



Utilisation des techniques de Deep Learning pour la
reconstruction des niveaux d'eau de l'estuaire de la Seine.

Julien DELOFFRE, Imad JANBAIN, Abderrahim JARDANI, Tan VU

Journées
REFMAR.shom.fr
Réseaux marégraphiques français



Observation du niveau marin
au profit des politiques publiques et de la recherche
Colloque international francophone

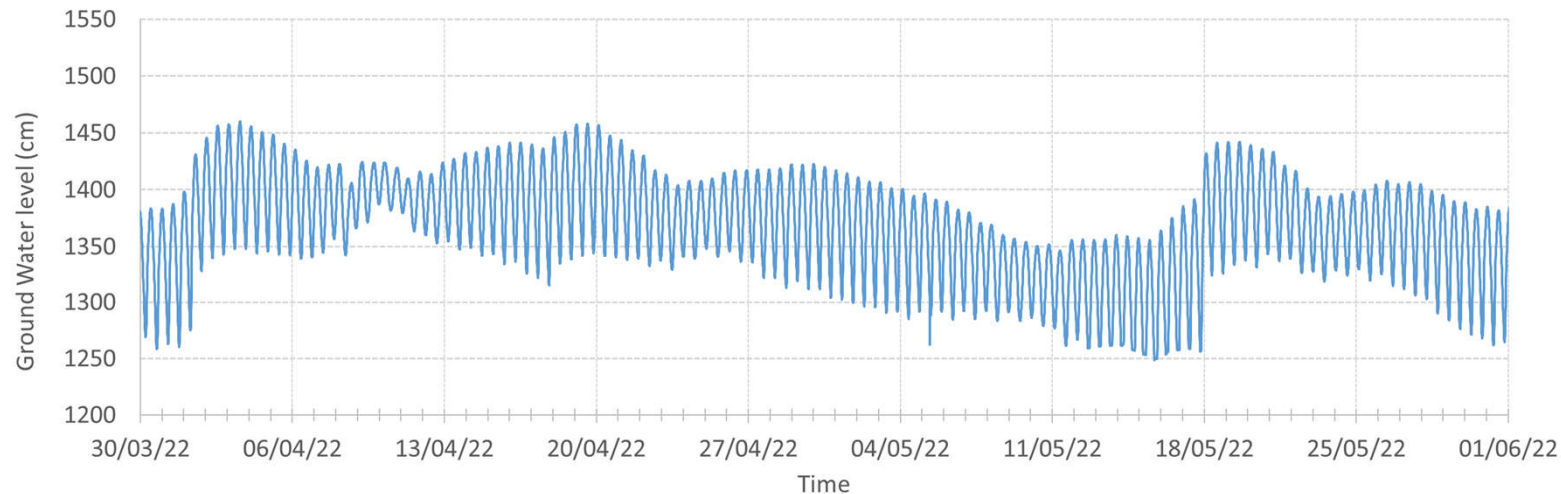
17 - 19 octobre 2022

à Océanopolis
BREST

#JR2022

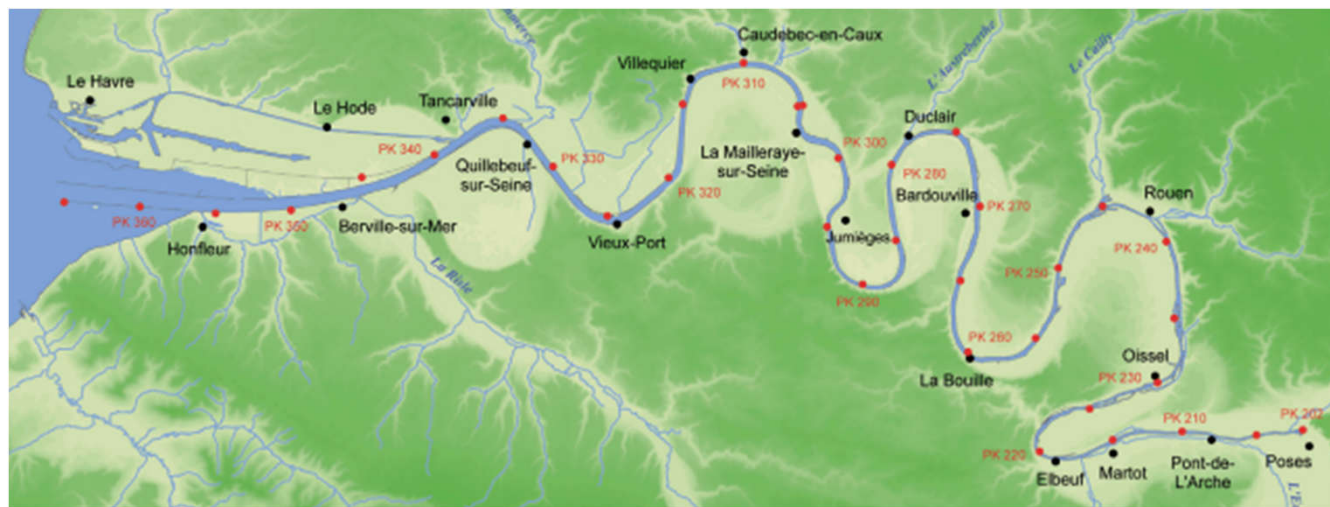
Problématique

- Estuaire => Relations entre les compartiments des eaux de surface (continentales & marines) & des eaux souterraines,
- Concomitance entre les marées, le cycle hydrologique (fleuve & affluents), les conditions hydro-météo et le rôle du sous-terrain.
- Quelles évolutions dans le cadre du changement climatique notamment sur lors des événements extrêmes (ressource en eau, inondation de la plaine alluviale, étiage extrêmement forts, remontée du biseau salé)?



