



## Utilisation de l'IA pour le traitement des données du continuum Terre-Mer.

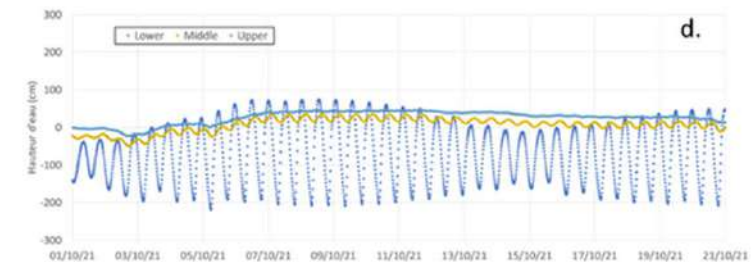
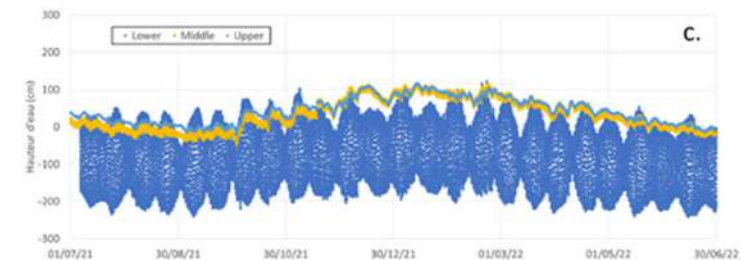
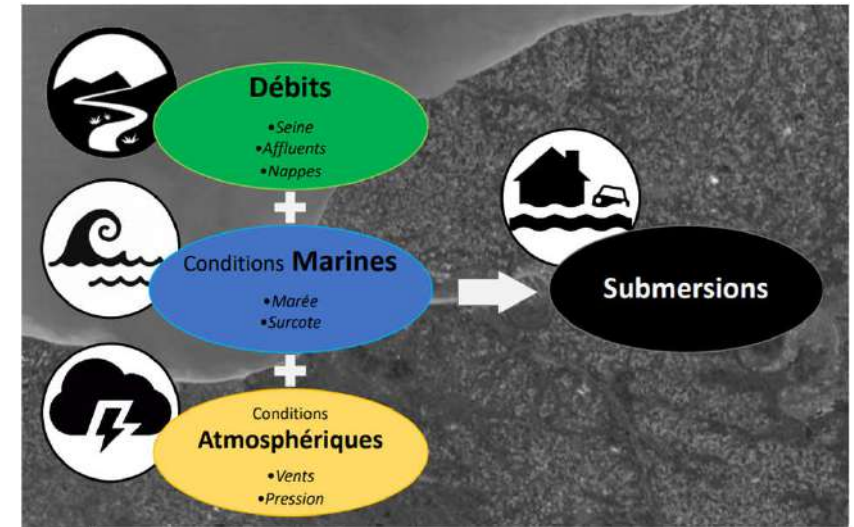
Deloffre, J., Vu, T., Janbain, I., Jardani, A., Gilbert, R., Huybrecht, N., Laborie, V., Lemoine, J.P., Massei, N., Salameh, E., Sivaramakrishnareddy, C., & Turki, I.

## 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 => problématique inondation/submersion
- Forte variabilité spatio-temporelle associée à de forts gradients,
- Quelles évolutions dans le cadre du changement climatique notamment sur lors des évènements extrêmes (ressource en eau, inondation de la plaine alluviale, étiage extrêmement forts, remontée du biseau salé)?

=> Complexité pour embrasser les processus physico-chimiques à l'échelle de l'ensemble de l'estuaire

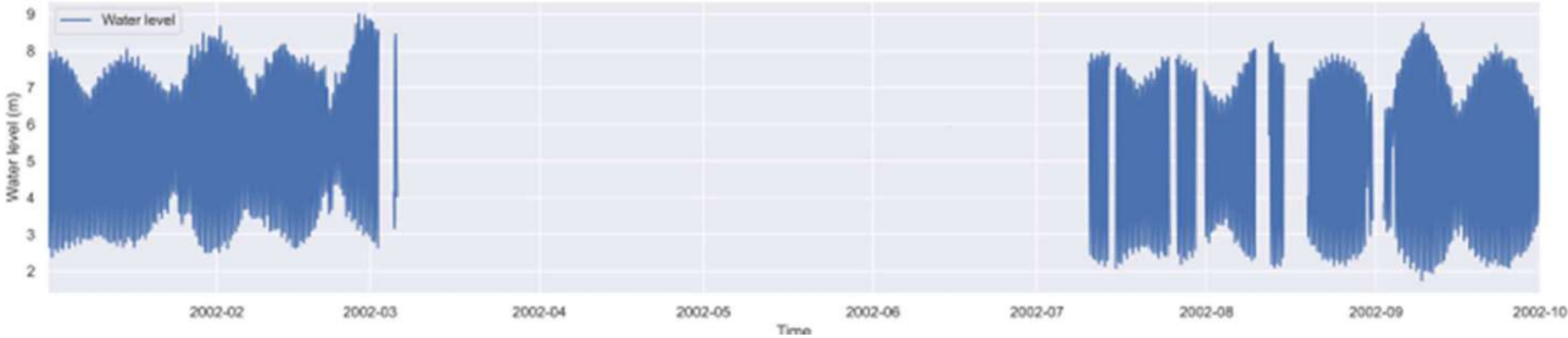
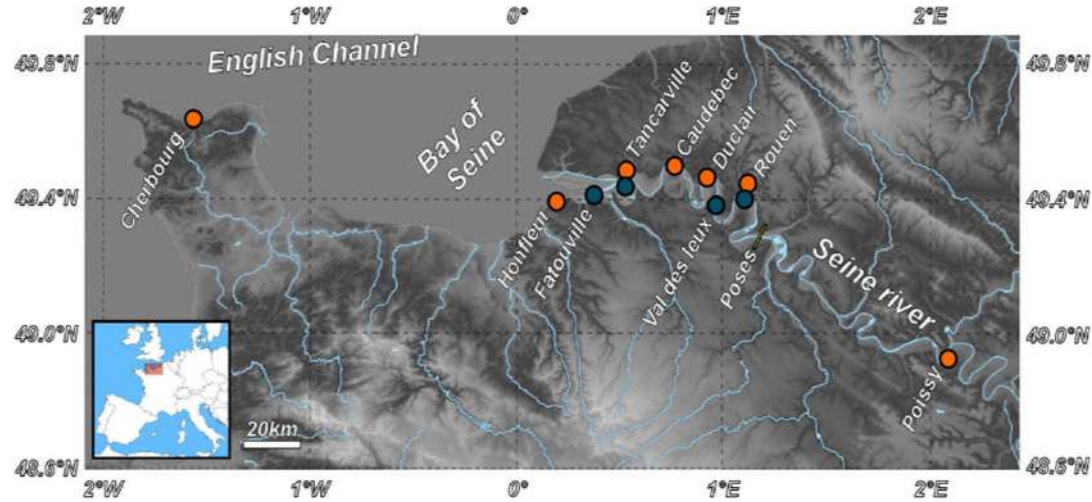
=> **Pertinence des approches statistiques?**



