



REPUBLIQUE DU BENIN

MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE (MESRS)

UNIVERSITE D'ABOMEY-CALAVI (UAC)

INSTITUT NATIONAL DE L'EAU (INE)

DEPARTEMENT D'HYDROLOGIE ET GESTION DES RESSOURCES EN EAU
(DHGRE)

Filière : Hydro-Informatique

MEMOIRE DE FIN DE FORMATION POUR L'OBTENTION DU DIPLOME DE
MASTER PROFESSIONNEL

THÈME :

**MODÉLISATION DES INONDATIONS CÔTIÈRES EN
ZONE URBAINE À L'AIDE DE CONVLSTM : ÉTUDE
DE LA RÉGION D'AKPAKPA À COTONOU**

Rédigé par Vedel-Johann AÏSSI

Sous l'encadrement de :

Dr. Frédéric BONOU
Maître-Assistant des Universités
de CAMES

Sous la supervision de :

Professeur Zacharie SOHOU
Professeur Titulaire des
Universités du CAMES

Mémoire soutenu le 29 Décembre 2023 en face du jury composé de :

Président : Prof. Eugène EZIN
Examineurs : Prof. Zacharie SOHOU
: Dr. Audace DOSSOU-OLORY
Rapporteur : Dr. Frédéric BONOU



| 1^{ère} Promotion

Année académique : 2022-2023

FICHE DE CERTIFICATION

Nous soussignés :

- Président du jury : **Prof. Eugène EZIN**, Professeur Titulaire des Universités du CAMES, Directeur de l'Institut de Formation et de Recherche en Informatique
- Examineurs :
 - **Prof. Zacharie SOHOU**, Professeur Titulaire des Universités du CAMES, Directeur de l'Institut de Recherche Halieutique et Océanologique du Bénin ;
 - **Dr. Audace DOSSOU-OLORY**, Maître Assistant des Universités du CAMES, Chef Filière Hydro-Informatique
- Rapporteur : **Dr. Frédéric BONOU**, Maître-Assistant des Universités de CAMES, Enseignant Chercheur associé au Laboratoire d'Hydrologie Marine et Côtière, l'Institut de Recherche Halieutique et Océanologique du Bénin ;

Jury chargé de juger le mémoire de fin de formation pour l'obtention du diplôme Master Professionnel en **Hydro-Informatique** de l'étudiant **Vedel-Johann Rodrigue C. AÏSSI**, certifient que les différents apports, contributions et corrections énoncées au cours de la soutenance tenue le 29 Décembre 2023 de 11h à 12h30, ont été intégrés au document pour l'amélioration de la qualité scientifique et esthétique de ce dernier.

En foi de quoi, le présent certificat lui est délivré pour servir et valoir ce que de droit.

Président du jury

Prof. Eugène EZIN

Examineurs

Rapporteur

Prof. Zacharie SOHOU

Dr. Audace DOSSOU-OLORY

Dr. Frédéric BONOU

DEDICACE

Ce travail est dédié :

- A mon père **Johannès Baptista AÏSSI**, pour le soutien majeur qu'il ne cesse de m'accorder dans le but de faire de moi un homme meilleur, recevez l'expression de ma profonde gratitude et que la prospérité vous soit donnée par le Tout Puissant ;
- A ma très chère mère **Judith NASSARAH**, pour ses efforts, son soutien indéfectible, ses conseils et ses différentes prières, infiniment merci et daigne le Seigneur vous combler de la santé afin de bénéficier des fruits qui découleront de ce travail.

REMERCIEMENTS

Nous ne saurions rédiger ce rapport sans remercier dans un premier temps, Dieu pour toutes ses grâces, et pour tout ce qu'il fait pour nous au jour le jour.

Dans un second temps, nous tenons à exprimer nos vifs et sincères remerciements aux personnes ressources et aux organismes qui ont contribué de près ou de loin à la réalisation de ce travail.

Nous remercions les autorités de l'Institut National de l'Eau :

- **Professeur Daouda MAMA**, Professeur Titulaire des Universités / CAMES, Directeur de l'Institut National de l'Eau. Il est toujours présent quel que soit le stade de la formation et n'a jamais cessé de tout faire pour nous favoriser de meilleures conditions d'études. Que le Seigneur tout-puissant veille toujours sur sa bonne personne et protégera sa famille ;
- **Professeur Luc SINTONDI**, Professeur Titulaire des Universités / CAMES, Directeur Adjoint de l'Institut National de l'Eau
- **Professeur Flavien DOVONOU**, Maître de conférences des Universités du CAMES, Chef Département Hydrologie et Gestion des Ressources en Eau (HGIRE);
- **Professeur Abdoukarim ALLASANE**, Maître de conférences des Universités du CAMES, Coordonnateur des programmes de Master à l'Institut National de l'Eau ;
- **Dr Jean HOUNKPE**, Enseignant-Chercheur, Chef Filière Hydrologie Quantitative, Université d'Abomey-Calavi.
- **Dr Audace DOSSOU OLORY** Maître Assistant des Universités du CAMES, Chef Filière Hydro-Informatique, Université d'Abomey-Calavi.

Mes remerciements vont également à l'endroit de tout le personnel enseignant de l'INE en général et celui du département d'Hydrologie et Gestion des Ressources en Eau (DHGRE) en particulier pour la qualité de l'enseignement et leur dévouement à nous transmettre le savoir malgré les difficultés rencontrées ;

A l'endroit du superviseur de nos travaux de recherche et de mon maître mémoire, nous remercions :

- **Professeur Zacharie SOHOU**, Professeur Titulaire des Université du CAMES, Directeur de l'Institut de Recherche Halieutique et Océanologique du Bénin ;

- **Dr Frédéric BONOU**, Maître-Assistant des Universités de CAMES, pour l'encadrement reçu pendant la période des travaux et pour sa disponibilité ;

A toutes les autres personnes, pour l'aide apportée et les conseils reçus ;

- Tous ceux qui m'ont apporté leurs aides d'une manière ou d'une autre dans l'élaboration de ce rapport ;
- Toute ma famille, mes sœurs, en particulier **Vanessa AÏSSI, Valdina AÏSSI**, qui n'ont ménagé aucun effort pour la réussite de ma formation en me soutenant sur tous les plans et en me prodiguant tous les conseils nécessaires malgré les moments difficiles ;
- Tous mes camarades de cette promotion et amis spécialement **Patricia HOUSSIONON, Sergio AGOSSOU, Freddy HOUNONDAHO, NDAYISABA Jean-Marie, IRAKIZA Gisele, Fulgence GBETABLE** ;
- Et à tous ceux que je n'ai pas eu le temps de citer, je vous remercie beaucoup pour tout le soutien.

RESUME

Les inondations côtières sont un phénomène naturel qui menacent de plus en plus les zones urbaines situées sur le littoral, en raison des changements climatiques et de la pression démographique. Dans ce contexte, notre étude intitulée : « **Modélisation des inondations côtières en zone urbaine à l'aide de ConvLSTM : étude de la région d'Akpakpa à Cotonou** » vise à concevoir un modèle de prévision des inondations côtières en zone urbaine à l'aide de ConvLSTM, qui est un modèle basé sur les réseaux de neurones récurrents et convolutifs, capables de traiter des données séquentielles spatiales.

Nous nous sommes intéressés à la région d'Akpakpa à Cotonou, située sur le littoral du golfe de Guinée, et nous avons utilisé des données journalières sur la période 1997-2021. Pour modéliser et prévoir les inondations côtières, nous avons utilisé un indice d'inondation côtière (IF) comme variable de sortie, qui représente le risque d'inondation. Nous avons utilisé quatre paramètres d'entrée : la précipitation, l'indice d'inondation côtière IF, la marée, et la SSH. Nous avons simulé quatre scénarios de prévision à 2, 4, 8 et 16 jours, et nous avons évalué la qualité de notre modèle avec des indicateurs de performance liés au modèle ConvLSTM, comme le coefficient d'efficacité de Nash-Sutcliffe (NSE), le coefficient de détermination (R^2), l'erreur quadratique moyenne (MSE), l'erreur absolue moyenne (MAE), le coefficient d'efficacité Legates et McCabe (LME), la racine carrée de la moyenne des erreurs quadratiques (RMSE) et le coefficient de corrélation (R). Nos résultats montrent que plus le modèle est capable de prédire l'IF avec une faible perte, plus il est fiable pour anticiper le risque d'inondation et précis. Et sur la base des indicateurs de performance liés au modèle en régression, qui indiquent que le modèle est capable de prédire jusqu'au 16^{ème} jours.

MOTS CLES : inondations côtières, ConvLSTM, indice d'inondation côtière, indicateurs de performance

ABSTRACT

Coastal flooding is a natural phenomenon that increasingly threatens urban areas located on the coast, due to climate change and demographic pressure. In this context, our study entitled: “**Modeling coastal flooding in urban areas using ConvLSTM: a case study of the Akpakpa region in Cotonou**” aims to design a forecasting model for coastal flooding in urban areas using ConvLSTM, which is a model based on recurrent and convolutional neural networks, capable of processing spatial sequential data.

We focused on the Akpakpa region in Cotonou, located on the coast of the Gulf of Guinea, and we used daily data over the period 1997-2021. To model and forecast coastal flooding, we used a coastal flooding index (FI) as an output variable, which represents the risk of flooding. We used four input parameters: precipitation, FI, tide, and SSH. We simulated four forecasting scenarios at 2, 4, 8 and 16 days, and we evaluated the quality of our model with performance indicators related to the regression model, such as NSE, R², MSE, MAE, LME, RMSE and R. Our results show that the more the model is able to predict the CFI with low loss, the more reliable and accurate it is to anticipate the risk of flooding. And based on the performance indicators related to the ConvLSTM model, which indicate that the model is able to predict up to the 16th day.