Données piézométrique plaine alluviale de la Seine

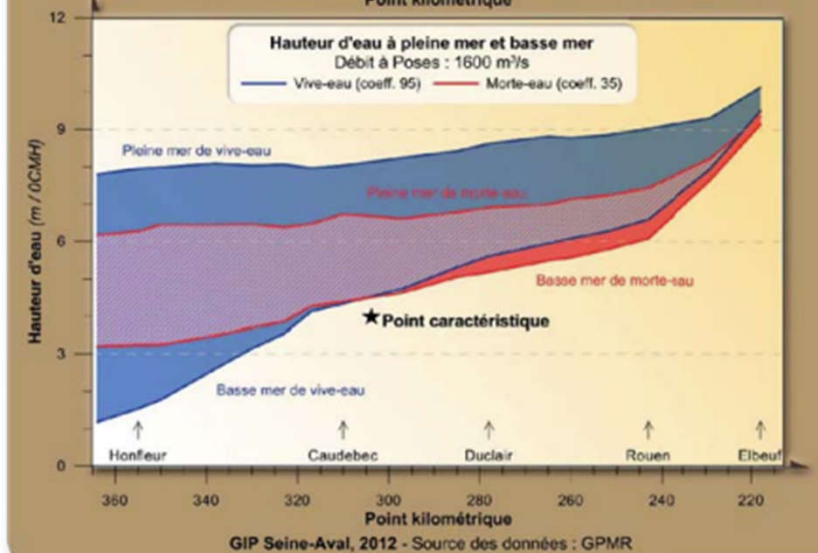
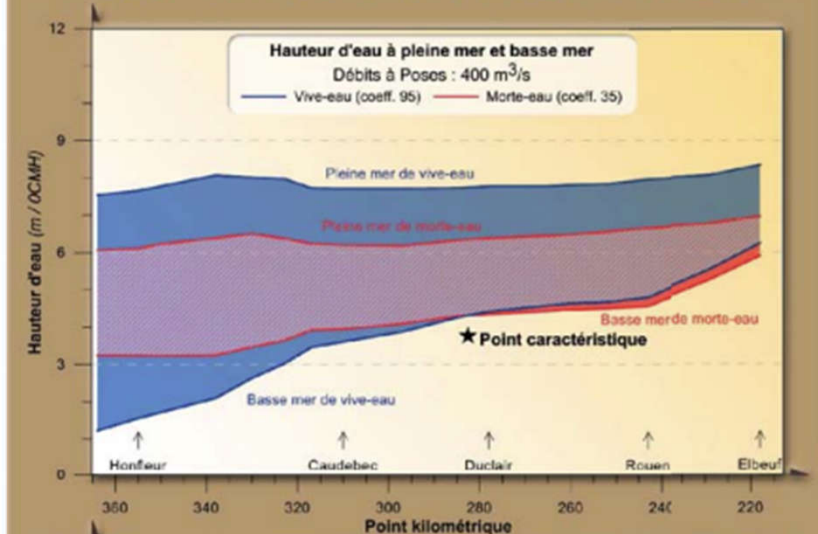
Site d'étude



Estuaire de la Seine

- ~ 160km de long
- Estuaire macrotidal (8.5m à l'embouchure)
- Débit moyen du fleuve $450 \text{ m}^3 \cdot \text{s}^{-1}$ (50-2,200 $\text{m}^3 \cdot \text{s}^{-1}$)
- Nappes puissantes et karstifiées (calcaires crétacé sup.)
- Fortement anthropisé (navigation)

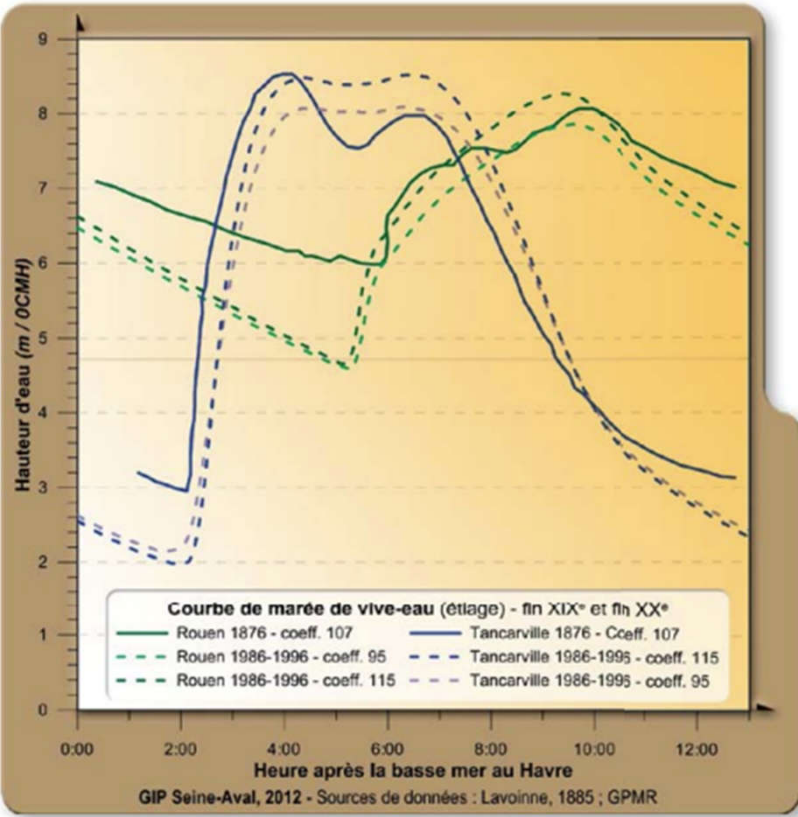
HAUTEURS D'EAU DANS L'ESTUAIRE DE LA SEINE SELON LE COEFFICIENT DE MARÉE ET LE DÉBIT DE LA SEINE