Évolution des niveaux piézométriques

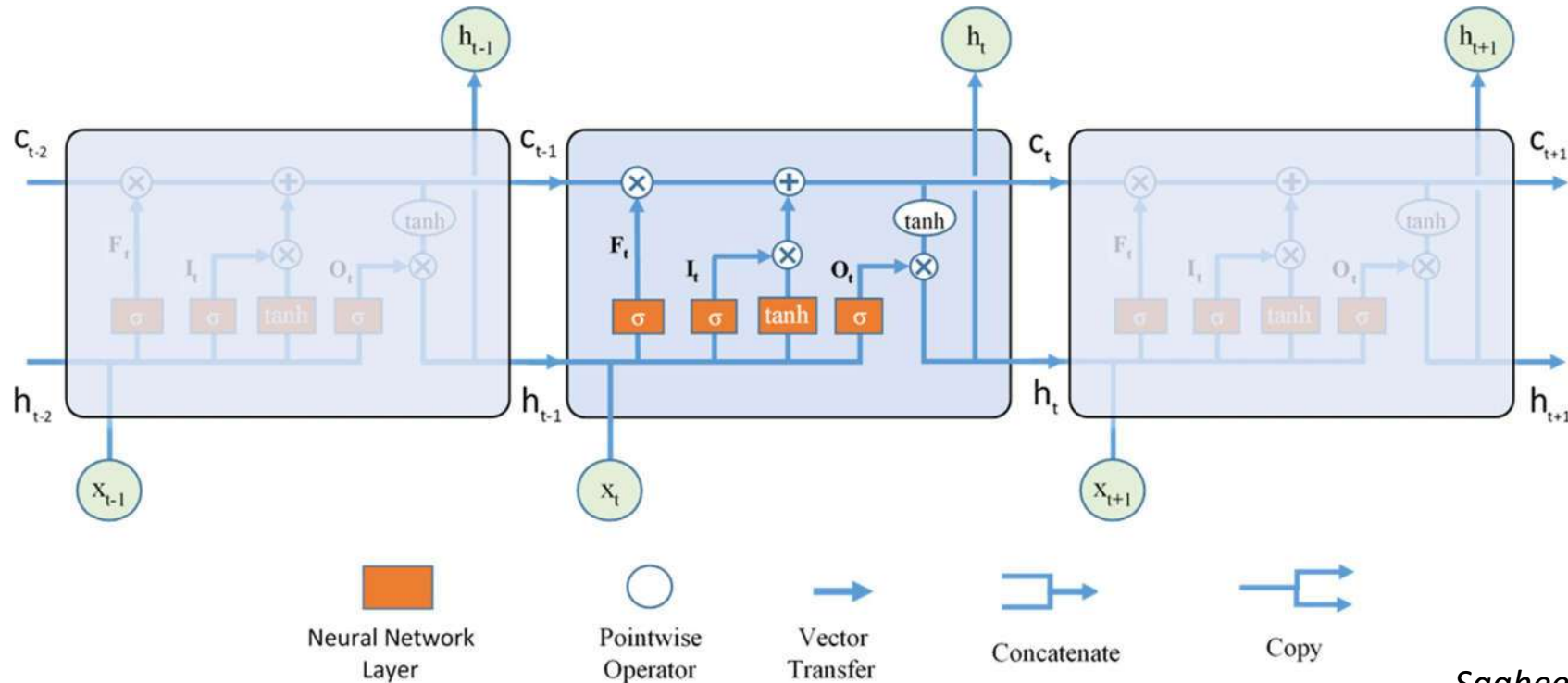
# Reconstruction des niveaux d'eau en Seine (Janbain et al., 2023a)

Problématique: Malgré l'importance de la donnée (sécurité de la navigation) il existe des données manquantes dans les marégraphiques (jusque 6 mois).



## Reconstruction des niveaux d'eau en Seine (Janbain et al., 2023a)

### Traitement de données : réseau de type LSTM

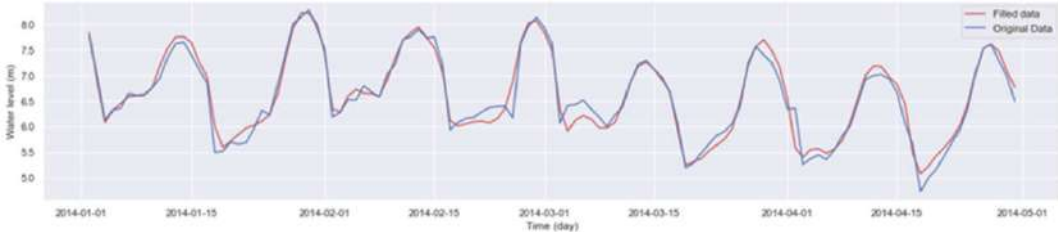
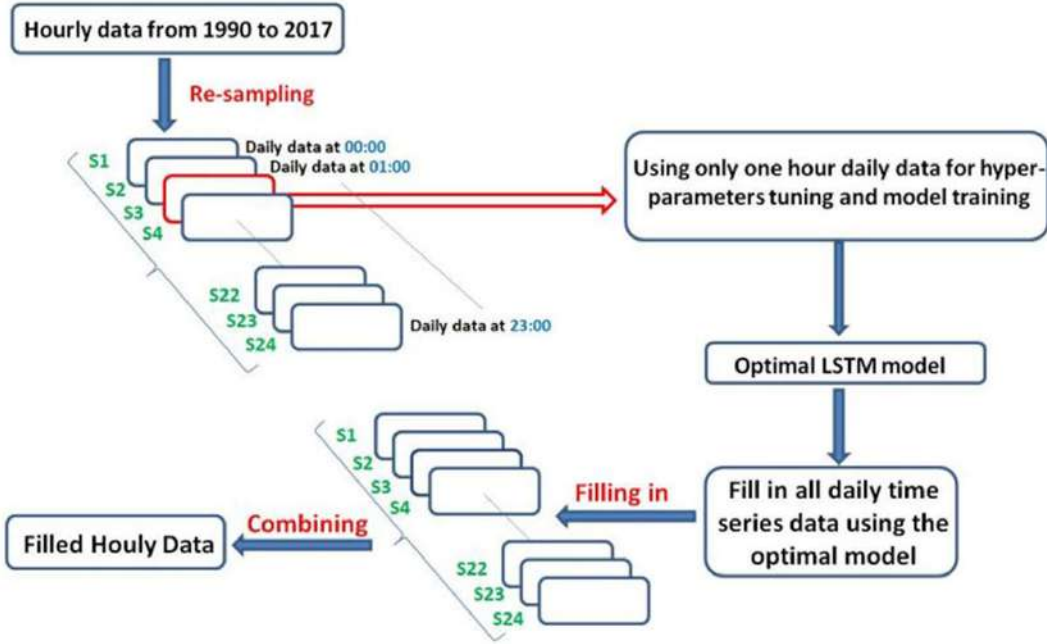


*Sagheer & Kotb, 2019*

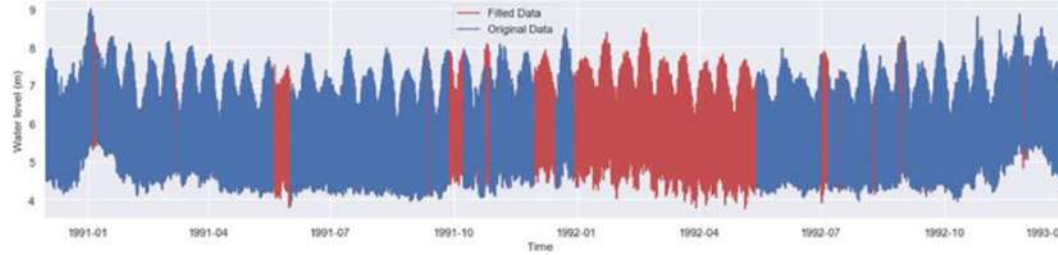
Type de RNN efficace pour traiter les variables dépendantes du temps des séries temporelles,

# Reconstruction des niveaux d'eau en Seine (Janbain et al., 2023a)

## Traitement de données : réseau de type LSTM

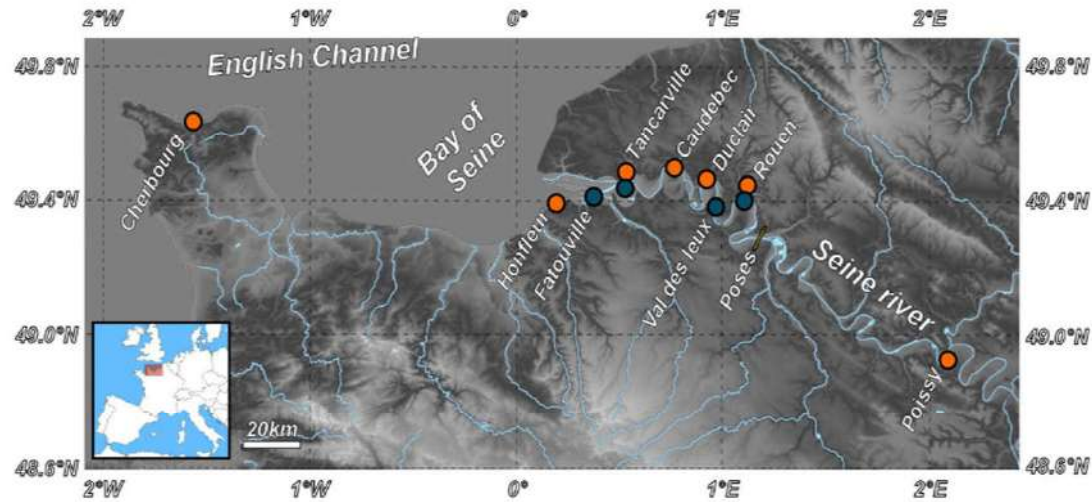


filling in four months of daily data in the case of Duclair by comparison with the original data (RMSE % = 5.8%).

