Date fin	(hh:mm)	d'hétérogénéité	totale (mm)	(mm/h)	5 mins
44	2019-07-11 07:40:00	2019-07-12 00:43:55	17:5	0.00	57.0	150.00	2 ans
45	2019-07-06 10:19:57	2019-07-06 14:15:15	3:55	0.00	10.3	69.00	3 mois
46	2019-06-30 08:30:00	2019-06-30 17:54:56	9:24	0.00	16.8	78.00	6 mois

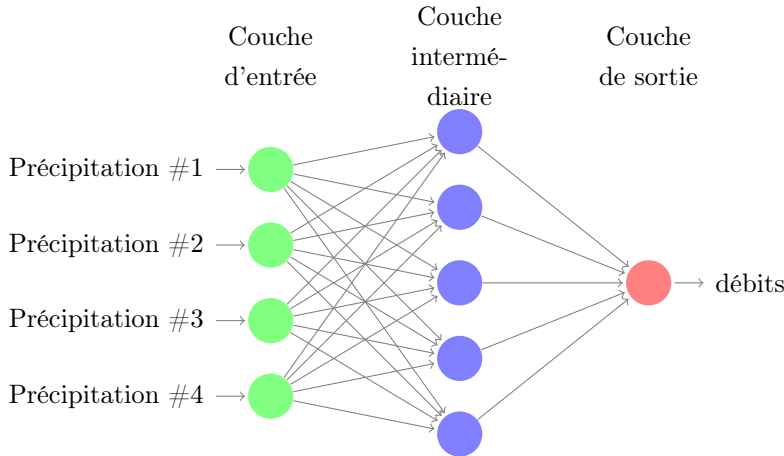
FIGURE 4 – Événement à récupérer dans la ligne 45.

3 Modèles et Résultats.

3.1 Motivation :

Les réseaux de neurones artificiels sont présentés comme une alternative à la modélisation et la prédiction traditionnelles. Depuis la fin des années 80, l'utilisation des réseaux de neurones en hydrologie a augmenté. Ils sont un outil très utile car, comme la plupart des processus hydrologiques, ce sont des opérateurs mathématiques qui peuvent être non-linéaires.

La majorité des études effectuées pour la prévision du débit en utilisant les réseaux de neurones partent des données de pluie. Les résultats obtenus dans la transformation pluie-débit sont réellement satisfaisants. Ce que nous faisons comprendre la prédiction des débits aux régulateurs à la sortie en fonction de débits et les précipitations observées.



3.2 Modèle utilisé :

J'ai utilisé les deux frameworks sur python Keras et tensorflow, car ils sont faciles et rapides à apprendre et à utiliser.

D'une part, pour prédire des débits aux régulateurs, j'ai utilisé un réseau de neurones avec comme entrée les précipitations dans les 51 stations avec 7 couches intermédiaires et pour sortie les 39 débits à prédire. Donc, je n'avais pas besoin de trouver la relation entre les débits. De plus, j'ai partitionné les données entre observations pour apprentissage(90%) et test (10%). D'autres part, les composantes utilisées sont :

- Fonction de perte (loss function) : Erreur quadratique moyenne (mse), puisqu'on a un problème de régression.
- Algorithme d'optimisation : Adam et c'est le plus utilisé.
- Taux d'apprentissage (learning rate, lr) : 10^{-3} . Pour déterminer le taux d'erreur, j'ai choisi deux méthodes :
- L'erreur relative :

$$\frac{1}{m} \sum_{n=1}^m \frac{| \text{valeur calculée} - \text{valeur théorique} |}{\text{valeur théorique}},$$

où m est la longueur de la liste de test.

- Coefficient d'efficacité de Nash-Sutcliffe :

$$\text{NSE} = 1 - \frac{\sum_{n=1}^m |Q_m - Q_o|}{\sum_{n=1}^m |Q_{mean} - Q_o|}.$$

3.3 Résultats

3.3.1 Résultats pour un seul débit

Dans un premier temps, j'ai utilisé ce réseau de neurones seulement pour prédire le débit BHFI0010B.

Figure 5 présente la fonction de perte pour l'ensemble d'entraînement et l'ensemble de test en fonction du nombre d'itérations pour le débit BHFI0010B.