GIP Seine-Aval, 2012 - Source des données : GPMR

Seine aval, 2012

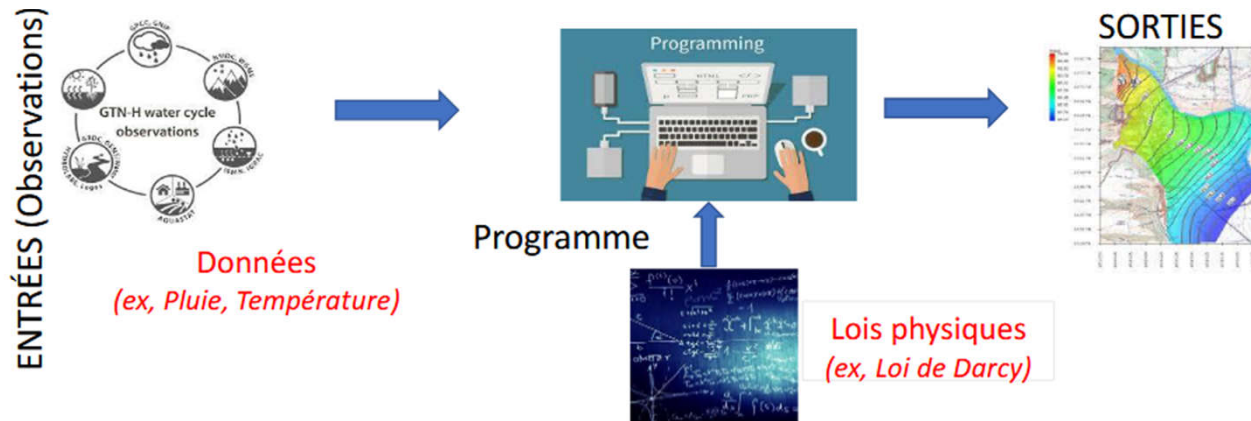
Site d'étude



Seine aval, 2012

Méthode d'apprentissage automatique

Approche classique de la modélisation



Machine learning

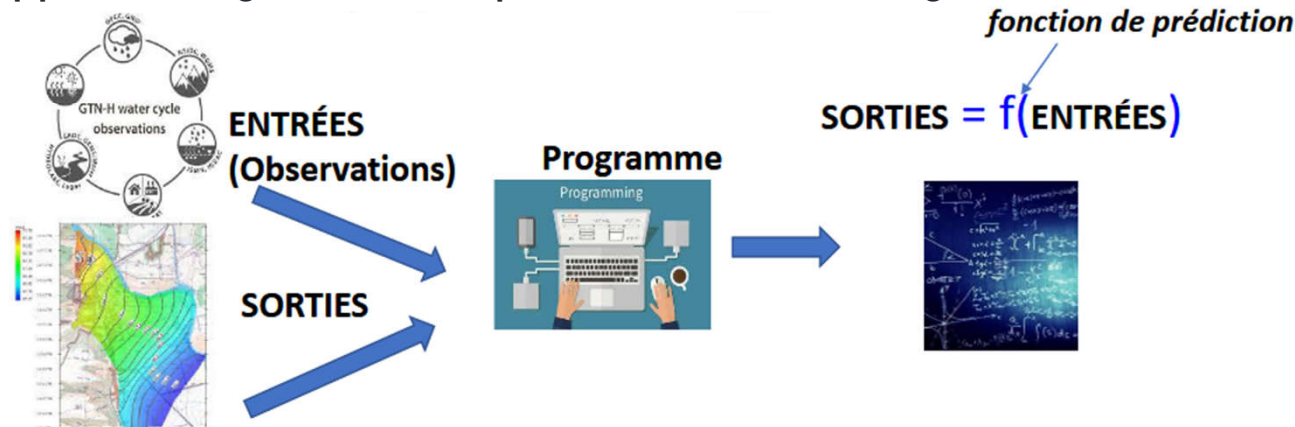
Avantages

- ✓ N'utilise aucune loi physique
- ✓ Possibilité d'intégrer plusieurs type de données
- ✓ Peut être exécuté en peu de temps lorsque le réseau est formé

Inconvénients

- ✓ Nécessite une grande base de données
- ✓ Nécessite des ressources informatiques performantes
- ✓ Inféodé aux données

Approche par apprentissage automatique (machine learning)



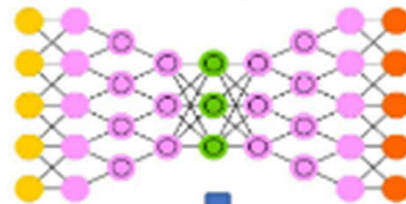
Méthode d'apprentissage automatique

Phase d'apprentissage: formation d'un réseau neuronal sur la base des données historiques