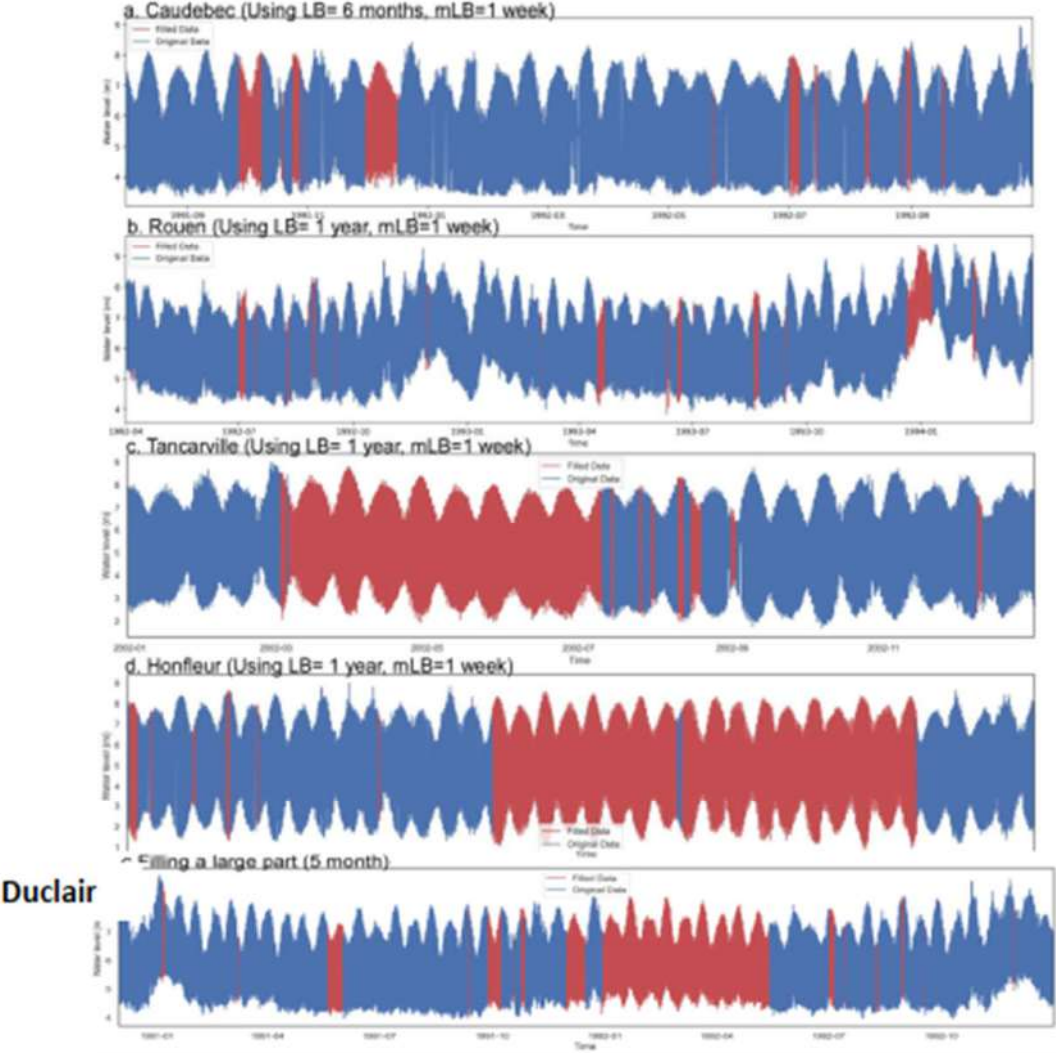


filling in a large part (5 months) of the missing data (RMSE %= 5.8%).

# Reconstruction des niveaux d'eau en Seine (Janbain et al., 2023a)



The model successfully filled all missing data sizes up to 4.5 months with low RMSE % values: one-week RMSE % = 1.9%, one-month RMSE % = 3.4, and 4.5-month RMSE % = 5.8%.



## Application à plus grande échelle (Vu et al., soumis)

Objectif: examiner l'évolution des débits des rivières recueillies au cours d'un siècle à l'échelle mondiale

Problématique: missing data=> reconstruction des débits de 136 rivières dans le monde (50% des apports)

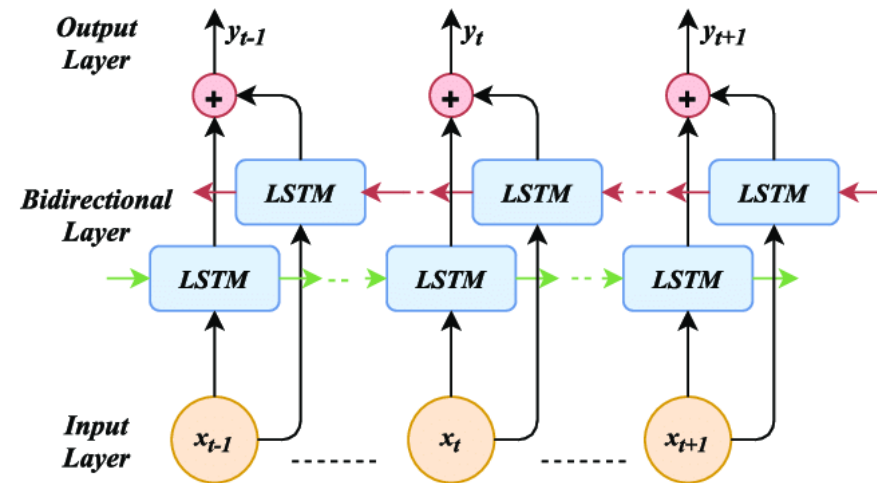
	River name	Outflow station	Lon	Lat	Obs. rate (%)	Recons. rate (%)	Change (%)
<b>Africa</b>							
1	Benue	Garoua	9.30	13.38	46.83	53.17	-1.86
2	Congo	Kinshasa	-4.30	15.30	91.00	9.00	-0.34
3	Limpopo	Beitbridge Pumpstation C/S	-22.22	29.98	28.25	71.75	0.85

Méthodologie de reconstruction :Bi-LSTM

deux couches cachées de LSTM dans des directions opposées

- des états vers l'avant pour les positions temporelles positives
- des états vers l'arrière pour les directions temporelles négatives

⇒ Prise en compte des états passés et futurs, ce qui facilite la reconstruction de séries temporelles



## Application à plus grande échelle (Vu et al., 2023)

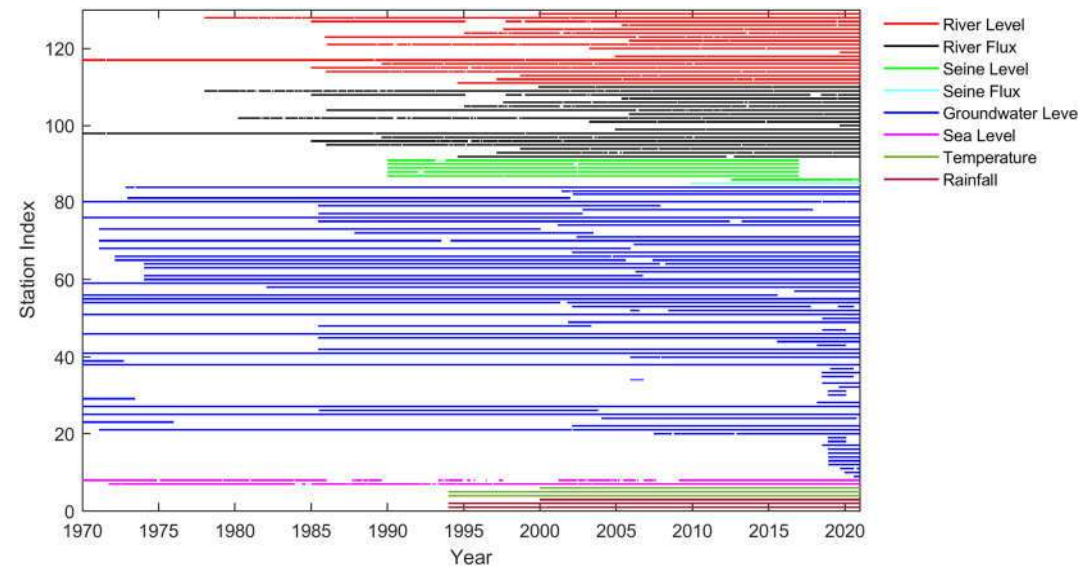
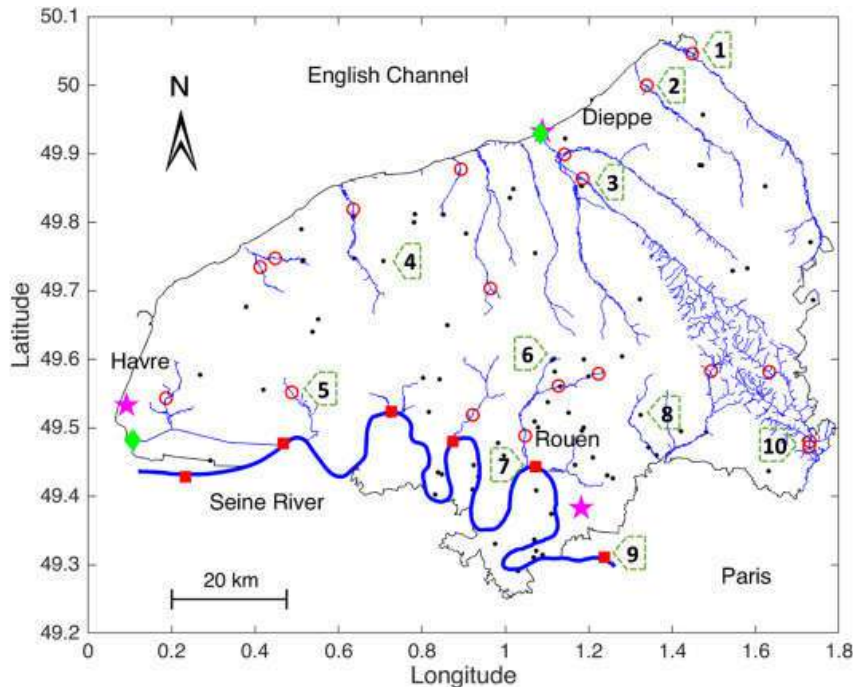
Objectif: intégrer une vision holistique de la masse d'eau estuarienne