KEYWORDS: coastal flooding, ConvLSTM, coastal flooding index, performance indicators

SOMMAIRE

FICHE DE CERTIFICATION	ii
DEDICACE.....	iii
REMERCIEMENTS.....	iv
RESUME.....	vi
ABSTRACT	vii
SOMMAIRE	viii
LISTE DES TABLEAUX	ix
LISTE DES FIGURES.....	x
LISTE DES IMAGES	xi
LISTE DES SIGLES ET ABREVIATIONS.....	xii
INTRODUCTION	1
CHAPITRE I : REVUE DE LITTERATURE	4
I.1 Clarification de certains concepts	4
I.2 Méthodes de modélisations existantes : Etat des connaissances.....	6
CHAPITRE II : DEMARCHE METHODOLOGIQUE	10
II.1 Cadre d'étude	10
II.2 Matériel et approche méthodologique.....	16
CHAPITRE III : RESULTATS ET DISCUSSIONS	43
III.1 Résultats.....	43
III.2 Discussions.....	50
CONCLUSION.....	57
REFERENCES BIBLIOGRAPHIQUES	58
TABLE DES MATIERES	60

LISTE DES TABLEAUX

<i>Tableau 1: Dataframe contenant les données d'entrée pour le modèle.....</i>	<i>21</i>
<i>Tableau 2: Récapitulatif indicateurs de performance et formules.....</i>	<i>28</i>
<i>Tableau 3: Seuils de risques d'inondation.....</i>	<i>45</i>
<i>Tableau 4: Message obtenu par horizon de prévision</i>	<i>45</i>
<i>Tableau 5: Récapitulatif des données des indicateurs de performances par horizon de prévision.....</i>	<i>48</i>
<i>Tableau 6: Différences observées des paramètres optimaux</i>	<i>51</i>
<i>Tableau 7: Avantages et inconvénients entre les méthodes utilisées</i>	<i>53</i>

LISTE DES FIGURES

Figure 1: Carte de localisation de la commune de Cotonou	13
Figure 2: Architecture du modèle	22
Figure 3: Processus du déroulement de la séquentialisation (taille de fenêtre : 2)	25
Figure 4- a: Prévission IF prédit Jour 2	44
Figure 4- b: Prévission IF prédit Jour 4	44
Figure 4- c: Prévission IF prédit Jour 8	44
Figure 4- d: Prévission IF prédit Jour 16	44
Figure 5- a : Perte au J2	46
Figure 5- b: Perte au J4	46
Figure 5- c: Perte au J8	46
Figure 5- d: Perte au J16	46
Figure 6- a :Dispersion des résidus au J2	47
Figure 6- b: Dispersion des résidus au J4	47
Figure 6- c: Dispersion des résidus au J8	47
Figure 6- d: Dispersion des résidus au J16	47
Figure 7: Indicateurs de performance pour différents horizons	49

LISTE DES IMAGES

<i>Image 1: Localisation de la commune de Cotonou avec Google Earth</i>	<i>19</i>
<i>Image 2: Page de connexion de l'application</i>	<i>31</i>
<i>Image 3: Page d'accueil de l'application</i>	<i>32</i>
<i>Image 4: Initialisation du nombre de paramètres.....</i>	<i>33</i>
<i>Image 5: Sélection des fichiers en fonction du nombre de paramètres fixés par l'utilisateur</i>	<i>33</i>
<i>Image 6: Affichage de la liste des données</i>	<i>34</i>
<i>Image 7: Traitement des données</i>	<i>35</i>
<i>Image 8: Affichage de l'ensemble des données combinées.....</i>	<i>35</i>
<i>Image 9: Page du Calcul de l'IF.....</i>	<i>36</i>
<i>Image 10: Résultat du calcul de l'IF sur l'application</i>	<i>37</i>
<i>Image 11: Prédiction avec le modèle ConvLSTM sur l'application.....</i>	<i>38</i>
<i>Image 12: Pendant le processus de prédiction</i>	<i>38</i>
<i>Image 13: Affichage de l'ensemble des graphiques et des indicateurs de performance.....</i>	<i>39</i>
<i>Image 14: Page Copernicus Data sur l'application.....</i>	<i>40</i>
<i>Image 15: Quelques paramètres figurant dans les services CMEMS</i>	<i>40</i>

LISTE DES SIGLES ET ABREVIATIONS

API	:	Application P rogramming I nterface
C2EA	:	C entre d' Excellence d' Afrique pour l' E au et l' A ssainissement
CMEMS	:	C opernicus M arine E nvironment M onitoring S ervice
ConvLSTM	:	C onvolutional L ong S hort- T erm M emory
CSV	:	C omma S eparated V alues
D	:	C oefficient de W illmott
DP	:	D eep L earning
ECMWF	:	E uropean C entre for M edium- R ange W eather F orecasts
IFRI	:	I nstitut de F ormation et de R echerche en I nformatique
IGN	:	I nstitut G éographique N ational
INE	:	I nstitut N ational de l' E au
INSAE	:	I nstitut N ational de la S tatistique et de la D émographie
IRHOB	:	I nstitut de R echerche H alieutique et O céanologique du B énin
LME	:	L egates- M cCabe E fficiency
MAE	:	M ean A bsolute E rror
ML	:	M achine L earning
MSE	:	M ean S quared E rror
NSE	:	N ash- S utcliffe E fficiency
PE	:	P récipitation E ffective moyenne
PMC	:	P erceptron M ultiple C ouche
R	:	C oefficient de C orrélation de P earson
R2	:	C oefficient de D étermination
RGPH	:	R ecensement G énéral de la P opulation et de l' H abitat
RMSE	:	R oot M ean S quared E rror
RNA	:	R éseaux d e N eurones A rtificiels
SSH	:	S ea S urface H eigh
UAC	:	U niversité d' A bomey C alavi

INTRODUCTION

L'une des conséquences directes de l'effet de serre est le réchauffement climatique avec la fonte rapide des glaciers dans l'antarctique. Presque tous les pays ayant une bordure maritime subissent de plein fouet des inondations à cause de la montée du niveau des mers des océans. Notre pays, le Bénin n'en est pas épargné. La région d'Akpakpa à Cotonou, située sur le littoral du golfe de Guinée, subit un double risque d'inondations : celle côtière (dus à la montée du niveau de la mer, aux marées, aux tempêtes, aux surcotes, etc), et les inondations pluviales, (dus à l'augmentation de la fréquence et de l'intensité des précipitations). Ces phénomènes peuvent se combiner et aggraver les conséquences des inondations sur les zones urbaines, comme l'érosion du littoral, la salinisation des sols et des nappes phréatiques, la contamination des sources d'eau potable, la dégradation des écosystèmes côtiers, etc. Ces impacts peuvent affecter la santé, la sécurité, le bien-être et la résilience des populations urbaines. Comment prévoir ce phénomène d'inondation côtière ? C'est à cela que s'intéresse ce présent travail. Pour faire face à ce défi, il est nécessaire de développer des stratégies d'adaptation efficaces pour réduire les vulnérabilités des populations urbaines ; c'est la région d'Akpakpa à Cotonou qui a été choisie pour servir de zone d'étude. Parmi ces stratégies d'adaptation utilisées, on peut citer l'établissement d'un système d'alerte précoce basé sur la modélisation et la prévision des inondations côtières, ainsi que la mise en place de plans d'urgence et de procédures visant à minimiser les pertes en vies humaines, les dommages sur les infrastructures, ainsi que les perturbations des activités économiques de la ville. C'est dans cette perspective que notre question de recherche se situe : Comment modéliser et prévoir les inondations côtières en zone urbaine et ces conséquences, sur la région d'Akpakpa à Cotonou ? Pour répondre à cette question, nous avons choisi comme thème de recherche : **Modélisation des inondations côtières en zone urbaine a l'aide de ConvLSTM : étude de la région d'Akpakpa à Cotonou.**

Objectif général

L'objectif général est de modéliser des inondations côtières en zone urbaine à l'aide de ConvLSTM : étude de la région d'Akpakpa à Cotonou.

Objectifs spécifiques

De façon spécifique, il s'agit :

- collecter et analyser les données climatiques, hydrologiques, et océanographiques pertinentes pour la région d'Akpakpa à Cotonou, en incluant les données de précipitations, des marées et des hauteurs du niveau de la mer ;
- concevoir et développer un système d'alerte précoce à partir du modèle ConvLSTM capable de prédire avec précision les inondations côtières à partir des données collectées sur une horizon de 4 jours ;
- évaluer et comparer les performances du modèle avec les méthodes de modélisation traditionnelles en utilisant des métriques appropriées telles que : le coefficient de détermination (R^2), l'erreur quadratique moyenne (MSE), l'erreur absolue moyenne (MAE), le coefficient de Nash-Sutcliffe (NSE), le coefficient de Legates et McCabe (LME), le coefficient de Willmott (d), la racine carrée de la moyenne des erreurs quadratiques (RMSE), le coefficient de corrélation de Pearson (R) ;
- mettre en place une application web pour des utilisateurs potentiels afin de tester le modèle construit ;

Compte tenu des objectifs ci-dessus cités, après avoir fait une revue littéraire, nous effectuerons les travaux à travers une démarche méthodologique qui nous emmèneront à un ensemble de résultats autour desquels s'établiront des discussions.

Chapitre I

CADRE DE L'ETUDE

CHAPITRE I : REVUE DE LITTERATURE

Ce chapitre présente la revue de littérature sur les concepts et méthodes liés à la modélisation des inondations côtières en zone urbaine. Nous commencerons par clarifier certains termes clés. Nous présenterons ensuite l'état des connaissances sur les différentes approches de modélisation hydrologique et des inondations côtières, en mettant en évidence les avantages et les limites de chacune. Nous nous appuierons sur des sources variées, telles que des articles scientifiques. Enfin, nous chercherons à identifier les lacunes dans ce domaine, afin de justifier le choix de notre méthode basée sur le ConvLSTM.