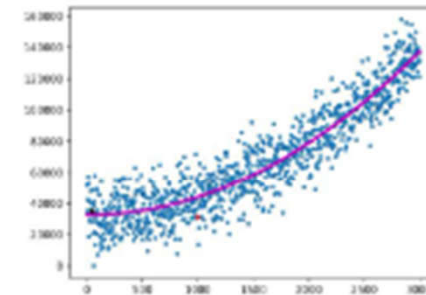
Données historiques



Deep Learning algorithm



Predictions

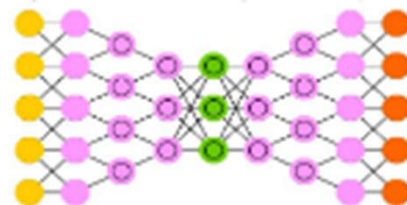


Phase de test

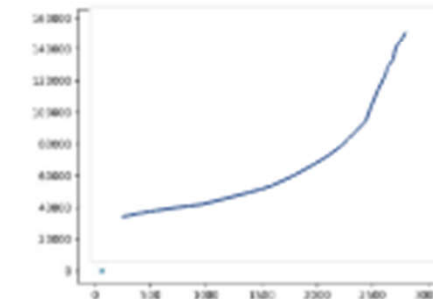
Données récentes



Réseau appris

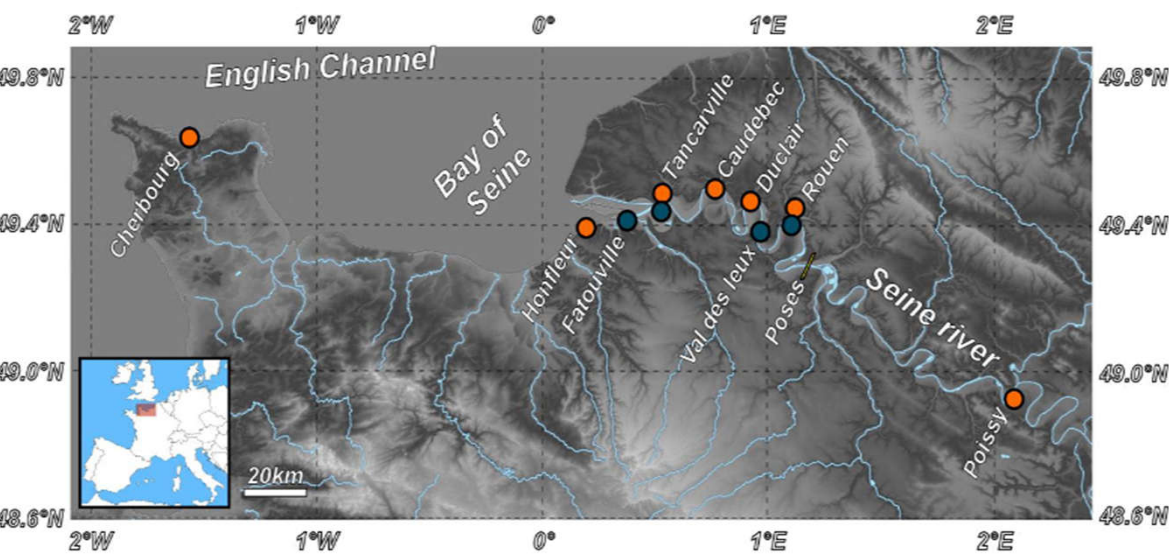


Predictions

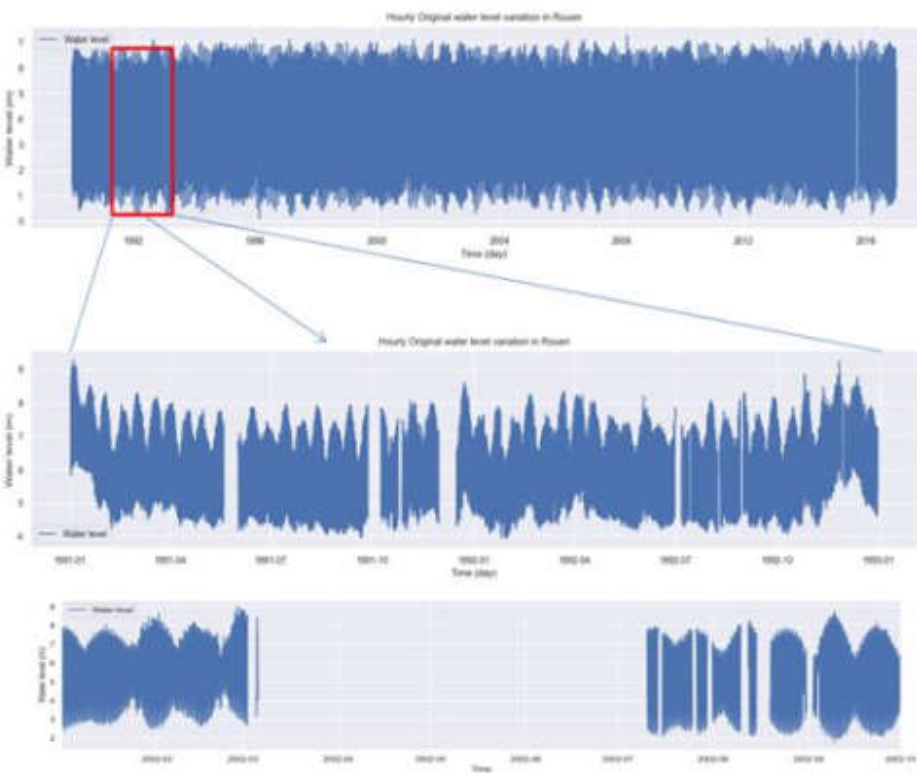


Reconstruction des niveaux d'eau en Seine (Janbain et al., soumis)

Problématique: Présence de données marégraphiques manquantes (jusque 6 mois).