Problématique: coupler des données de piézométrie, marégraphie, débit, pluviométrie,; météorologiques (pluie et température)



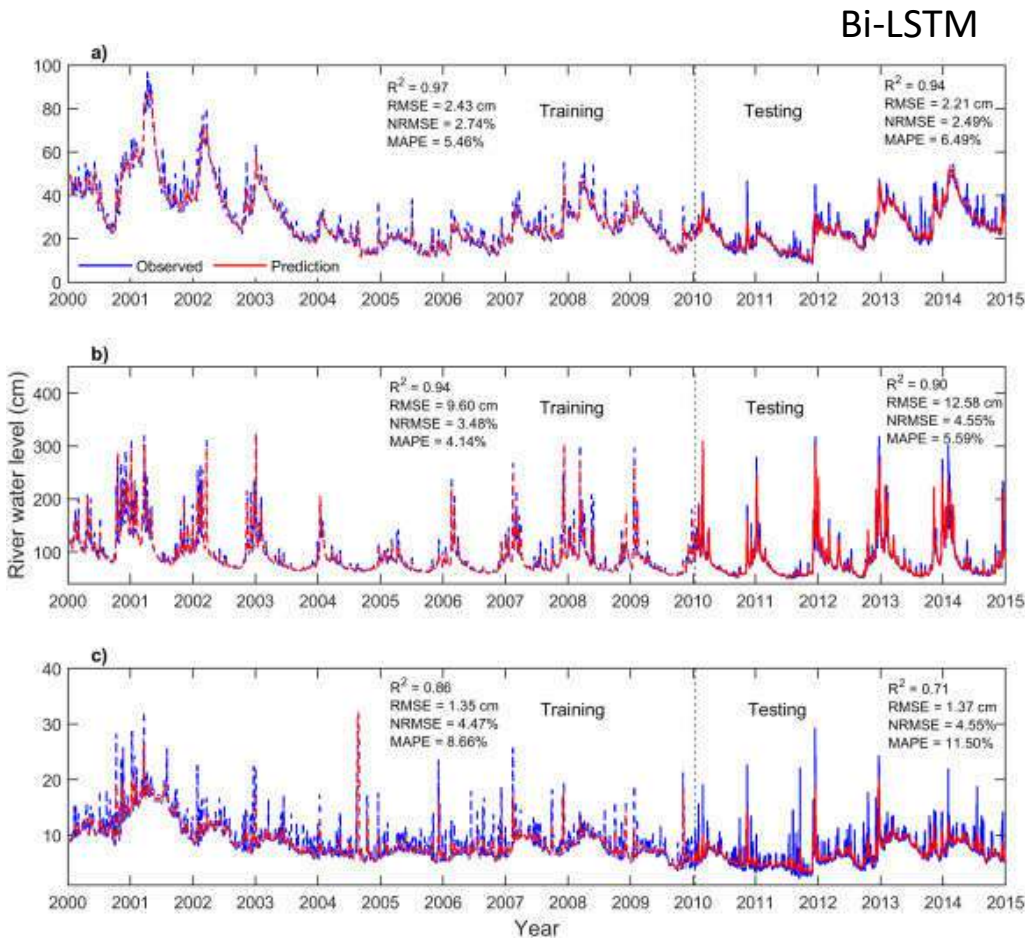
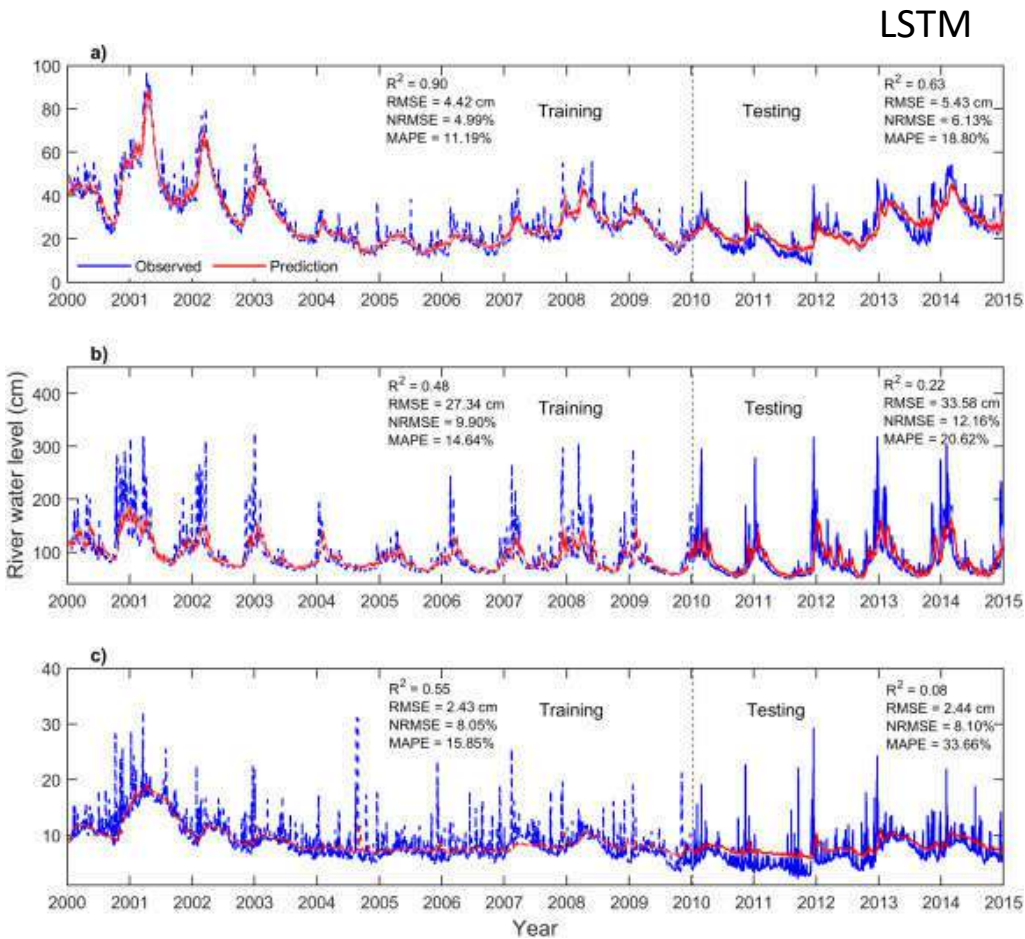
### Legend:

- Piezometer
- River gauge
- ★ Wather station
- Seine river
- River
- Seine station
- ◆ Sea gauge