I.1 Clarification de certains concepts

Dans le but de mieux comprendre les contours du sujet, quelques termes se doivent d'être clarifiés.

- **Hydroinformatique** : est une discipline qui combine la modélisation informatique de l'eau avec les technologies de l'information et de l'informatique pour l'ingénierie, la gestion et l'aide à la décision de l'eau (*Michael Abbott, 2022*).
- **Inondations côtières en zone urbaine** : un type d'inondation qui se produit le long des côtes, généralement lorsque l'eau de mer envahit une région densément peuplée comprenant des bâtiments, des infrastructures, et une population importante. Elle est généralement le résultat d'une combinaison de marées maritimes, de vents violents et de pression barométrique. Dans un journal d'article rédigé par *Déborah Idier et al. en 2020*, les inondations côtières font référence à l'inondation des zones côtières en raison d'une combinaison de facteurs tels que l'élévation du niveau de la mer, les marées, les ondes de tempête et les conditions de vagues.
- **Modélisation hydrologique** : permet en effet de simuler le risque d'inondation futur dans une zone exposée et d'adapter le développement immobilier en conséquence. Elles fournissent des projections sur la fréquence des débordements et permettent, ce faisant, d'intégrer le risque d'inondation dans le dessin du bâti, d'orienter la surélévation des constructions et les mesures de protection (tels des digues et dragages de rivières) en fonction (*Gralepois et al., 2015*). Ces outils sont mobilisés pour aiguiller le développement de projets en zone inondable, dans un contexte où le foncier disponible est rare (*Haghe, 2010 ; Moulin et al, 2013*). De plus de façon spécifique, elle s'applique dans la prédiction des inondations dans les zones côtières.

- **Modélisation des inondations côtières** : consiste à utiliser des méthodes mathématiques et informatiques pour simuler et prédire les inondations dans les zones côtières. Ces modèles prennent en compte différents facteurs tels que les conditions météorologiques, les niveaux d'eau, les marées, la topographie et l'utilisation des terres pour prédire l'étendue et la profondeur des inondations (*Smith et al., 2012*)
Selon *Wahl et al. (2015)*, il existe plusieurs méthodes pour modéliser les inondations côtières, notamment l'utilisation de modèles statistiques pour étudier la dépendance entre les facteurs d'inondation ou la probabilité conjointe, et les modèles numériques pour simuler l'inondation des inondations côtières.
- **Marées** : phénomènes naturels, résultent des interactions gravitationnelles entre la lune et le soleil avec les océans. Elles se traduisent par des oscillations périodiques de la surface de la mer, alternant entre les phases de marée haute, où la mer atteint son point culminant, et de marée basse, où elle atteint son point le plus bas. L'écart vertical entre ces deux niveaux d'eau est appelé amplitude de marée. Les marées sont également accompagnées de courants de marée, des mouvements horizontaux de l'eau qui se produisent en conjonction avec le cycle de montée et de descente de la marée. (*River and tides, 2008*)
- **Pression barométrique** : la pression barométrique est la mesure du poids de l'atmosphère qui exerce une pression sur la surface de la Terre. Elle est aussi nommée pression relative, car elle prend comme point zéro la pression atmosphérique. La pression barométrique varie en fonction de l'altitude, de la météo et de l'attraction gravitationnelle sur les gaz dans l'atmosphère (*Yakaranda Magazine, 2021*).
- **ConvLSTM** : est une méthode d'apprentissage profond qui combine des réseaux de neurones convolutifs (CNN) et des réseaux de mémoire à long et court terme (LSTM). Le ConvLSTM est utilisé pour traiter les données de séries chronologiques et réaliser une reconstruction secondaire des vecteurs de caractéristiques, ce qui permet d'obtenir une grande précision de prédiction (*Shi et al., 2015*).
- **Fonction d'activation** : elle joue un rôle central dans un réseau de neurones car elle permet d'appliquer une transformation non linéaire aux données qui traversent le modèle et permet de modifier leur représentation spatiale. Néanmoins, seulement une

petite fraction d'entre elles est utilisée comme fonction d'activation en apprentissage profond. La différence entre ces fonctions d'activation peut parfois sembler négligeable. Cependant, certaines de ces fonctions doivent être choisies avec précision si l'on souhaite créer un réseau de neurones fonctionnel. Nous pouvons citer : ***ReLU***, ***Sigmoid***, ***Softmax***, ***Softplus***, ***Softsign***, ***Tanh***, ***Elu***, ***Selu***

- **Epoch** : la fonction d'époque est une notion utilisée dans les réseaux de neurones artificiels pour désigner le nombre de fois que le réseau parcourt l'ensemble des données d'apprentissage. Une époque correspond à un cycle complet d'apprentissage, où le réseau ajuste ses paramètres (poids et biais) en fonction de l'erreur entre la sortie désirée et la sortie obtenue.
- **Fonction d'optimisation** : dans le même article (*DataScientest, 2022*), la fonction d'optimisation dans les réseaux de neurones est une procédure qui vise à ajuster les paramètres du réseau (poids et biais) afin de minimiser une fonction de coût qui mesure l'erreur entre la sortie du réseau et la sortie désirée. La fonction d'optimisation permet au réseau de neurones d'apprendre à partir des données et d'améliorer ses performances. Il existe plusieurs méthodes d'optimisation, comme la descente de gradient, la rétropropagation, les algorithmes génétiques.
- **Perte d'entraînement d'un modèle** : la perte d'entraînement d'un modèle est une mesure qui indique à quel point le modèle est capable de prédire correctement les données qu'il n'a pas encore vues. Plus la perte est faible, plus le modèle est performant. La perte d'entraînement d'un modèle ConvLSTM dépend de la fonction de perte que l'on choisit pour entraîner le réseau. Il existe différentes fonctions de perte possibles, comme la perte quadratique moyenne (MSE), la perte logarithmique (log loss), la perte de Kullback-Leibler (KL divergence), (*QuinnRadich et al., 2023*).

I.2 Méthodes de modélisations existantes : Etat des connaissances

La présente partie a pour objectif de situer l'état des connaissances sur les méthodes de modélisation à partir de l'apprentissage automatique dans notre contexte d'inondation.

- (*Campolo, et al., 2020*) ont développés un modèle de prévision des crues fluviales de réseau neuronal illustrant des résultats prometteurs à de courtes échelles de temps.

Cependant, une diminution rapide du rendement des prévisions était évidente avec un horizon temporel plus long.

- Une autre approche de prévision des inondations à court terme a été présentée par (Nayak *et al.*, 2021) : utilisation de la technique neuro-floue. Les résultats ont illustré la viabilité de leurs modèles pour la prévision à court terme du débit des rivières.
- (Han *et al.*, 2022) dans l'application de la machine à vecteurs de support (SVM) pour la prévision des inondations ont mentionné que, bien que leur modèle objectif ait donné de meilleurs résultats que les modèles de référence, il a fallu déployer des efforts considérables pour assurer une meilleure performance du modèle objectif.
- (Sit *et Demir*, 2023) ont exploré l'utilisation de réseaux neuronaux profonds artificiels pour la prévision des inondations et ont mentionné l'utilité des réseaux neuronaux pour la prévision des inondations à l'aide de données de séries chronologiques.

Ces exemples montrent que l'apprentissage automatique et l'apprentissage profond sont des méthodes de modélisation innovantes et performantes pour la prévision des inondations à différentes échelles de temps. Ces méthodes permettent de traiter des données complexes et hétérogènes, de capturer des relations non linéaires et de s'adapter aux changements environnementaux.

Cependant, ces méthodes présentent aussi des limites, comme la nécessité de disposer de données suffisantes et de qualité, le manque d'interprétabilité des modèles ou la difficulté de validation des résultats.

Les approches présentées jusqu'à présent ont utilisé l'apprentissage automatique conventionnel pour prévoir les situations d'inondation. Une étude de (Tran *et Song*, 2023) a toutefois utilisé des algorithmes d'apprentissage profond, c'est-à-dire RNN et LSTM, pour prévoir les niveaux d'eau comme moyen pratique de développer une solution pour la prévision des inondations dans les zones urbaines. Leurs résultats ont indiqué que tous les modèles d'apprentissage profond avaient une grande précision. Par conséquent, dans cet article, des approches hybrides d'apprentissage profond sont utilisées pour prévoir les inondations à des échelles de temps courtes et longues, en s'attendant à ce que ces modèles nouvellement développés constituent un pas en avant dans les systèmes d'alerte précoce aux inondations basés sur des données. (Mohammed Moishin *et al.*, 2021).

Ainsi, nous pouvons dire que les algorithmes d'apprentissage profond sont des outils performants et innovants pour la prévision des inondations, qui peuvent s'adapter à différents

contextes, données et objectifs. L'approche hybride d'apprentissage profond, qui combinent les RNN et les LSTM avec d'autres techniques, comme les machines à vecteurs de support, sont prometteuses pour améliorer la précision et la robustesse des prévisions, et pour réduire les incertitudes associées.

Nous nous proposons donc d'explorer l'apport d'une approche hybride d'apprentissage profond, qui combine les RNN et les LSTM pour prévoir les inondations côtières à Cotonou. Nous faisons l'hypothèse que cette approche permettra de réaliser une prévision plus précise du risque d'inondation, en tenant compte des données météorologiques, hydrologiques et océanographiques.