Station : Rouen



Préparation des données

1. Données brutes

- ✓ Données de hauteurs d'eau sur 5 stations (Tancarville, Caudebec, Duclair, Rouen, Honfleur) entre 1990 et 2017.
- ✓ Pas d'échantillonnage: 1h.

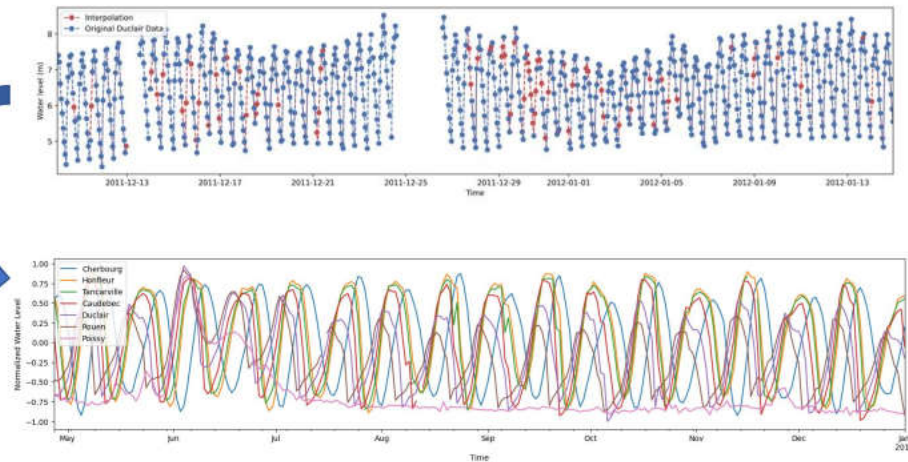
2. Nettoyage de données

- ✓ Remplir les petites parties manquantes (6 heures maximum) en utilisant l'interpolation linéaire afin d'obtenir des données plus complètes pour entraîner le modèle LSTM.

3. Normalisation de données

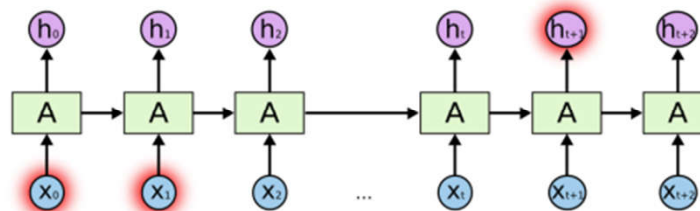
$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

	Honfleur	Tancarville	Caudebec	Duclair	Rouen
Counts	223039	224317	228886	225370	228326
Mean	5.05	5.23	5.73	6.07	6.37
Standard deviation	2.02	1.92	1.32	1.02	1.03
Min	0.76	1.36	3.03	3.95	3.92
25%	3.14	3.36	4.54	5.26	5.58
50%	5.23	5.35	5.74	6.12	6.4
75%	6.92	7.02	6.88	6.86	7.14
Max	9.1	9.98	9.3	9.82	9.86

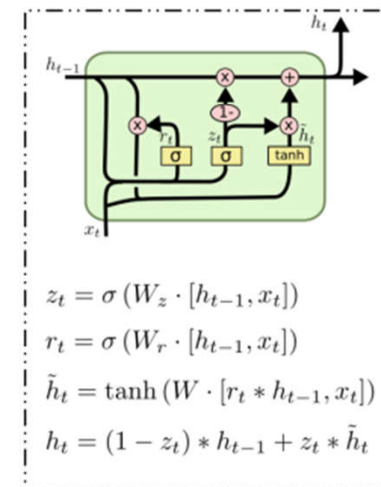


Traitement de données : réseau de type LSTM

$$f(t) = f(t-1, t-2, t-3, \dots, \text{parameters})$$

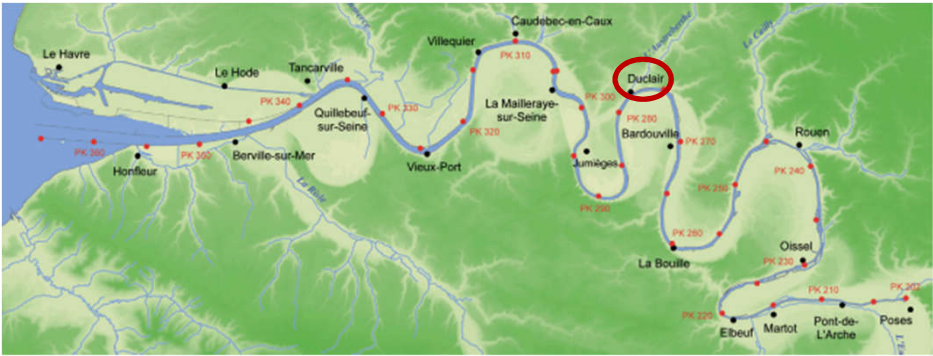


- un modèle simple en forme de boîte noire conçu pour établir une relation entre deux ou plusieurs variables dépendantes du temps.