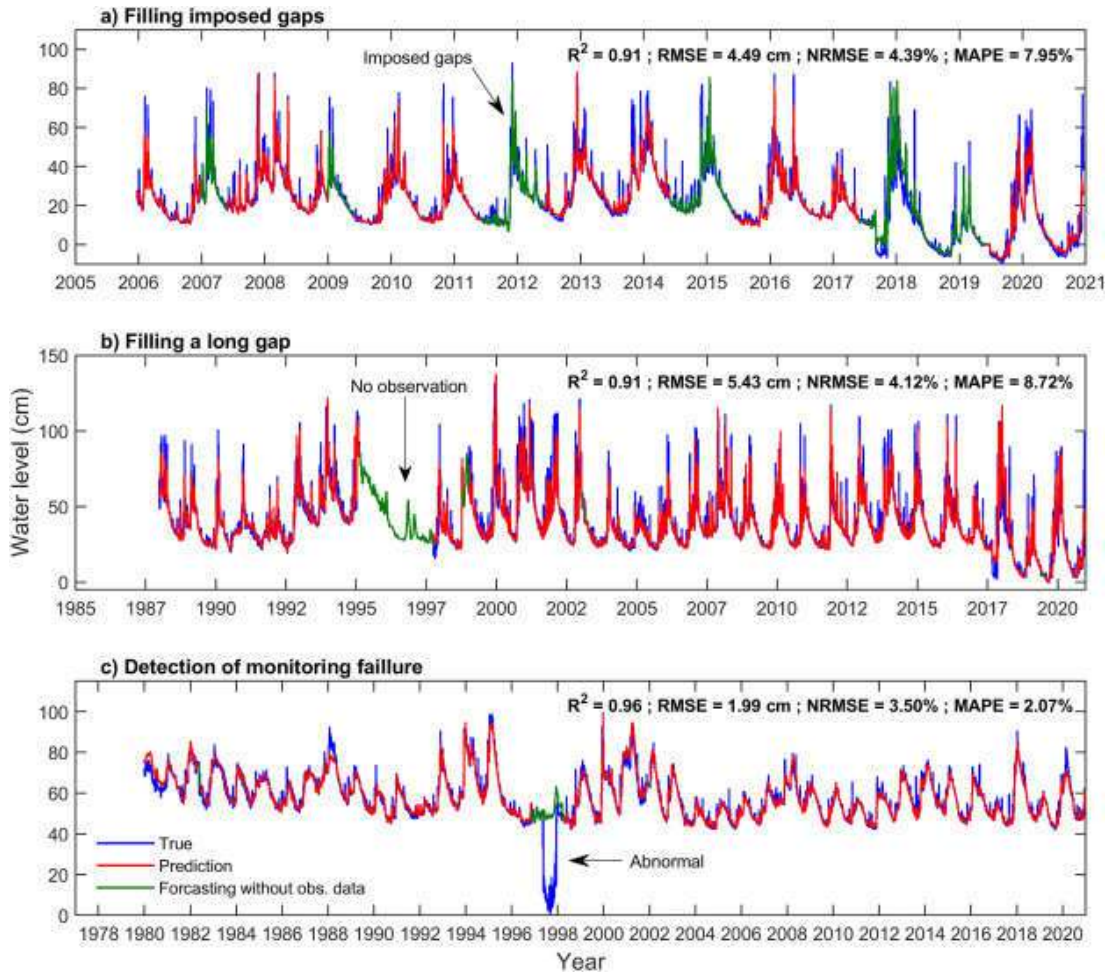




### Méthodologie bi-LSTM and LSTM



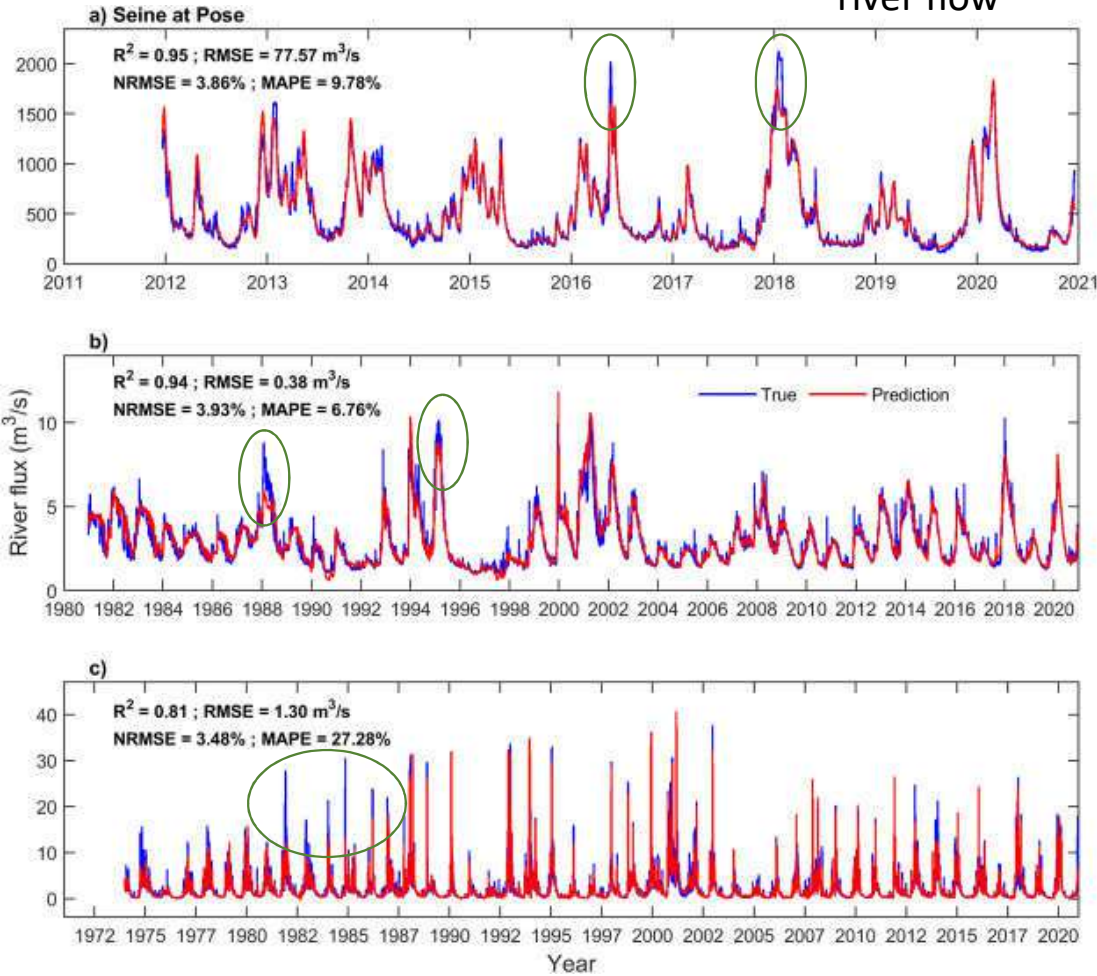
## Résultats



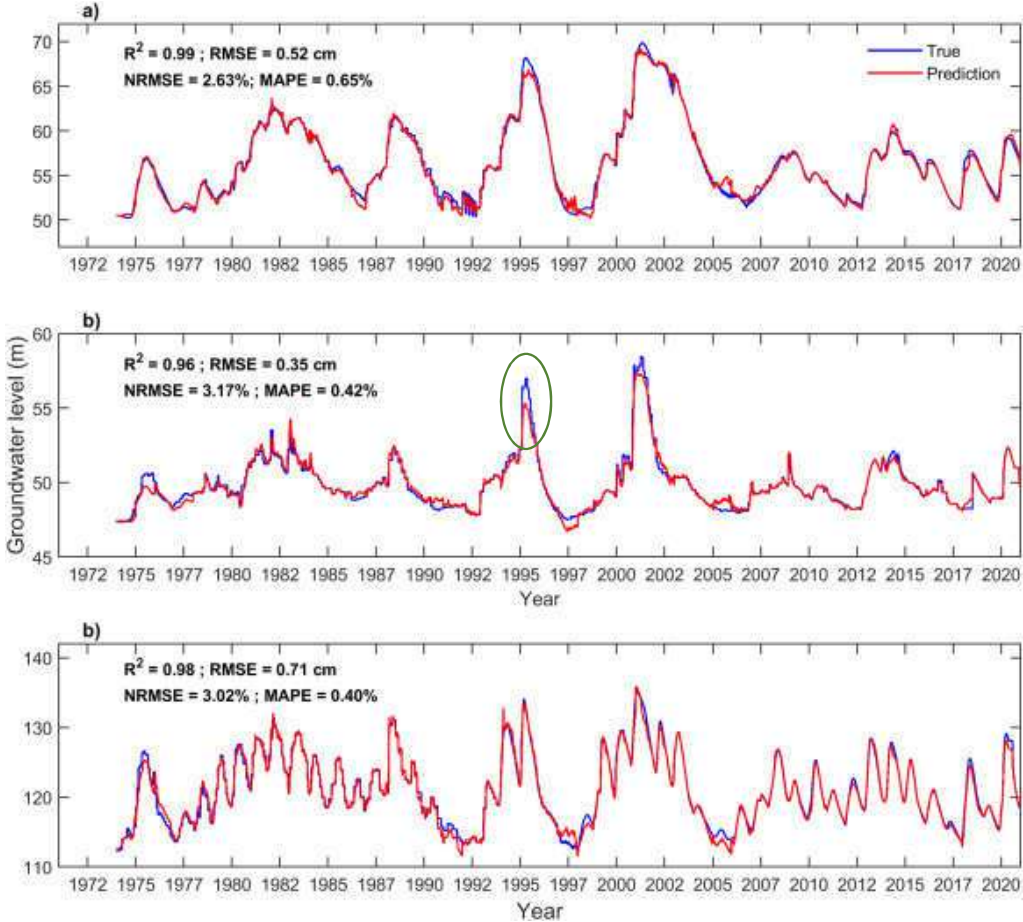
Application de la méthodologie pour combler les lacunes des série de mesure. Le système peut combler les lacunes et détecter les anomalies avec de longues périodes d'absence.