Chapitre II

DEMARCHE METHODOLOGIQUE

CHAPITRE II : DEMARCHE METHODOLOGIQUE

Ce chapitre présente la démarche méthodologique adoptée pour étudier la dynamique spatio-temporelle des inondations côtières à Cotonou, en utilisant le modèle ConvLSTM. Nous commencerons par décrire le cadre d'étude, en situant la commune de Cotonou sur le plan géographique, administratif et physique. Nous présenterons ensuite le matériel et les sources de données utilisés pour alimenter le modèle. Nous expliquerons enfin les différentes étapes de la modélisation et de la simulation du risque d'inondation, ainsi que la mise en place d'une interface pour visualiser les résultats.

II.1 Cadre d'étude

2.1 Localisation de la zone d'étude : COMMUNE DE COTONOU

La République du Bénin est entièrement située dans la zone intertropicale de l'Afrique de l'Ouest, entre 6° 10' et 12° 25' de latitude Nord et 0° 45' et 3° 55' de longitude est. Elle couvre une superficie de 114,763 km² et est limitée au nord par le fleuve et la République du Niger, au nord-ouest par le Burkina Faso, au sud par l'océan Atlantique, à l'est par la République fédérale du Nigeria et à l'Ouest par le Togo. Elle abrite en 2013 une population de 10 008 749 habitants avec une densité de population de 87,2 hab/Km² (INSAE, RGPH4, 2013). Selon (Global Digital Reports, 2023) dans l'article NUMÉRIQUE 2023 : BÉNIN, La population totale du Bénin était de 13,53 millions d'habitants en janvier 2023.

Du point de vue découpage administratif, le pays compte 12 départements : Alibori, Atacora, Atlantique, Borgou, Collines, Couffo, Donga, Littoral, Mono, Ouémé, Plateau et Atlantique.

Le département du Littoral concerné par l'étude compte 01 commune : Cotonou. Elle est subdivisée en 13 arrondissements qui sont composés de 143 quartiers de ville.

L'étude s'est déroulée dans la région d'Akpakpa au niveau des côtes, dans la commune de Cotonou, située sur le cordon littoral entre le Lac Nokoué et l'Océan Atlantique dans le département du Littoral à l'Est de la frontière du Togo. Sa position géographique est de 6°21'55'' de latitude Nord, 2°25'05'' de longitude Est et de l'altitude par rapport au niveau de la mer : 7 m. Sa Superficie est de 79 km².

Elle est limitée :

- au Nord par le lac Nokoué;

- au Sud par l'Océan Atlantique ;
- à l'Ouest par les Communes de Ouidah et d'Abomey-Calavi;
- à l'Est par la commune de Sèmè-Podji

2.2 Cadre administratif et physique

Au plan administratif, la commune de Cotonou comporte 13 arrondissements dont 143 quartiers de ville à savoir :

- 1^{er} Arrondissement (10 quartiers) :
Avotrou, Finagnon, Dandji, Donatin, N'vènamèdé, Suru-Léré, Tanto, Tokplegbe, Tchanhoukpamè, Yagbé
- 2^{ème} Arrondissements (12 quartiers) :
Ahouassa, Djèdjèlayé, Gankodo, Irédé, Lom'nava, Minontchou, Sènadé 1 et 2, Kowégbo, Kpondéhou 1 et 2, Yénawa
- 3^{ème} Arrondissements (13 quartiers) :
Adjégoulè, Adogléta, Agbato, Agbodjedo, Ayélawadjè 1 et 2, Fifatin, Hlakonmè, Hounonkpo, Kpankpan, Midombo, Sègbéya nord, Sègbéya sud
- 4^{ème} Arrondissements (11 quartiers) :
Abokicodji centre, Abokicodji lagune, Dédokpo, Enagnon, Fifadji Houto, Gbèdjèwin, Missessin, Ohè, Sodjèatinmè Centre, Sodjèatinmè Est, Sodjèatinmè Ouest
- 5^{ème} Arrondissements (15 quartiers) :
Bocossi Tokpa, Dota, Gbedokpo, Gbéto, Guinkomè, Joncquet, Missèbo, Missité, Mifongu, Nouveau pont, Tokpa hoho, Wlacodji plage, Wlacodji Kpodji, Zongo éhuzu, Zongo Nima
- 6^{ème} Arrondissements (18 quartiers) :
Aïdjèdo 1, 2, 3, 4, Ahouansori Agata, Ahouansori Agué, Ahouansori Toweta 1 et 2, Dantokpa, Djidjè 1 et 2, Gbèdromèdé 1 et 2, Hindé 1 et 2, Jericho, Ladji, Vossa
- 7^{ème} Arrondissements (12 quartiers) :
Dagbédji, Enagnon, Fignon, Gbenan, Gbewa (Batito), Gbedomidji (Maro militaire), Missité, Sedami, Sedjro, Sehogan, Todoté, Yevedo
- 8^{ème} Arrondissements (8 quartiers) :
Agbodjedo, Agontikon, Gbédagba, Houéhoun, Houénoussou, Médedjro, Tonato, Minonkpo

- 9^{ème} Arrondissements (4 quartiers) :
Fifadji, Vossakpodji, Zogbo, Zogbohoulé
- 10^{ème} Arrondissements (7 quartiers) :
Gbènonkpo, Kouhounou, Midédji, Missékplé, Missogbé, Vèdoko, Yénawa
- 11^{ème} Arrondissements (13 quartiers) :
Alobatin, Ayidoté, Finagnon, Gbediga 1 et 2, Gbégamey 1, 2, 3, 4, Houéyiho 1 et 2, Saint Jean, Vodjè centre.
- 12^{ème} Arrondissements (13 quartiers) :
Aibatin 1, Cadjèhoun 1,2,3,4, Cadjèhoun Kpota, Vodjè Kpota, Fidjrossè centre, Fidjrossè kpota, Fiyegnon 1 et 2, Cocotier, Ahouanléko
- 13^{ème} Arrondissements (6 quartiers) :
Agla, Aibatin 2, Ahogbohoulé , Gbedegbè, Houénoussou, Missité

(Lionel, 2019).

De tous ces arrondissements, ceux qui requièrent notre attention en ce qui concerne le cadre de ce travail sont dans le 1^{er}, 2^{ème}, 3^{ème} et 6^{ème} arrondissements situés au sud où certains quartiers se trouvent au niveau des côtes de la commune de Cotonou.

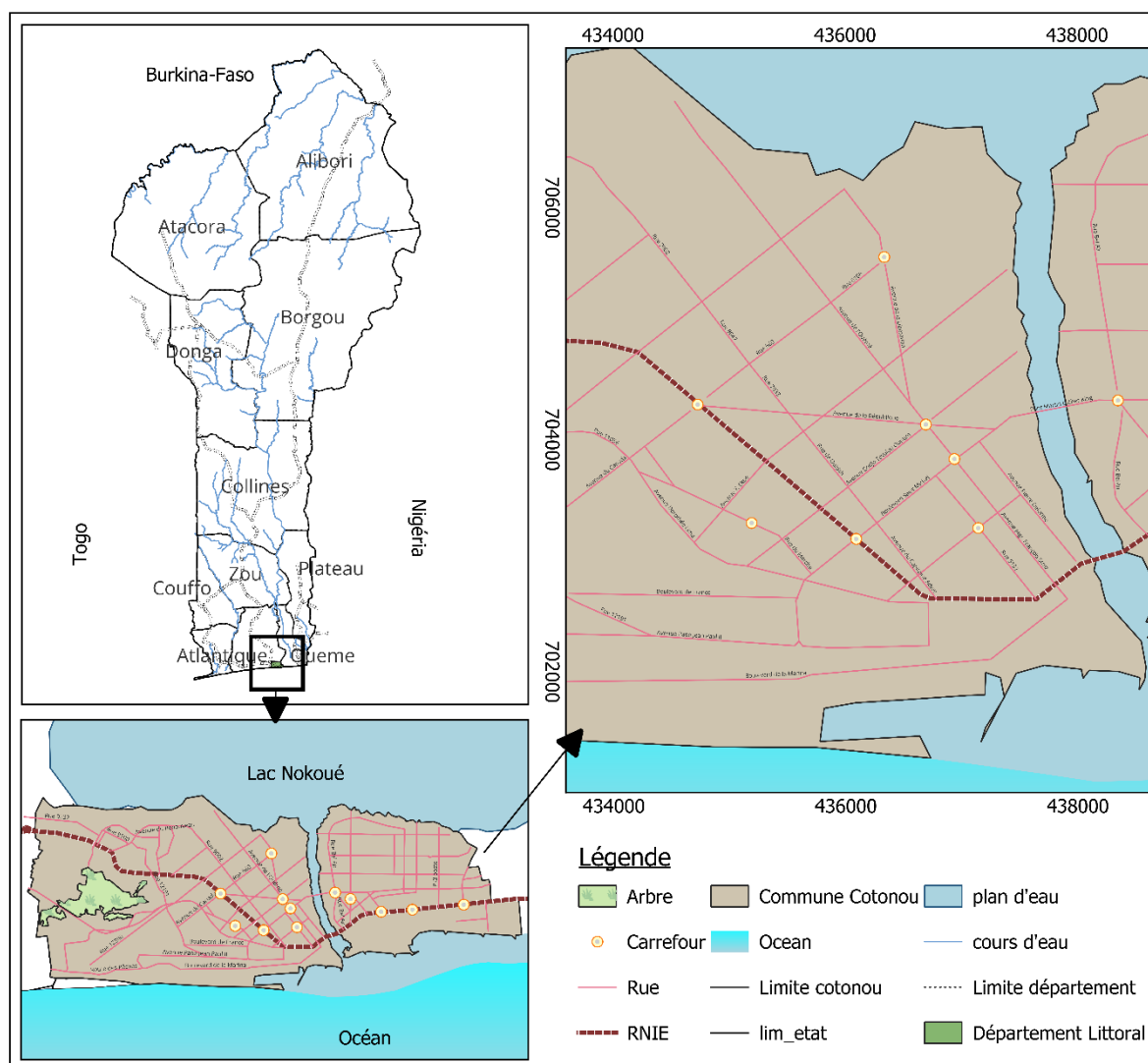


Figure 1: Carte de localisation de la commune de Cotonou

La figure ci-dessus montre la localisation de la zone d'étude au sein de la commune de Cotonou, située dans le département du Littoral, au Bénin. La zone d'étude se situe dans le quartier d'Akpakpa, contenant le chenal qui lie le lac Nokoué et l'océan. Elle a été choisie compte tenu de sa position vis-à-vis de la côte et des inondations côtières, d'après les données historiques.