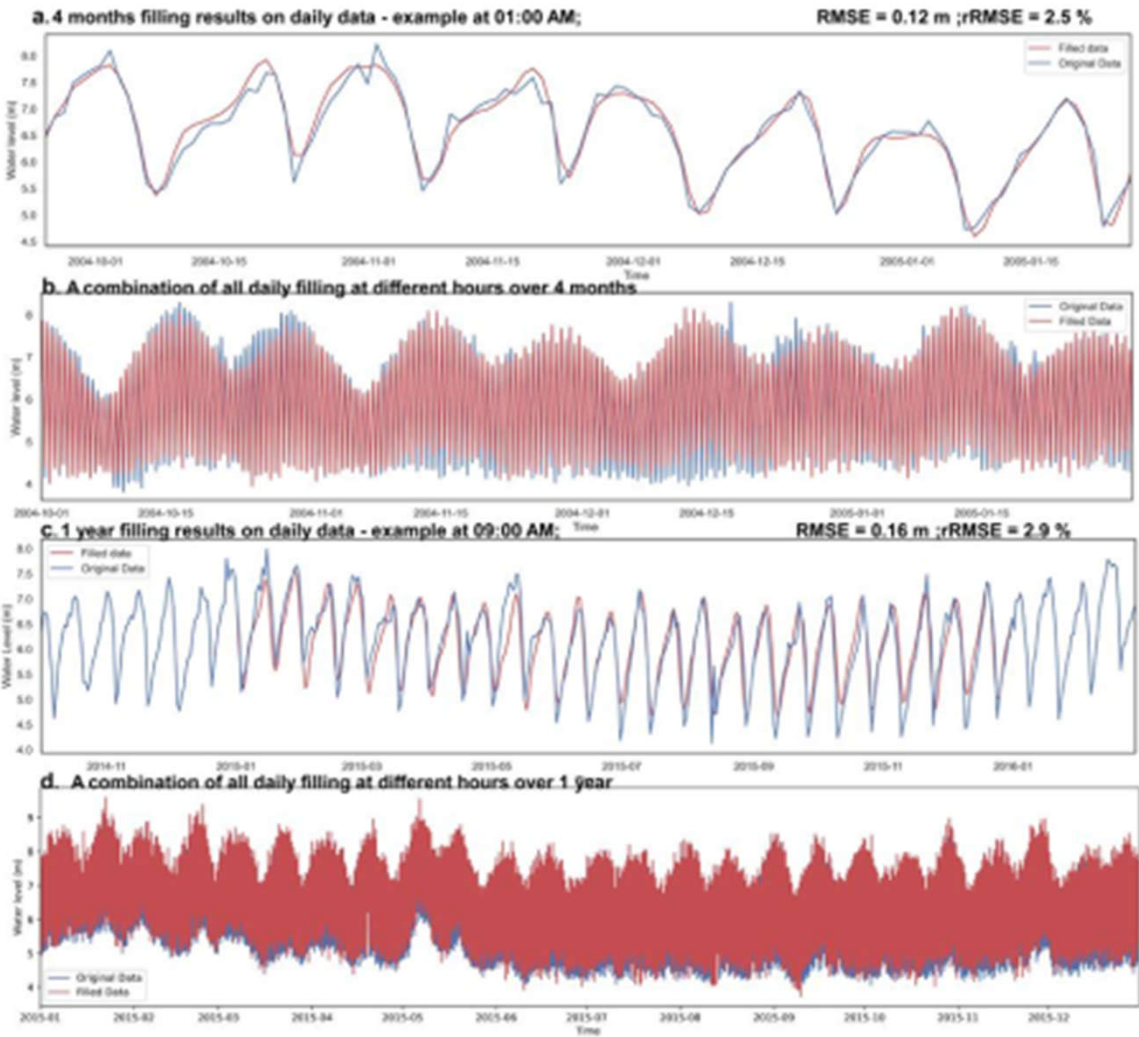


Reconstruction des niveaux d'eau en Seine (Janbain et al., soumis)

Apprentissage & validation

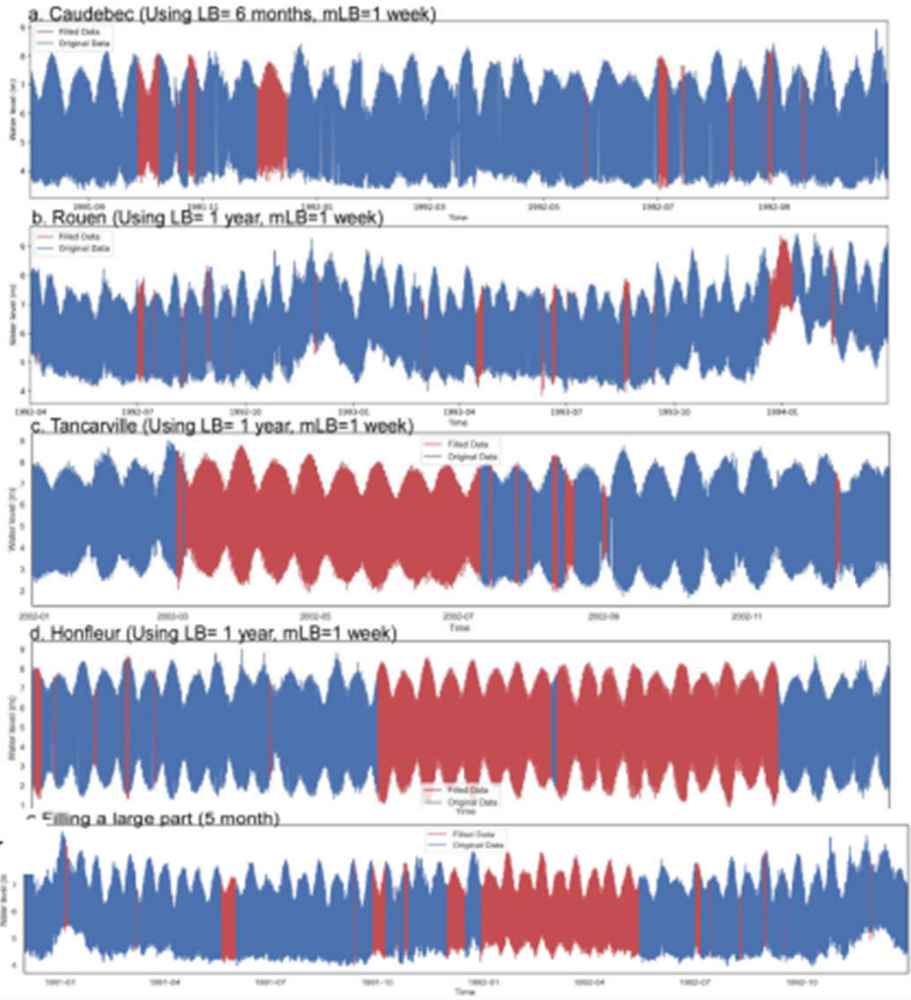
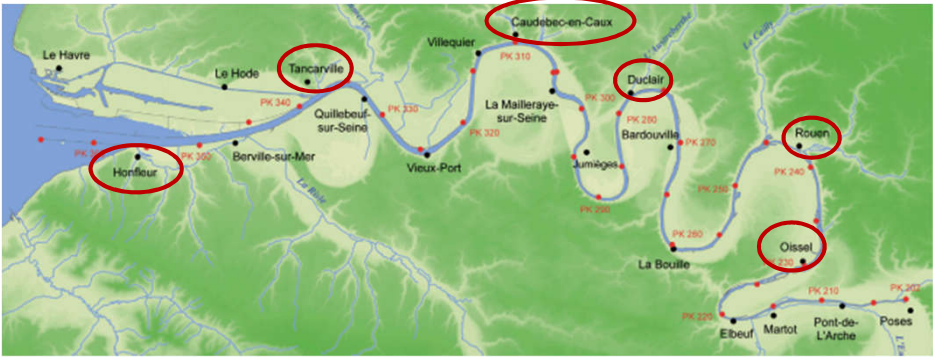