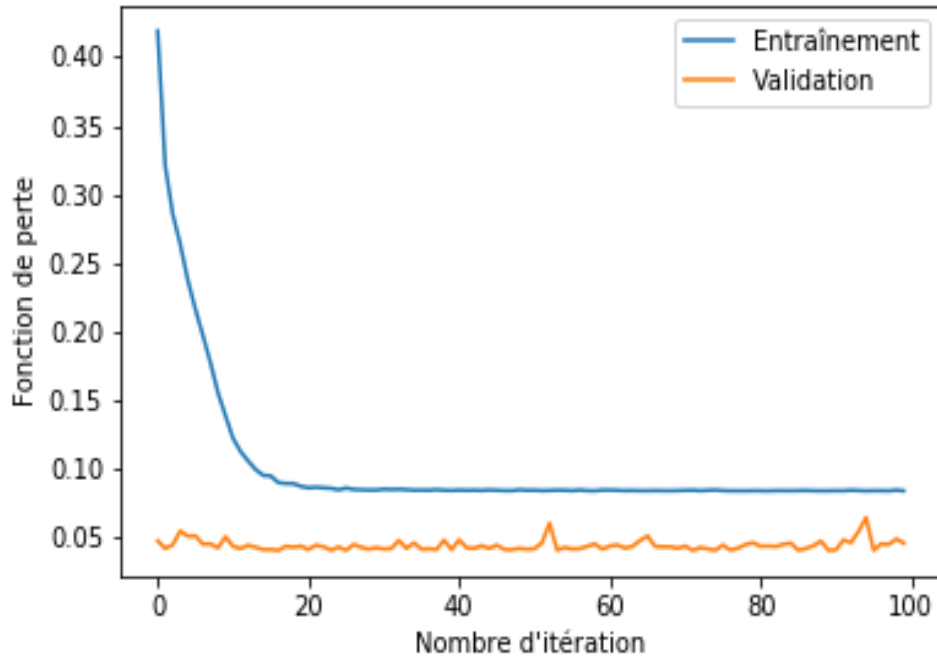


FIGURE 5 – Erreur d'entraînement et de test en fonction de Nombre d'itérations pour le débit BHFI0010B.

Tableau 3.3.1 montre une comparaison entre la série des valeurs observées pour le débit BHFI0010B et la série de valeurs prédites pour le débit BHFI0010B.

Débit BHFI0010B	Prédictions
0,38270545	0.3800919
0,38258636	0.3800919
0,3824673	0.3810919
0,38234821	0.3800919
0,382229149	0.3800919
0,382110089	0.3805919
0,381990999	0.3800919
0,381871939	0.3800919
0,381752878	0.3800919
0,381633788	0.3800919
0,381514728	0.3800919
0,381395668	0.3800919
0,381276578	0.3800919
0,379712999	0.3800919
0,377734035	0.3800919
0,375755072	0.3800919
0,373776108	0.3800919
0,371797144	0.3800919
0,369818181	0.3800919
0,367839217	0.3800919
0,365860283	0.3800919
0,36388132	0.3800919
0,361902356	0.3800919
0,359923393	0.3800919
0,357944429	0.3800919
0,355965465	0.3800919
0,353986502	0.3800919
0,352007538	0.3800919
0,350028574	0.3800919
0,348049611	0.3800919
0,346070647	0.3800919
0,344091684	0.3800919
0,34211272	0.3800919
0,340133786	0.3800919
0,338154823	0.3800919

Débit BHFI0010B	Prédiction de ce débit
0,336175859	0.3800919
0,334196895	0.3800919
0,332217932	0.3800919
0,330238968	0.3800919
0,328260005	0.38016197
0,326281041	0.3800919
0,324302077	0.3800919
0,322323114	0.3800919
0,32034415	0.3800919
0,318365186	0.3800919
0,316386223	0.3800919
0,314407259	0.3800919
0,312428325	0.3800919
0,310449362	0.3800919
0,308470398	0.3800919
0,306491435	0.3800919
0,304512471	0.3800919
0,302533507	0.3800919
0,300554544	0.3800919
0,29857558	0.3800919
0,296596617	0.3800919
0,294617653	0.3800919
0,292638689	0.3800919
0,290659726	0.3800919
0,288680762	0.3800919
0,250212848	0.3800919
0,250914514	0.3800919
0,25161621	0.3800919
0,252317876	0.3800919
0,253019542	0.3800919
0,253721237	0.3800919
0,254422903	0.3800919
0,255124599	0.3800919

TABLE 1: une comparaison entre la série réelle et la série prédite pour le débit BHFI0010B.