2.2.1 Cadre administratif

Après la refondation de l'Etat béninois devenu depuis 2002 un Etat décentralisé, la commune de Cotonou est subdivisée en 13 arrondissements composés de 144 quartiers.

Le Maire est à la fois Président du Conseil Municipal (Assemblée Locale) et Premier Responsable de l'administration municipale. Il est assisté de trois Adjoints dans la gestion des affaires locales, tandis qu'au niveau des arrondissements, ce sont les chefs d'arrondissement qui le représentent.

La Municipalité de Cotonou est administrée par deux grands blocs :

- Le cabinet du Maire qui s'occupe exclusivement de la vie politique ;
- Le secrétariat général qui coordonne tous les services des directions centrales de la Maire.

Constitutionnellement, la ville de Porto-Novo distante de 30 km de Cotonou, est et demeure la Capitale politique du Bénin. Nonobstant, Cotonou abrite les principaux centres de décision pour l'administration du pays : la Présidence de la République, la quasi-totalité des départements ministériels, la Cour Constitutionnelle, la Haute Autorité de l'Audiovisuel et de la Communication (HAAC), le Conseil Économique et Social (CES), la Cour Suprême etc.

Sont également domiciliées à Cotonou, la totalité des représentations diplomatiques de pays étrangers et d'institutions internationales près le Bénin et la direction de grands projets socio-économiques. (*Monographie de Cotonou, Afrique Conseil, 2006*)

En somme, le Bénin est administré et dirigé depuis la commune de Cotonou.

2.2.2 Milieu physique

La commune de Cotonou est située sur le cordon littoral qui s'étend entre le lac Nokoué et l'Océan Atlantique, constitué de sables alluviaux d'environ cinq mètres de hauteur maximale. Elle représente la seule commune du département du Littoral et est limitée au nord par la commune de Sô-Ava et le lac Nokoué, au Sud par l'Océan Atlantique, à l'Est par la commune de Sèmè-Kpodji et à l'Ouest par celle d'Abomey-Calavi. Elle couvre une superficie de 79 km²,

dont 70% sont situés à l'Ouest du chenal. Les quartiers de l'Est sont reliés à la partie Ouest par trois ponts.

A l'Ouest de Cotonou, se trouvent le Port Autonome et l'Aéroport International qui font de la ville, la plus importante porte d'entrée et de sortie du Bénin, tandis que l'Est dispose d'une vaste zone industrielle.

Concernant le climat qu'offre le milieu d'étude s'inscrit dans un climat de type sub-équatorial avec une alternance de deux saisons pluvieuses (Avril à Juillet et Septembre à Novembre) et de deux saisons sèches (Décembre à Mars et Août). Pendant la grande saison des pluies, la ville menacée par de graves inondations, offrant ainsi aux cotonnois le spectacle d'un gros village lacustre (niveau bas fortement influencé par les variations du niveau des plans d'eau; niveau maximal des crues : 1,5 m (IGN). La pluviométrie varie entre 900 et 1200 mm, alors que la température moyenne est de 27 °C environ. Le vent le plus remarquable dans la commune est l'harmattan qui se manifeste généralement courant Novembre à Décembre.

En ce qui concerne la topographie, elle se situe dans la plaine côtière, possède des sols sableux qui sont généralement pauvres en matière organique avec une faible capacité d'échange et un faible pouvoir de rétention en eau, ce qui a pour corollaire les inondations répétées observées çà et là dans la ville de Cotonou pendant la saison pluvieuse.

La commune de Cotonou a un relief peu accidenté avec des marécages. Il a deux caractéristiques principales : des dépressions longitudinales parallèles à la côte et des bas-fonds érodés par l'écoulement des eaux pluviales qui communiquent avec le lac. Le site est coupé en deux par le chenal appelé « Lagune de Cotonou », communication directe entre le lac et la mer, creusée par les Français depuis 1894. Les épis du port ont contribué à l'érosion de toute la côte Est de la ville. Cette érosion s'opère à une vitesse moyenne de 16,8 mètres par an dans la critique.

Le réseau hydrographique à Cotonou est alimenté par trois principaux plans d'eau : l'Océan Atlantique, la Lagune et le Lac Nokoué, pourvoyeurs de ressources halieutiques non négligeables.

La commune de Cotonou dispose d'un certain nombre de formations végétales bien réparties :

- en bordure de la côte, les sables du cordon littoral sont couverts de plantations de cocotiers;

- une zone à végétation rare et clairsemée formée essentiellement d'halophytes sur le cordon littoral.

Concernant la géologie de la commune, la nappe phréatique se trouve à proximité de la surface du sol dont la perméabilité élevée accélère l'infiltration des eaux l'infiltration des eaux pluviales et usées ce qui pourrait générer des risques de pollution. (Monographie de Cotonou, Afrique Conseil, 2006)

II.2 Matériel et approche méthodologique

2.1. Matériel

Dans l'optique d'atteindre les différents objectifs définis, nous avons utilisé comme matériel :

Matériels de sources de collecte des données :

- + ECMWF Copernicus Climate Data Store;
- + Port autonome de Cotonou;
- + Climate Data;

Matériels de traitement de données :

- + QGIS pour la cartographie de la zone d'étude ;
- + Langage Python pour le traitement.

Matériels de modélisation des inondations :

- + Jupyter notebook

Matériels pour la conception de l'application web :

- + Flask;
- + Python;
- + HTML & CSS;
- + JavaScript;
- + PyCharm.

2.2. Méthodologie de travail

La procédure méthodologique peut se résumer en quelques étapes telles que :

2.2.1 Recherche documentaire

Cette première partie consiste à faire des recherches sur les travaux antérieurs qui abordaient des sujets similaires ou connexes à ce thème de travail, et tout autre document pouvant fournir des informations relatives à la géographie, l'hydrologie, la climatologie,

l'océanographie, l'apprentissage profond, le modèle ConvLSTM et la sociologie des zones d'études. Nous avons utilisé des mots-clés tels que "inondations côtières", "zone urbaine", "modélisation", "ConvLSTM", etc. pour consulter des bases de données scientifiques comme scispace, Web of Science. L'ensemble de ces informations réunies nous ont permis de faire une synthèse bibliographique des études réalisées sur les inondations côtières en zone urbaine, les méthodes de modélisation existantes et le modèle ConvLSTM.

2.2.2 Collecte des données

Cette partie nous a permis de rassembler toutes les informations pouvant servir à l'accomplissement de notre travail. Elle s'est déroulée comme suit :

- collecte de données spatiales et temporelles du milieu d'étude au niveau de différentes sources, telles que le site de la **ECMWF Copernicus Climate Data Store**, le **Port autonome de Cotonou** et d'autres sites web. Nous avons collecté des données sur la période de **1997 à 2021**, avec une fréquence journalière, concernant les variables suivantes : **précipitation journalière** (mm), **précipitation maximal journalière** (mm) sur un pas d'un an, **marée** (m) et **hauteur de la surface de la mer ou SSH** (m) ;

2.2.3 Prétraitement des données

Cette partie nous a permis de préparer les données collectées avant de les utiliser pour entraîner et tester notre modèle ConvLSTM. Elle s'est déroulée en plusieurs étapes :

- Réorganisation des données collectées : qui consiste à disposer les données dans une organisation pour faciliter le traitement des données
- Concaténation des données spatiales et temporelles en un seul DataFrame, qui contient les variables nécessaires pour modéliser les inondations côtières en zone urbaine : la marée, la SSH, la précipitation et la précipitation maximale ;
- Filtration des données selon la période d'étude : qui consiste à conserver les données uniquement comprises dans la période d'étude et à supprimer les autres ;
- Visualisation des valeurs aberrantes et limitation à un intervalle défini à travers la règle des 1,5 écart interquartile (EIQ) : qui consiste à repérer les données qui s'écartent fortement de la tendance générale de la série statistique dans l'optique d'exclure les valeurs aberrantes en utilisant le concept d'écart interquartile, qui est la différence entre le troisième quartile (Q3) et le premier quartile (Q1) d'une série ordonnée.;

- Vérification des valeurs manquantes et remplacement par une valeur appropriée par la méthode de l'interpolation linéaire : qui consiste à préserver la continuité et la tendance des données ;
- Normalisation des données entre 0 et 1 : qui consiste à effectuer une transformation linéaire aux données numériques à cause de la différenciation au niveau des unités de ces données.

2.2.4 Modélisation et simulation du risque d'inondation avec le modèle ConvLSTM

Cette partie nous a permis de construire, d'entraîner et de tester notre modèle ConvLSTM, qui est un type de réseau de neurones récurrents capable de traiter des séries temporelles spatialisées. Elle s'est déroulée en plusieurs étapes :

- Définition de l'architecture du modèle ConvLSTM, en choisissant le nombre de couches cachées, le nombre de filtres, le type d'activation, la fonction de coût et l'optimiseur ;
- Entraînement du modèle ConvLSTM sur les échantillons d'entraînement, en choisissant le nombre d'époques, la taille du lot et la validation croisée ;
- Evaluation du modèle ConvLSTM sur les échantillons de test, en utilisant des indicateurs de performance tels que : le degré de détermination (d), l'erreur quadratique moyenne (MSE), l'erreur absolue moyenne (MAE), le coefficient de détermination (R2), le coefficient de corrélation (R), l'efficacité de Nash-Sutcliffe (NSE), le logarithme de la vraisemblance marginale (LME)
- Visualisation des résultats du modèle ConvLSTM, en affichant les différents graphiques tels que : la comparaison entre les valeurs prédites et les valeurs réelles de l'IF (indice d'inondation côtière), l'évolution de la fonction de perte (loss) pendant l'entraînement du modèle ConvLSTM, les indicateurs de performance du modèle et enfin la courbe de régression pour les différents horizons.