Duclair

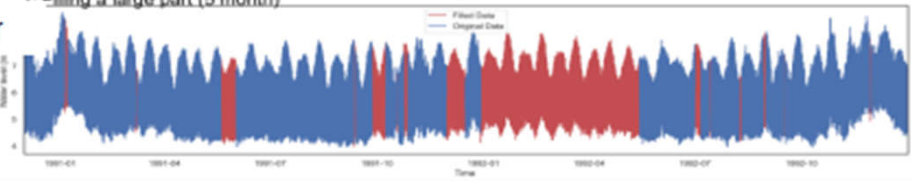


Reconstruction des niveaux d'eau en Seine (Janbain et al., soumis)

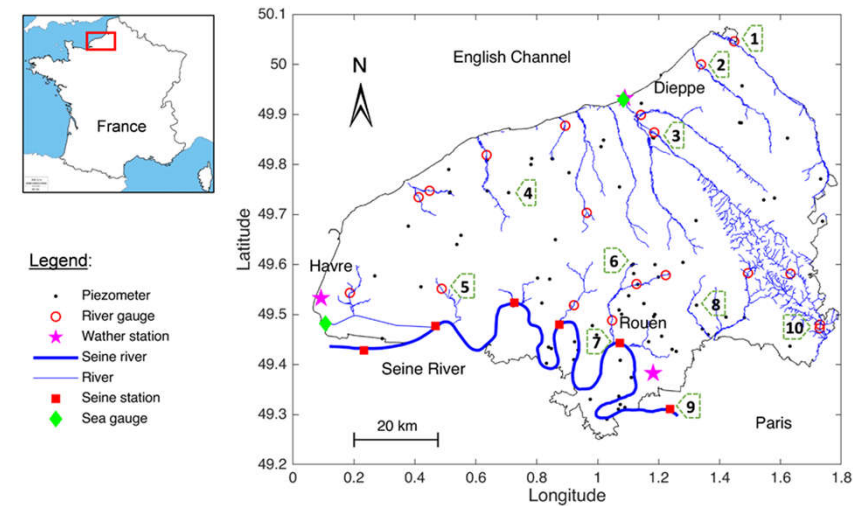
Résultats



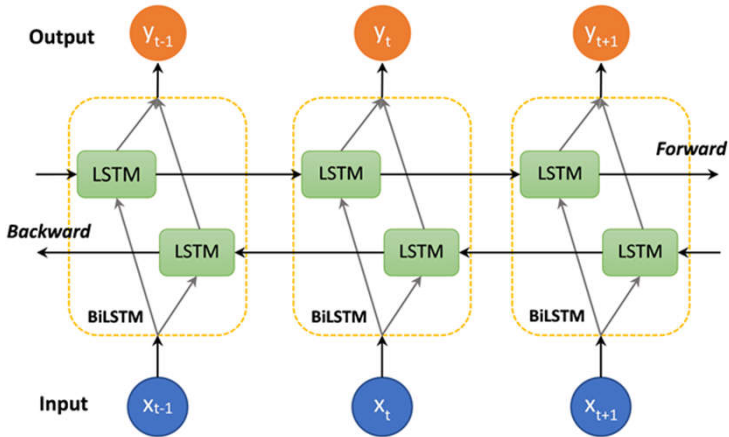
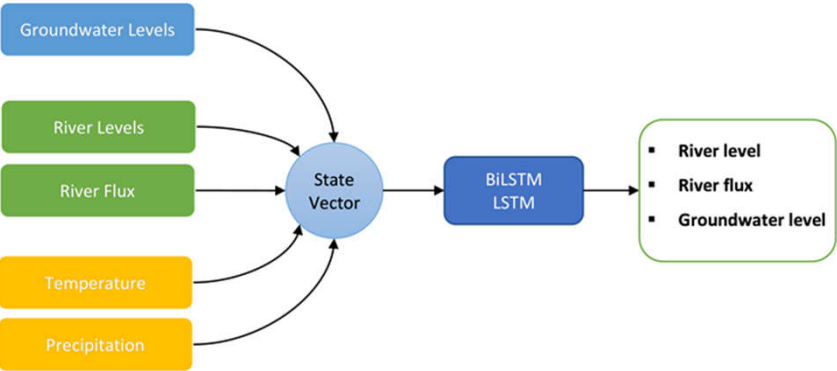
Duclair