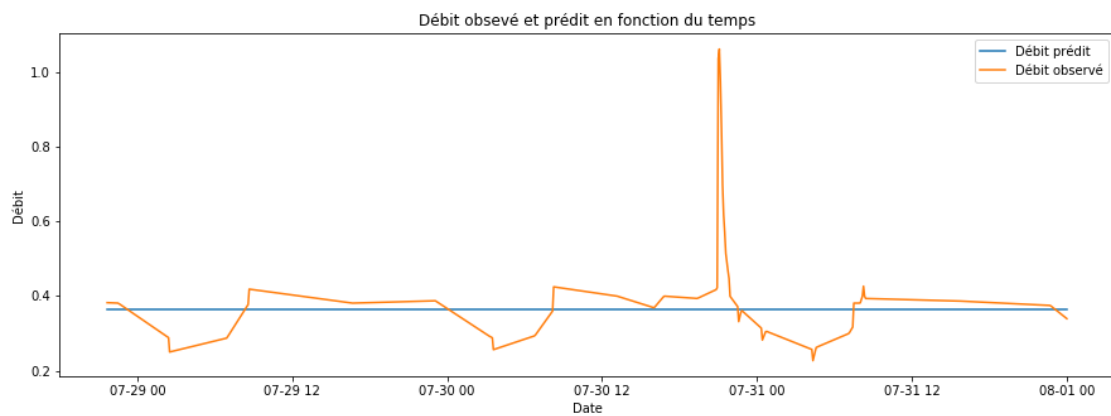


FIGURE 6 – Le débit BHFI0010B observé et prédit en fonction du temps.

Conclusion : On remarque bien dans la comparaison dans le tableau 3.3.1, que le débit prédit est proche au débit observé. De plus, la figure 5, montre que le débit observé est proche de celui prédit.

3.3.2 Résultats pour tous les débits, calcul d'efficacité et d'erreur relative :

Dans l'étude suivante, j'ai utilisé le réseau de neurones choisi mais qui a pour sorties tous les 39 débits à prédire.

Pour évaluer l'efficacité de ce réseau de neurones, j'ai utilisé les deux critères définis précédemment, l'un calcule l'erreur relative et l'autre calcule le coefficient d'efficacité de Nash-Sutcliffe.

Tableau 3.3.2 montre la performance de ces deux méthodes pour les 39 débits aux sites.

Débits	Métrique 1 .	Métrique de Nash-Sutcliffe
BCFI0010	10.0982738	-0.67176238
BSFI0010	26.19356808	0.26650279
ANFI0010	46.12171844	0.19706359
APFI0010	9.60710792	0.34992166
BAFI0010	9.74805827	0.1238022
AFFI0010	110.78539051	0.29329501
ACFI0010	15.55332694	0.01580603
BLFI0010	11.62553201	0.40657086
BNFI0010	42.95768456	0.17581864
AGFI0010	32.25782896	0.36887674
BMFI0010	12.19556142	0.4008361
BJFI0010	12.40049762	0.12091534
BKFI0010	21.75338187	0.26874943
BOFI0010	22.91510007	0.37766768
AIFI0010	17.07001503	0.3469332
BIFI0010A	17.53474521	0.01937515
BIFI0010B	23.15438299	-0.10704477
ALFI0010	30.251724	0.28140009
AKFI0010	22.28824753	0.13689202
BQFI0010	22.40831938	0.34118189
AHFI0010	53.33571095	0.21404942
AMFI0010	20.85670191	0.28203708
BGFI0010	12.34231628	-0.04249652
ADFI0010	89.02577287	0.01474383
ABFI0010	18.60234241	0.36642983
BHFI0010A	9.59344852	0.02900637
BHFI0010B	12.76927412	0.05992261
AAFI0010	19.02625276	0.40867358
BDFI0010	3.74054447	-0.01333693
BPFI0010	166.98244995	0.07862896
BFFI0010	47.14105012	-0.29782724
AOFI0010	27.48194569	0.19772419
AJFI0010	33.06219629	0.14545399
ARFI0010	71.77131577	-1.13409083
BRFI0010	40.62216818	0.34421999
BBFI0010	25.12007986	0.00612329
AEFI0010	11.49912722	0.2856344
BEFI0010A	18.96470496	-0.00288484
BEFI0010B	12.56477098	0.09317228

TABLE 2: Calcul d'efficacité et d'erreur relative pour les 39 débits.