2.2.5 Interface web

Cette partie nous a permis de mettre une petite application web contenant le modèle ConvLSTM pour utilisateur, construit à partir de Flask, JavaScript, HTML, CSS et Python.

2.3. Déroulement proprement des travaux

- Localisation de la zone d'étude



Image 1: Localisation de la commune de Cotonou avec Google Earth

Pour développer notre modèle ConvLSTM, nous avons utilisé le langage de programmation Python, qui est un langage populaire et puissant pour le traitement des données et l'apprentissage automatique. L'ensemble du code a été implémenté avec Jupyter notebook. Nous avons importé plusieurs bibliothèques Python qui nous ont permis de manipuler, de visualiser, et de modéliser les données. Nous avons subdivisé les déroulements des travaux en 3 volets.

2.3.1. Volet 1 : Chargement des données - Réorganisation – Pré-traitement

Comme paramètre d'entrée de notre modèle, nous avons eu besoin de cinq types de données: la **précipitation journalière**, la **précipitation maximale journalière**, la **marée**, la **SSH**, et l'indice de risque d'inondation côtière (**IF**). Le choix de ces données sont pertinentes pour notre sujet, car elles influencent le phénomène des inondations côtières en zone urbaine. Nous avons obtenu ces données à partir de différentes sources que nous avons lus à l'aide de la bibliothèque **Pandas**. Voici les sources de données que nous avons utilisées :

- Un fichier qui contient les données de la précipitation journalière, en millimètres par jour (mm/j), sur une période de 1996 à 2022 provenant de la plateforme Copernicus ECWMF, sous forme de fichiers NetCDF qui est ensuite convertis en fichier xls.

- Fichier qui contient les données de la précipitation maximale sur un pas d'un an, en millimètres par jour, pour la période de 1996 à 2022, provenant de la plateforme Copernicus ECWMF, sous forme de fichiers NetCDF qui est ensuite converti en fichier xls.
- Un fichier qui contient les données de la marée, en mètres, pour la période de 1990 à 2022, provenant du port autonome de cotonou en format csv qui a été converti en format.xlsx.
- Un fichier qui contient les données de la SSH, en mètres, pour la période de 1997 à 2021, provenant de la plateforme Copernicus ECWMF, sous forme de fichiers NetCDF qui est ensuite convertis en fichier xls.

Nous avons ensuite :

- filtré les données selon la période d'étude, qui est de 1997 à 2021, en utilisant la *fonction loc* de **Pandas**, qui permet de sélectionner des lignes ou des colonnes selon des critères ;
- concaténé les quatre fichiers en un seul dataframe en utilisant la *fonction merge* de **Pandas**, qui permet de joindre des dataframes selon une ou plusieurs colonnes communes ;
- utilisé les colonnes year, month, et day comme clés de jointure ;
- vérifié que les données ne comportaient pas de valeurs manquantes avec la fonction *fillna()* de **panda**, ce qui est un critère important pour la fiabilité des analyses.

Ayant rassemblé l'ensemble de tous les paramètres en une seule dataframe, nous avons procédé au calcul de la précipitation effective moyenne par jour **PE** et de l'indice de risque d'inondation côtière **IF** à partir des données de précipitation, en utilisant la formule suivante :

$$PE = \frac{1}{n} \sum_{i=1}^n P_i$$

Où :

n est le nombre de jours dans l'année (365j)

P_i est la précipitation journalière au jour *i*

Cette formule permet de calculer la moyenne de la précipitation journalière sur une année, en sommant les valeurs de précipitation pour chaque jour, et en divisant par le nombre de jours.

$$IF = \frac{PE - \mu_{Precip_max}}{\sigma_{Precip_max}}$$

Où :

PE est la précipitation effective moyenne par jour

μ_{Precip_max} est la moyenne de la précipitation maximale sur un pas d'un an

σ_{Precip_max} est l'écart – type de la précipitation maximale

Cette formule permet de normaliser la précipitation effective par rapport à la précipitation maximale, et de définir une échelle de risque d'inondation. Après avoir calculé l'IF, nous avons ajouté le tableau IF au dataframe data comme une nouvelle colonne, en utilisant la méthode *insert* de *Pandas*, qui permet d'insérer une colonne dans un dataframe selon un nom et une position. Elle se présente comme suis :

Tableau 1: Dataframe contenant les données d'entrée pour le modèle

	year	month	day	Precip(mm/jour)	Precip_max(mm/jour)	Maree(m)	sea_surface_height_above_geoid(m)	IF
0	1997	1	1	-0.136270	-0.323666	-0.483044	-1.939817	-0.084556
1	1997	1	2	-0.909873	-0.916793	-0.602620	-1.898495	-1.352527
2	1997	1	3	-0.906264	-0.992648	-0.731714	-1.895471	-1.771241
3	1997	1	4	-0.918265	-1.047900	-0.849336	-1.820386	-1.990432
4	1997	1	5	-0.378662	0.053166	-0.929834	-1.764450	-1.768174
...	...	...	...	...	...	...	...	...
9126	2021	12	27	-0.745295	-0.876029	-0.378193	0.301142	0.165205
9127	2021	12	28	-0.917352	-0.891613	-0.552999	0.395880	0.157442
9128	2021	12	29	-0.918265	-1.047900	-0.713666	0.438714	0.149714
9129	2021	12	30	-0.918265	-1.047900	-0.816791	0.489610	0.142028
9130	2021	12	31	-0.918265	-1.047900	-0.822762	0.495657	0.134384

9131 rows × 8 columns

2.3.2. Volet 2 : Construction du modèle, entraînement du modèle

Précédemment, un aperçu du calcul de l'IF est fourni. Dans cette partie, nous avons défini une architecture globale de notre modèle proposé. Comme le montre la figure 2 ci-dessous, les principales informations nécessaires à la construction de notre modèle de prédiction sont : la précipitation journalière, la marée, la SSH, l'indice de risque d'inondation côtière IF. Les données sont utilisées comme source d'entrées au sein de l'algorithme. L'algorithme aura pour objectif de prévoir IF pour les horizons de prévision à 2, 4, 8 et 16 jours, et de fournir également les pertes du modèle à l'horizon de prévisions.

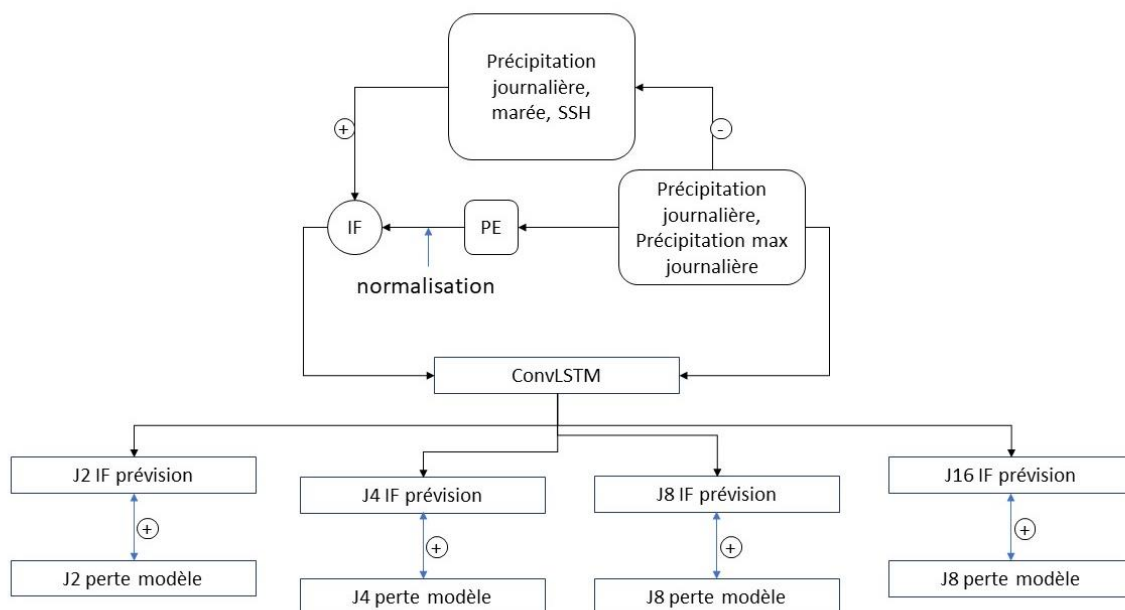


Figure 2: Architecture du modèle

Pour concevoir notre modèle ConvLSTM, nous avons :

- utilisé la bibliothèque **Keras**, qui offre une interface de haut niveau pour **TensorFlow**, qui est un environnement de calcul numérique basé sur des graphes, qui permet de créer, d'entraîner, et d'évaluer des modèles d'apprentissage automatique, notamment des réseaux de neurones profonds ;
- choisi d'utiliser un modèle **ConvLSTM**, qui est un type de réseau de neurones récurrents, qui combine les propriétés des couches convolutionnelles **CNN** et des couches **LSTM** (Long Short-Term Memory). Ce type de modèle est adapté pour traiter des données séquentielles et spatiales, comme celles des inondations côtières, car il peut capturer les dépendances temporelles et spatiales entre les variables d'entrée et de sortie.