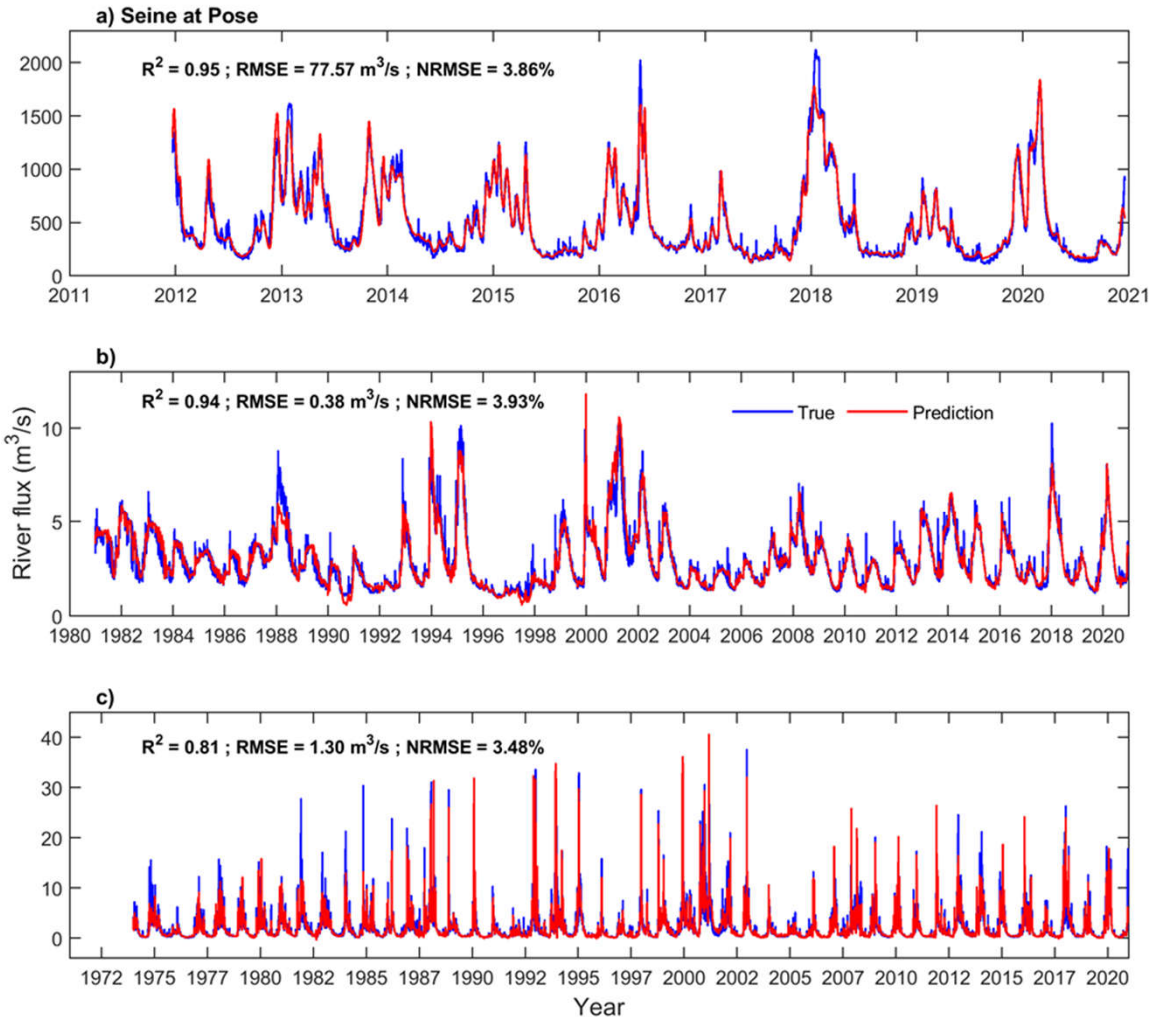
Traitement de données : réseau de type bi-LSTM



- ✓ 76 Piezometers,
- ✓ 18 Stations pour la surveillance du débit et du niveau des rivières.
- ✓ 3 Stations météo (Pluie+Temp)
- ✓ 2 Marégraphes
- ✓ 6 Stations sur la Seine



Résultats



Conclusions

- ✓ Approche particulièrement pertinente sur les estuaires,
 - Bases de données souvent riches (pour les 'grands estuaires')
 - Zones où les variations spatio-temporelles sont fortes,
 - Echelles de temps s'inscrivent dans des dynamiques long et court terme (e.g. houles vs cycles climatiques)
- ✓ Approche complémentaire à la modélisation numérique à base physique
- ✓ Prédictions : possible à 'court terme' et possibilité de travailler sur des prédictions 'long terme' sur plusieurs années en utilisant les projections climatiques du GIEC.