Figure 7 présente la fonction de perte pour l'ensemble d'entraînement et l'ensemble de test en fonction du nombre d'itérations pour les 39 débits.

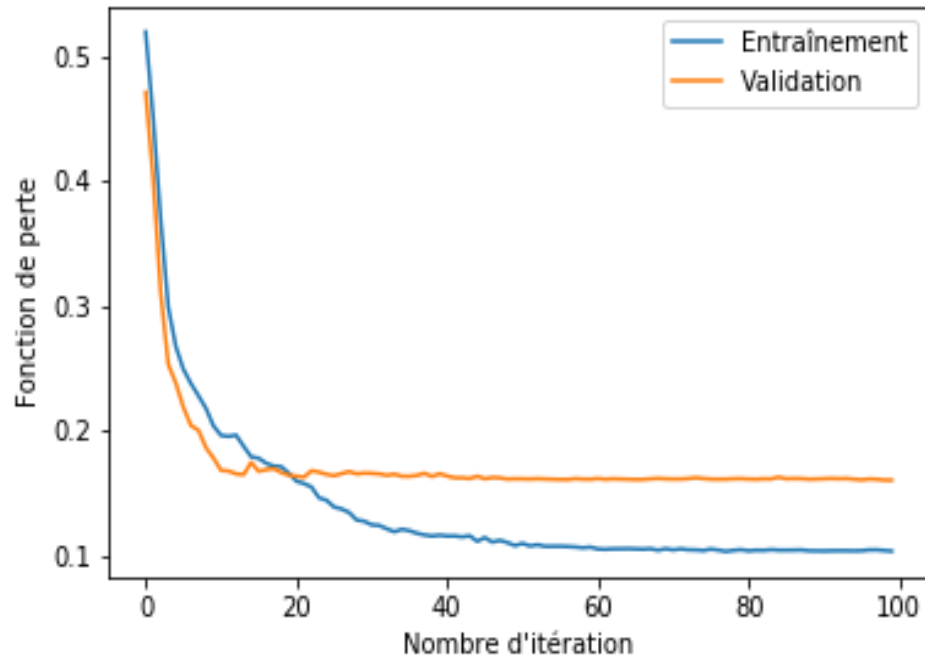


FIGURE 7 – Erreur d'entraînement et de test en fonction de Nombre d'itérations.

Les résultats concernant l'erreur relative : On remarque que 18 prédictions de débits ont un pourcentage d'erreur relative inférieure à 20% et 12 prédictions ont un pourcentage d'erreur entre 20% et 30%.

Les résultats concernant le coefficient d'efficacité Nash–Sutcliffe : On remarque qu'on a 30 valeurs strictement positives et 17 valeurs supérieures à 0.2.

Donc, on peut conclure que ce réseau de neurones peut nous permettre de prédire plus que la moitié de débits avec un taux d'erreur faible et avec une coefficient d'efficacité Nash–Sutcliffe supérieur à 0.2.

4 Conclusion

Avec le critère de Nash et l'erreur relative j'ai pu décider quel type d'architecture était le plus adéquat pour ce projet.

Dans l'ensemble, les résultats obtenus dans cette étude sont satisfaisants. Si nous comparons les prévisions que nous avons effectuées avec le débit observé, nous voyons que la simulation s'approche d'une manière satisfaisante de la réalité. En d'autres termes, le modèle est valorisé pour la prévision.

5 Bibliographie

Les liens suivants sont accessibles :

- Coefficient d'efficacité de nash–sutcliffe.
- keras et tensorflow.
- Une étude faite pour un problème de classification en utilisant keras et tensorflow.

6 Remerciements

Mes remerciements particuliers à Prof. Jonathan Jalbert pour la suggestion de ce sujet, son aide et sa supervision continues. Je tiens également à remercier ma soeur, Sanae Lotfi, de m'avoir mis en contact avec Prof. Jalbert.