# Résultats

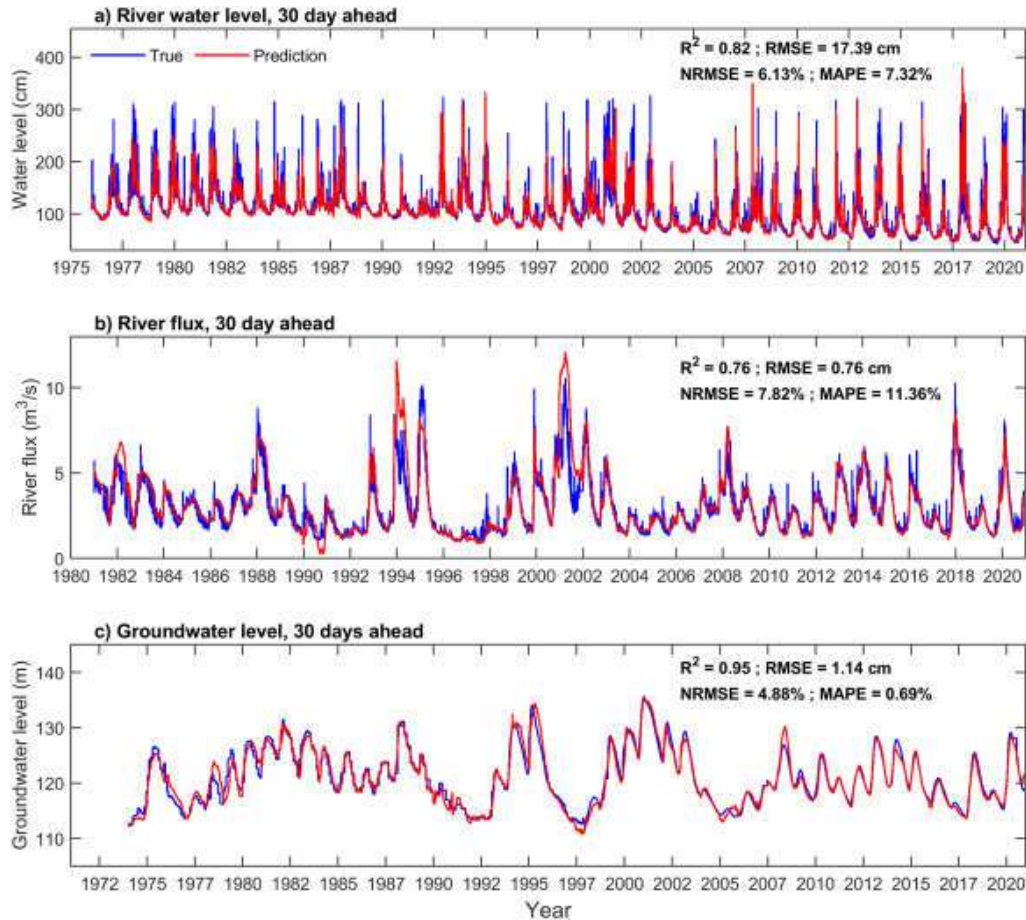
river flow



Groundwater level



## Application à plus grande échelle (Vu et al., 2023)



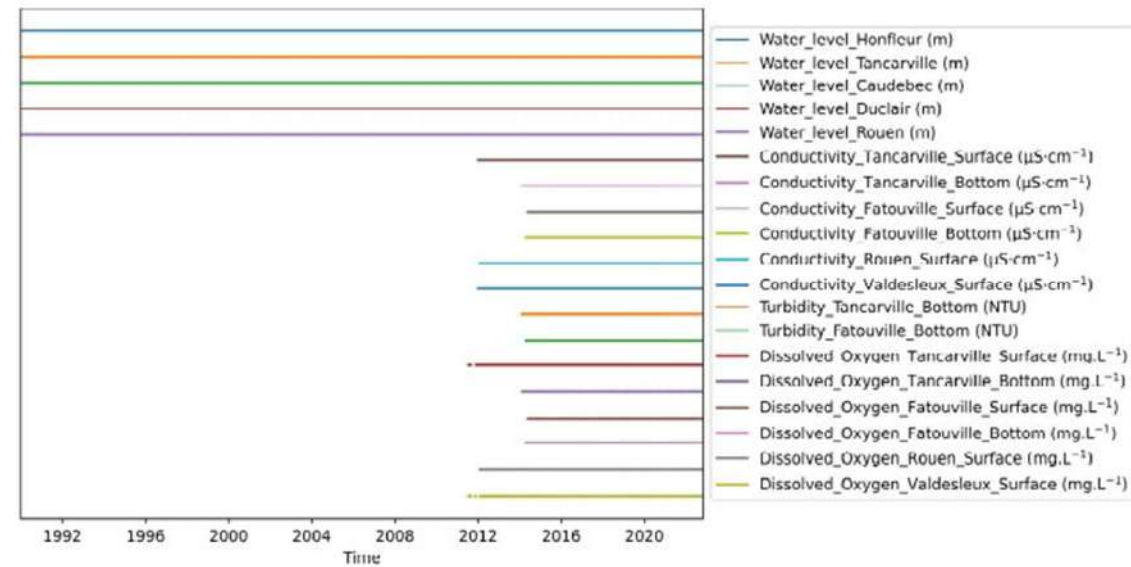
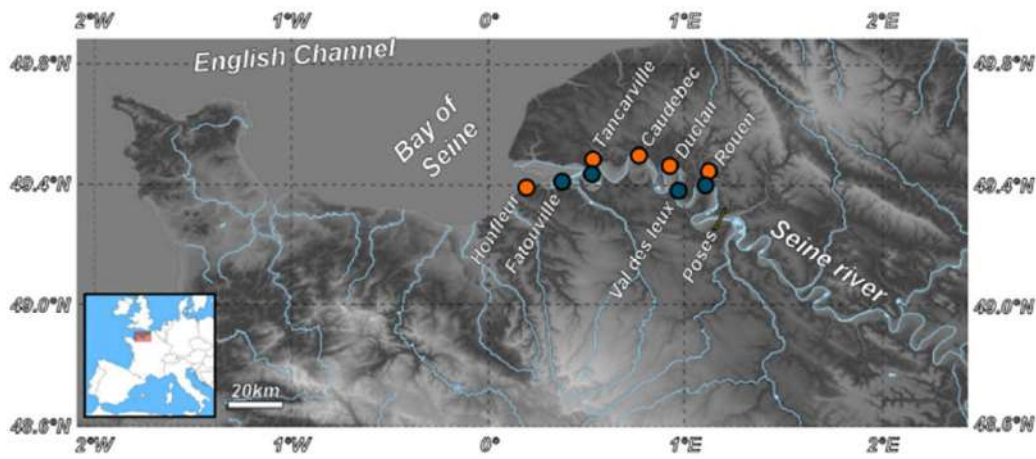
Prévisions à 'long terme' (30 jours) pour différentes dynamiques de l'eau. Les prévisions à plus long terme réduisent la précision de la prédiction, mais les résultats restent fiables.

Une prévision à long terme favorise le signal à basse fréquence, tandis que le signal à haute fréquence subit une certaine perte.