Après avoir, importé les différentes bibliothèques, nous avons dans un premier temps divisé les données en deux ensembles : un ensemble d'entraînement (1), qui contient 80% des données, et un ensemble de test (2), qui contient 20% des données. Nous avons utilisé l'ensemble d'entraînement pour ajuster les paramètres du modèle, et l'ensemble de test pour évaluer les performances du modèle. Nous avons normalisé les données entre 0 et 1, en utilisant la formule suivante :

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

Où :

x est la valeur originale

x_{min} est la valeur minimale

x_{max} est la valeur maximale

x_{norm} est la valeur normalisée

Cette formule nous a permis de mettre les données à la même échelle, et de faciliter l'apprentissage du modèle.

Nous avons :

- créé des séquences d'entrée et de sortie pour le modèle, en utilisant différentes tailles de fenêtres temporelles, qui sont de 2, 4, 8, et 16 jours. Une fenêtre temporelle est un intervalle de temps qui contient les données d'entrée et de sortie pour le modèle. Par exemple, si la taille de la fenêtre temporelle est de 2 jours, le modèle prendra les données des 2 jours précédents comme entrée, et prédira la donnée du jour suivant comme sortie ;
- utilisé différentes tailles de fenêtres temporelles pour comparer l'effet de la longueur de la séquence sur les performances du modèle. Nous avons parcouru les données d'entraînement et de test, et nous avons créé des séquences d'entrée et de sortie selon la taille de la fenêtre temporelle, en utilisant des boucles **for**, qui permettent de répéter une instruction selon un nombre de fois ou une condition. Nous avons converti les séquences en tableaux NumPy, en utilisant la fonction **array** de **NumPy**, qui permet de créer un tableau à partir d'une liste ou d'un autre objet itérable.

Ensuite, nous sommes passés à la structure du modèle ConvLSTM en utilisant la classe **Sequential** de **Keras**, qui permet de créer un modèle en empilant des couches.

Nous avons utilisé :

- deux couches ConvLSTM2D pour une meilleure complexité du modèle, qui sont des couches qui appliquent une convolution 2D à l'entrée et à l'état caché, et qui utilisent une cellule LSTM pour maintenir une mémoire à long terme. Nous avons réduit la taille du noyau et utilisé un mode de remplissage différent dans les couches ConvLSTM2D, pour éviter le surapprentissage et améliorer la généralisation du modèle ;

- la *fonction d'activation relu*, qui est une fonction non linéaire qui renvoie la valeur d'entrée si elle est positive, et zéro dans le cas contraire ;
- une *couche Flatten*, qui est une couche qui transforme un tenseur multidimensionnel en un vecteur unidimensionnel, pour faciliter la connexion avec la couche de sortie. Nous avons ajouté une *couche Dense*, qui est une couche qui effectue une multiplication matricielle entre l'entrée et les poids, suivie d'une addition avec le biais, pour produire la sortie ;
- la *fonction d'activation tanh*, qui est une fonction non linéaire qui renvoie une valeur entre -1 et 1, pour normaliser la sortie du modèle.

Nous avons ensuite compilé le modèle, en utilisant la *fonction compile de Keras*, qui permet de configurer le modèle pour l'entraînement.

- l'*optimiseur adam*, qui est un algorithme d'optimisation adaptatif, qui ajuste le taux d'apprentissage en fonction de la pente du gradient.
- la *fonction de perte huber*, qui est une fonction qui mesure l'erreur entre la sortie prédite et la sortie réelle, en utilisant une fonction quadratique pour les petites erreurs, et une fonction linéaire pour les grandes erreurs. Cette fonction permet de réduire l'impact des valeurs aberrantes sur l'apprentissage du modèle.

Une fois cette étape achevée, nous avons positionné le *plot model de Keras* qui nous montre le diagramme à chaque taille de la fenêtre. Nous pouvons visualiser l'ensemble de ces diagrammes qui nous montrent comment le processus se déroule du début de l'activation de la séquence jusqu'au niveau de la compilation ci-dessous :

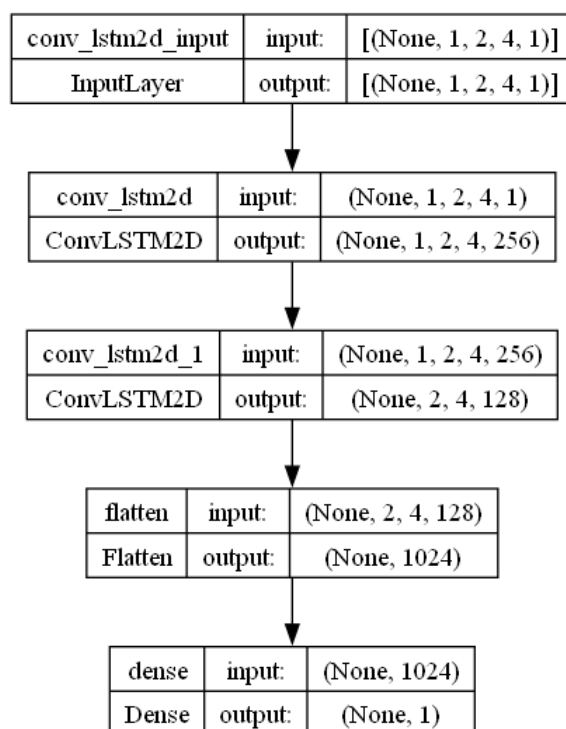


Figure 3: Processus du déroulement de la séquentialisation (taille de fenêtre : 2)

Ce processus est décrit comme suis :

- La première étape est la couche d'entrée, qui reçoit une entrée multidimensionnelle de forme (None, 1, 2, 4, 1). Cela signifie que le modèle peut traiter un nombre indéfini de séquences (None), que chaque séquence contient une seule image (1), que chaque image a deux canaux (2) qui désigne ici l'horizon de prévision, que chaque canal a quatre lignes (4) qui désigne le nombre de paramètres et que chaque ligne a une colonne (1) ;
- La deuxième étape est la première couche ConvLSTM2D, qui applique une convolution bidimensionnelle à chaque image de la séquence, en utilisant des filtres de taille (1, 2) et un nombre de filtres égal à 256. Cette couche transforme les données et produit une sortie de forme (None, 1, 2, 4, 256). Cela signifie que le modèle conserve le même nombre de séquences (None) et le même nombre d'images par séquence (1), mais que chaque image a maintenant 256 canaux (256), et que chaque canal a deux lignes (2) et quatre colonnes (4) ;
- La troisième étape est la deuxième couche ConvLSTM2D, qui applique une convolution bidimensionnelle à chaque image de la séquence, en utilisant des filtres de taille (3, 3) et un nombre de filtres égal à 128. Cette couche réduit la dimensionnalité des données et produit une sortie de forme (None, 2, 4, 128). Cela signifie que le modèle conserve

le même nombre de séquences (None), mais que chaque séquence contient maintenant deux images (2), que chaque image a 128 canaux (128), et que chaque canal a quatre lignes (4) et quatre colonnes (4) ;

- La quatrième étape est la couche Flatten, qui aplatit les données multidimensionnelles en un vecteur unidimensionnel. Cette couche produit une sortie de forme (None, 1024). Cela signifie que le modèle conserve le même nombre de séquences (None), mais que chaque séquence est représentée par un vecteur de 1024 éléments (1024) ;
- La cinquième et dernière étape est la couche Dense, qui prend le vecteur de chaque séquence et produit une sortie unique. Cette couche utilise une fonction d'activation sigmoïde, qui renvoie une valeur entre 0 et 1. Cette valeur peut être interprétée comme la probabilité que la séquence appartienne à une certaine classe. Cette couche produit une sortie de forme (None, 1). Cela signifie que le modèle conserve le même nombre de séquences (None), mais que chaque séquence a une seule sortie (1) qui dans notre cas désigne l'IF prédit.

Ce processus est répété suivant chaque horizon. Après cette étape, nous avons utilisé :

- la fonction **fit de Keras**, qui permet d'ajuster les paramètres du modèle en minimisant la fonction de perte, en utilisant l'optimiseur choisi.
- la fonction **evaluate de Keras**, qui permet de calculer la valeur de la fonction de perte sur les données de test, en utilisant le modèle entraîné.
- la fonction **predict de Keras**, qui permet de générer des prédictions pour les données de test, en utilisant le modèle entraîné. Vu l'énorme temps que va prendre la prédiction pour éviter des surapprentissage, d'arrêter l'entraînement du modèle lorsque la valeur de la fonction de perte sur les données de validation ne diminue plus, et des temps de calcul important, nous avons positionné le **early stopping** qui jouera ce rôle. Puis s'en est suivi l'utilisation de la classe **ReduceLROnPlateau de Keras**, qui permet d'implémenter la réduction du taux d'apprentissage, en spécifiant le critère de réduction, comme la fonction de perte sur les données de validation, et le facteur de réduction, qui est le rapport entre le nouveau taux d'apprentissage et l'ancien taux d'apprentissage.