## Paramètres physico-chimiques (Janbain et al., 2023b)

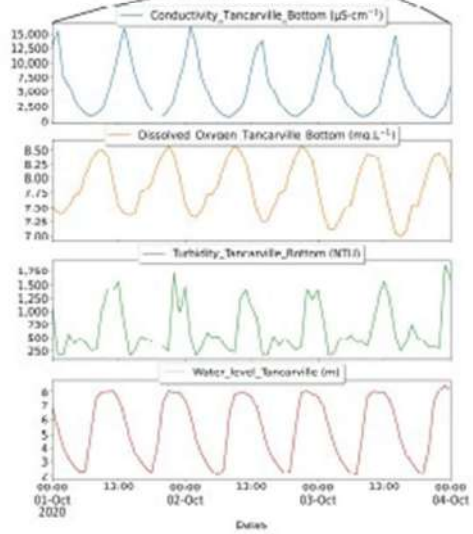
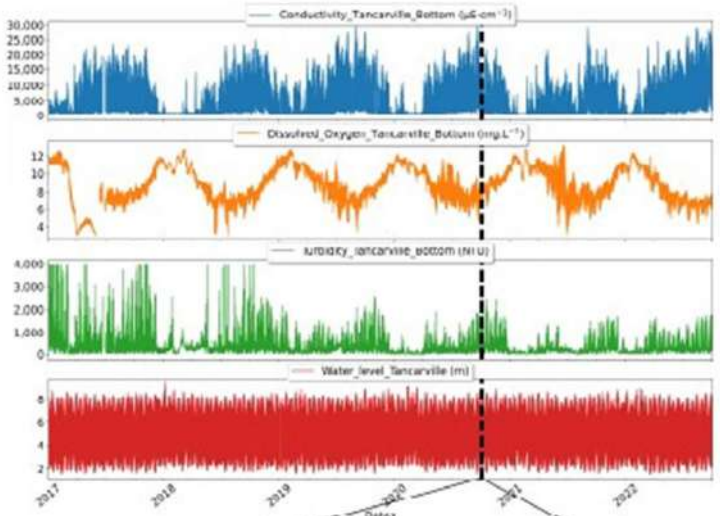
Objectif: reconstruire les données historiques de qualité de l'eau (turbidité, conductivité, oxygène dissous)

Problématique: relations entre les paramètres?

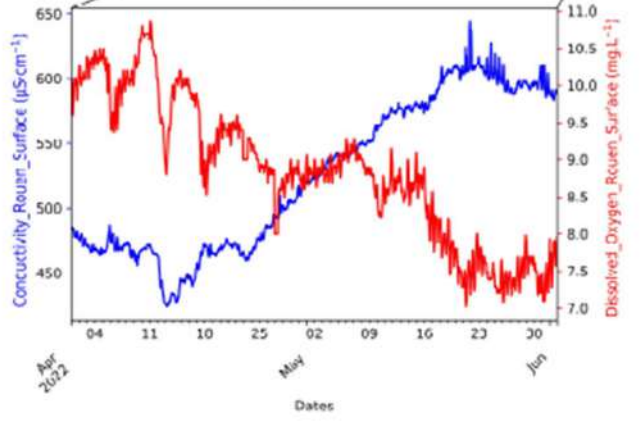
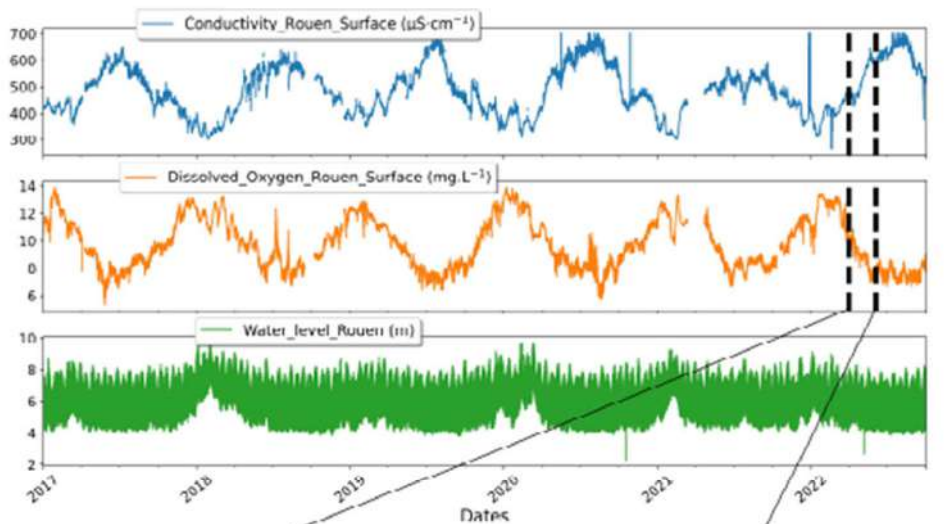


Méthologies: utilisation de 4 modèles

# Paramètres physico-chimiques (Janbain et al., 2023b)



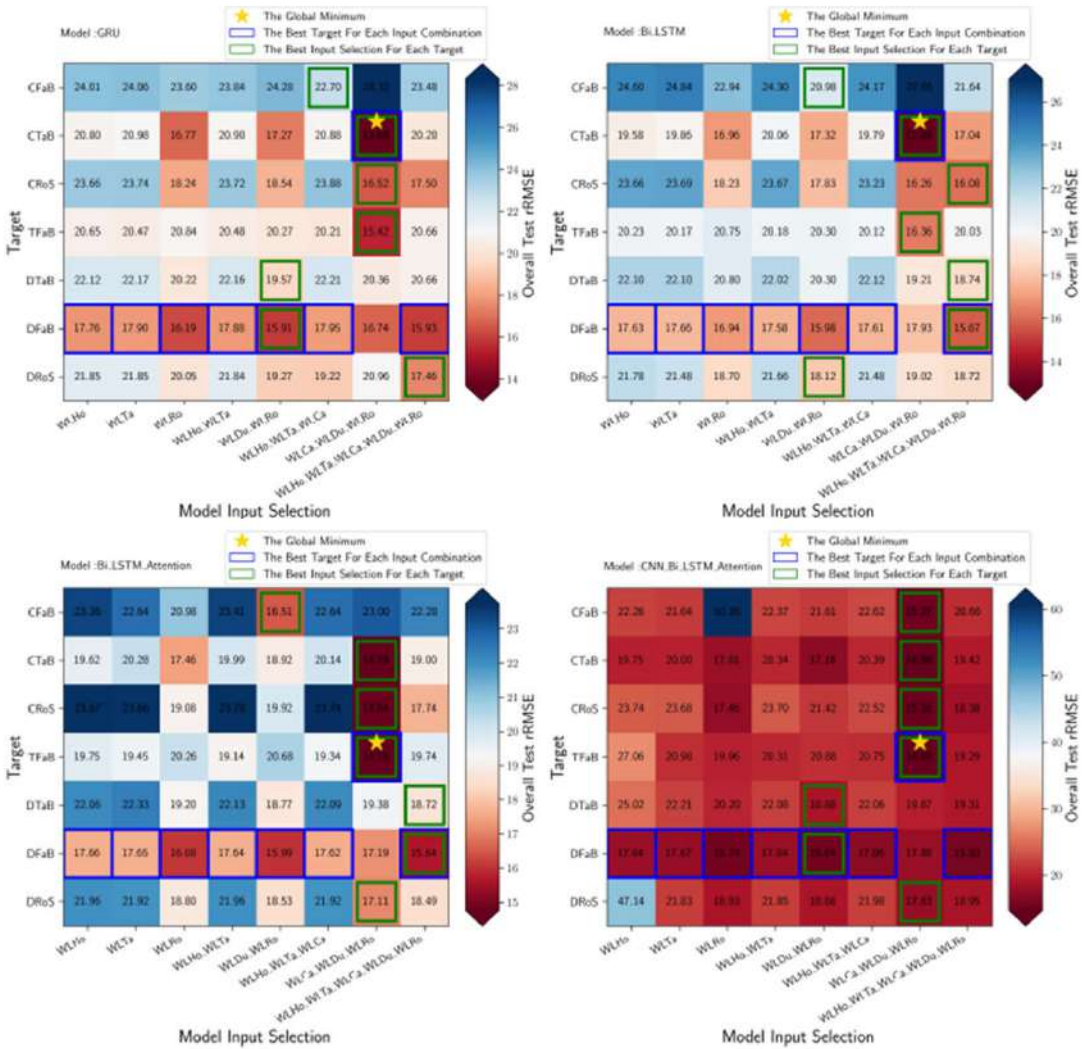
AVAL



AMONT

# Paramètres physico-chimiques (Janbain et al., 2023b)

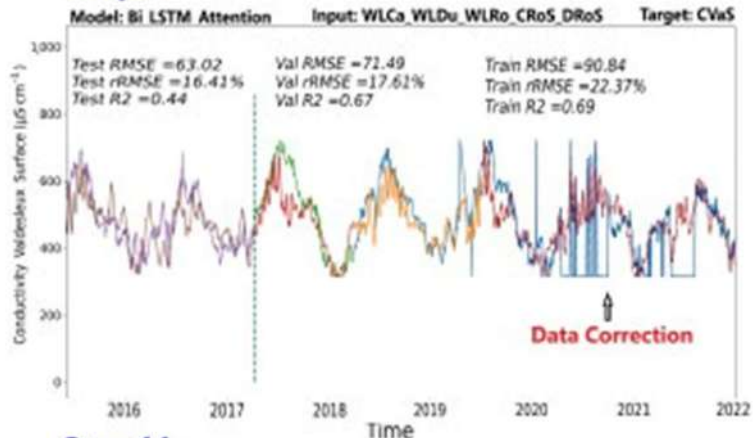
The results of evaluating the four proposed models for predicting water quality data with different input selections, including only water level data. For each model the rRMSE on the overall test dataset are presented. The combination of the water level WLCa\_WLDu\_WLRo emerges as the best input combination, in most of the case



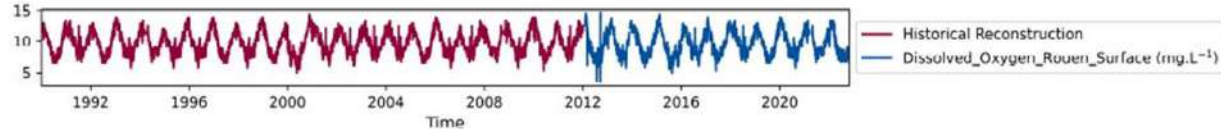
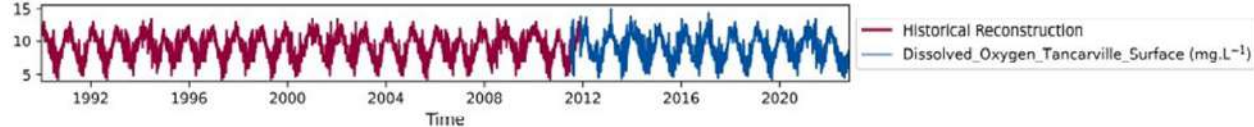
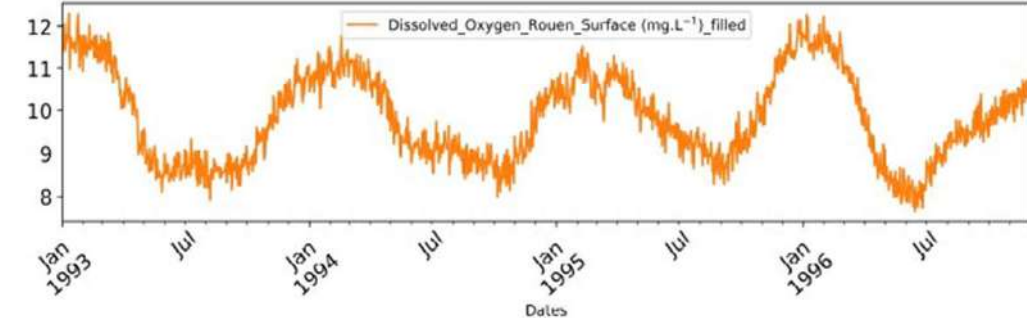
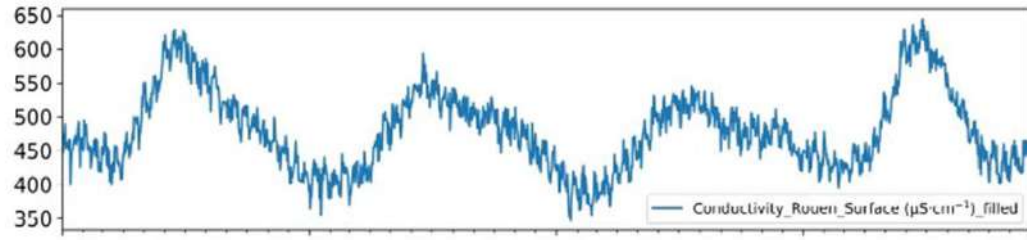
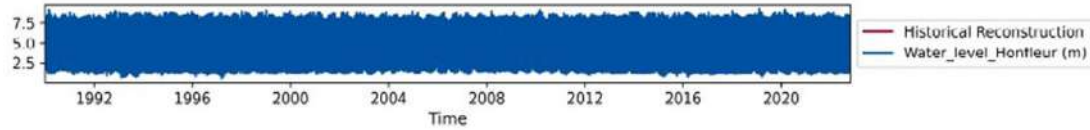
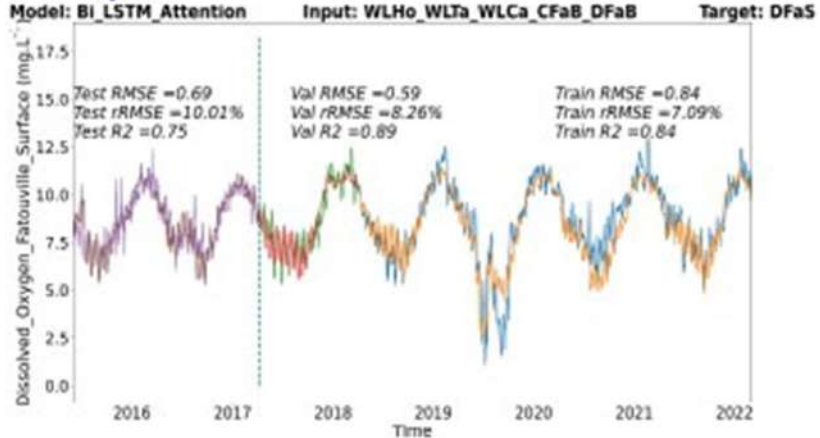
# Paramètres physico-chimiques (Janbain et al., 2023b)



## Step 8



## Step 11





## Lesson learned....

**Connaitre ses données :** comprendre les tendances, la saisonnalité, les problèmes d'acquisition (dérive de capteurs, donnée hors gamme....)

**Données manquantes:** Traiter les valeurs manquantes, supprimer les valeurs aberrantes et normaliser les données si nécessaire.

**Connaitre les forçages:** Divisez les données en ensembles de training et test => nécessite de prendre en compte les processus qui déterminent les variations des variables.

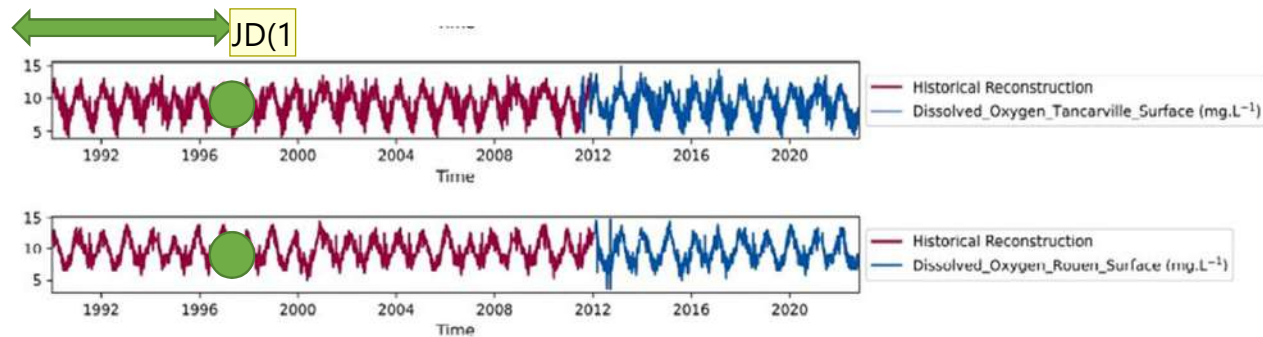
**Complexité à reproduire les extrêmes** => souvent les périodes que l'on cherche à simuler (e.g. inondation, hypoxie...)

Points de vigilance sur **l'interprétation:**

Plus les périodes de temps prédites (passé ou futur) sont longues et plus l'incertitude augmente

Par définition sans **bouleversement majeur/évolution** de l'environnement (complémentarité avec les modèles numériques)

Valide???



## Diapositive 17

---

JD(1

JULIEN DELOFFRE (Personnel); 05/06/2024

## 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
- ✓ Prédictions : possible à 'court terme' et possibilité de travailler sur des prédictions 'long terme' sur plusieurs années en utilisant les projections climatiques du GIEC.





## Références

Chidepudi et al., 2023. <https://doi.org/10.1016/j.scitotenv.2022.161035>

Chidepudi et al., 2024. <https://doi.org/10.1016/j.ejrh.2023.101632>

Janbain et al., 2023a <https://doi.org/10.1080/02626667.2023.2221791>

Janbain et al., 2023b <https://doi.org/10.3390/w15091773>

Vu et al., 2023a. <https://doi.org/10.1016/j.scitotenv.2023.163338>

Vu et al., 2023b. <https://doi.org/10.1016/j.scitotenv.2023.165494>

*Janbain et al., soumis. Deep Learning-Based Approaches for Extreme Flood Event Prediction: A Case Study in the Seine River Basin, France*

*Vu et al., soumis Nature. Human Impact: A Century of Altering Global Rivers.*