Après cette étape, s'en est effectué la **validation croisée**, qui est une technique qui permet de diviser les données en plusieurs sous-ensembles, et d'utiliser chaque sous-ensemble comme données de validation, pour évaluer la performance du modèle de manière plus robuste et plus fiable. Nous avons utilisé la fonction **train_test_split de Scikit-learn**, qui permet de diviser les

données en deux sous-ensembles, selon un ratio donné. Puis nous avons ensuite inversé la normalisation des données, en utilisant la formule inverse :

$$x = x_{norm} \times (x_{max} - x_{min}) + x_{min}$$

Où :

x est la valeur de la variable x après l'inversion

x_{min} est la valeur minimale

x_{max} est la valeur maximale

x_{norm} est la valeur normalisée

Nous avons ensuite comparé les valeurs prédites avec les valeurs réelles pour différents horizons de prédiction, en utilisant la **bibliothèque Matplotlib** , Nous avons obtenu :

- un graphique qui montre les valeurs réelles et les valeurs prédites de l'IF, en fonction du temps, pour chaque taille de fenêtre temporelle. Nous avons défini des seuils, selon ces seuils, il affiche un message, tel que : “Risque nul”, “Risque faible”, “Risque modéré”, “Risque très élevé” ;
- un graphique qui montre l'évolution de la fonction de perte sur les données d'entraînement et de validation, en fonction du nombre d'époques, pour chaque taille de fenêtre temporelle. A ce niveau, nous avons créé et utilisé la fonction **interpreter_graphique** pour l'analyse de nos courbes de perte d'entraînement et de validation de notre modèle. Il permet de comparer les valeurs finales et les tendances des deux courbes et renvoie un commentaire qui résume la performance du modèle. En fonction de chaque cas, nous avons défini des commentaires.
- un diagramme de dispersion des résidus à partir de la fonction **sns.residplot** sur chaque horizon accompagné d'une fonction qui analyse le graphique et affiche des commentaires en fonction de chaque seuil atteint.

Pour évaluer notre modèle, il est important de calculer les indicateurs de performance, qui sont des mesures qui permettent d'évaluer la qualité des prédictions du modèle. L'ensemble de ces indicateurs de performance sont également représentés sous forme de graphique qui compare les performances du modèle pour différentes tailles de fenêtres temporelles (voir partie résultat du Chapitre III). Nous pouvons résumer dans ce tableau, chacune des formules.

Tableau 2: Récapitulatif indicateurs de performance et formules

Indicateurs de performance	Formules
Le <u>coefficient de détermination (R^2)</u> , qui est une mesure qui indique la proportion de la variance de la variable dépendante qui est expliquée par la variable indépendante. Plus le R^2 est proche de 1, plus le modèle est performant.	$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}$ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite $\bar{y}$ est la moyenne des valeurs réelles n est le nombre d'observations.
L' <u>erreur quadratique moyenne (MSE)</u> , qui est une mesure qui indique la différence moyenne au carré entre les valeurs réelles et les valeurs prédites. Plus le MSE est faible, plus le modèle est performant.	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite n est le nombre d'observations.
L' <u>erreur absolue moyenne (MAE)</u> , qui est une mesure qui indique la différence moyenne absolue entre les valeurs réelles et les valeurs prédites. Plus le MAE est faible, plus le modèle est performant.	$MSE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite n est le nombre d'observations.
Le <u>coefficient de Nash-Sutcliffe (NSE)</u> , qui est une mesure qui indique le degré de correspondance entre les valeurs réelles et les valeurs prédites, par rapport à la moyenne des valeurs réelles. Plus le NSE est proche de 1, plus le modèle est performant.	$NSE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite $\bar{y}$ est la moyenne des valeurs réelles n est le nombre d'observations.
Le <u>coefficient de Legates et McCabe (LME)</u> , qui est une mesure qui indique le degré de correspondance entre les valeurs réelles et les valeurs prédites, par rapport à la somme des	$LME = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y} + \hat{y}_i - \bar{y})^2}$ Où : y_i est la valeur réelle

écarts absolus entre les valeurs prédites et la moyenne des valeurs réelles, et entre les valeurs réelles et la moyenne des valeurs réelles. Plus le LME est proche de 1, plus le modèle est performant.	$\hat{y}_i$ est la valeur prédite $\bar{y}$ est la moyenne des valeurs réelles n est le nombre d'observations.
Le <u>coefficient de Willmott (d)</u> , qui est une mesure qui indique le degré de correspondance entre les valeurs réelles et les valeurs prédites, par rapport à la somme des écarts absolus entre les valeurs prédites et les valeurs réelles, et entre les valeurs réelles et la moyenne des valeurs réelles. Plus le d est proche de 1, plus le modèle est performant.	$d = 1 - \frac{\sum_{i=1}^n y_i - \hat{y}_i }{\sum_{i=1}^n (y_i - \bar{y} + \hat{y}_i - \bar{y})}$ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite $\bar{y}$ est la moyenne des valeurs réelles n est le nombre d'observations.
L' <u>erreur quadratique moyenne racinée (RMSE)</u> , qui est une mesure qui indique la racine carrée de la différence moyenne au carré entre les valeurs réelles et les valeurs prédites. Plus le RMSE est faible, plus le modèle est performant.	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite n est le nombre d'observations.
Le <u>coefficient de corrélation de Pearson (r)</u> , qui est une mesure qui indique le degré de relation linéaire entre les valeurs réelles et les valeurs prédites. Plus le r est proche de 1, plus le modèle est performant.	$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}$ Où : y_i est la valeur réelle $\hat{y}_i$ est la valeur prédite $\bar{y}$ est la moyenne des valeurs réelles $\bar{\hat{y}}$ est la moyenne des valeurs prédites n est le nombre d'observations.

L'ensemble des graphiques sont présents dans la partie Résultats du chapitre III.

2.3.3. Volet 3 : Mise en place d'une petite interface

Nous avons ensuite créé une petite application web pour contenir notre modèle, à partir de PyCharm en utilisant les outils et les langages suivants :

- **Flask**, qui est un framework web léger et flexible, écrit en Python, qui nous a permis de créer notre application web dynamique et interactive, en utilisant des routes, des templates, et des formulaires.
- **Python**, comme nous l'avons défini plus haut, un langage de programmation avec lequel nous avons fait le code de notre modèle
- **HTML**, qui est un langage de balisage qui nous a permis de structurer et de présenter le contenu de notre page web.
- **JavaScript**, qui est un langage de programmation interprété, orienté objet, et événementiel, qui nous a permis de créer des effets dynamiques et interactifs sur notre page web, le survol des boutons, des effets et bien d'autres éléments.

L'objectif de créer une application web pour notre modèle est de permettre aux utilisateurs potentiels de tester et d'utiliser notre modèle de manière simple et intuitive, sans avoir besoin de connaître le code ou les détails techniques du modèle. L'application web offre les fonctionnalités et les éléments suivants :

- Une page de connexion où l'utilisateur doit d'abord se connecter, en utilisant un email et un mot de passe. Nous avons :
 - simulé une base de données avec un dictionnaire Python, qui contient les informations des utilisateurs autorisés à titre expérimental ;
 - utilisé la *fonction request de Flask*, qui permet de récupérer les données envoyées par l'utilisateur, et la *fonction render_template de Flask*, qui permet de renvoyer une page HTML en utilisant un template Jinja2.

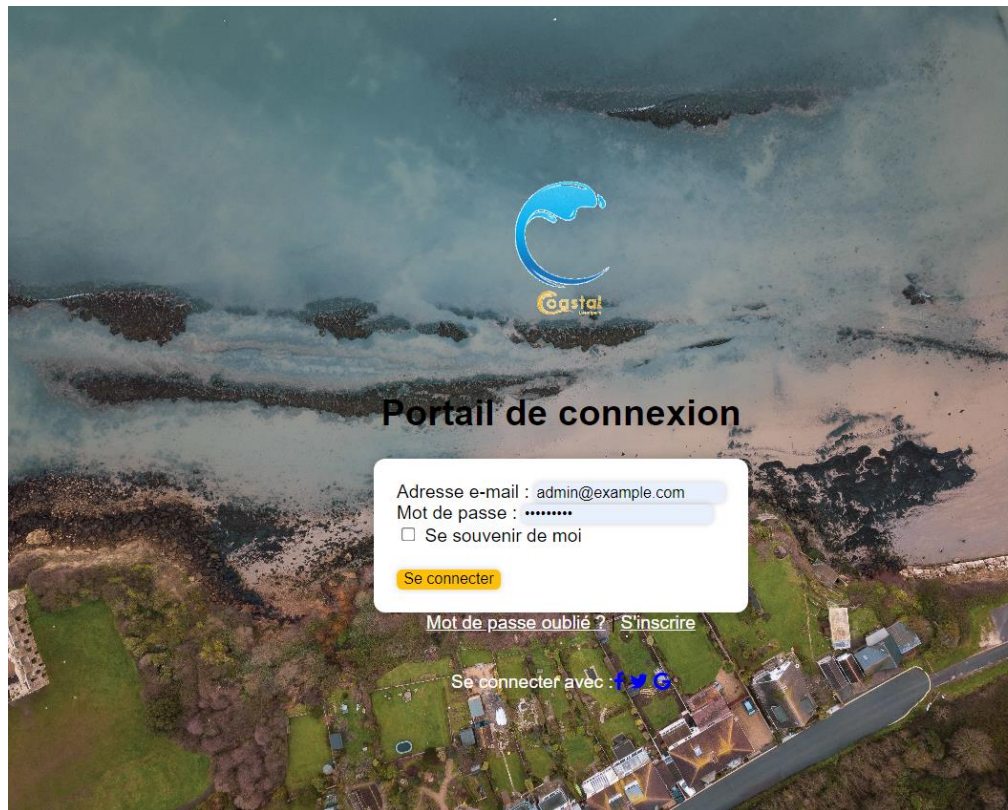


Image 2: Page de connexion de l'application

- Cette fonction **request** de Flask, permet de récupérer les données envoyées par l'utilisateur, en utilisant la méthode POST ;
 - utilisé la fonction **redirect** de Flask, qui permet de rediriger l'utilisateur vers la page d'accueil, en spécifiant l'URL ;
 - utilisé la fonction **url_for** de Flask, qui permet de générer l'URL d'une route, en spécifiant le nom de la fonction qui la définit.
- Une fois connecté, une page d'accueil s'affiche qui contient une barre de navigation avec des boutons qui mènent à différents pages. La route défini pour l'obtention de la page d'accueil se présente comme suis :

```
@app.route('/')  
def home():  
    return render_template('Page accueil.html')
```

La page d'accueil de l'application se présente comme suit :

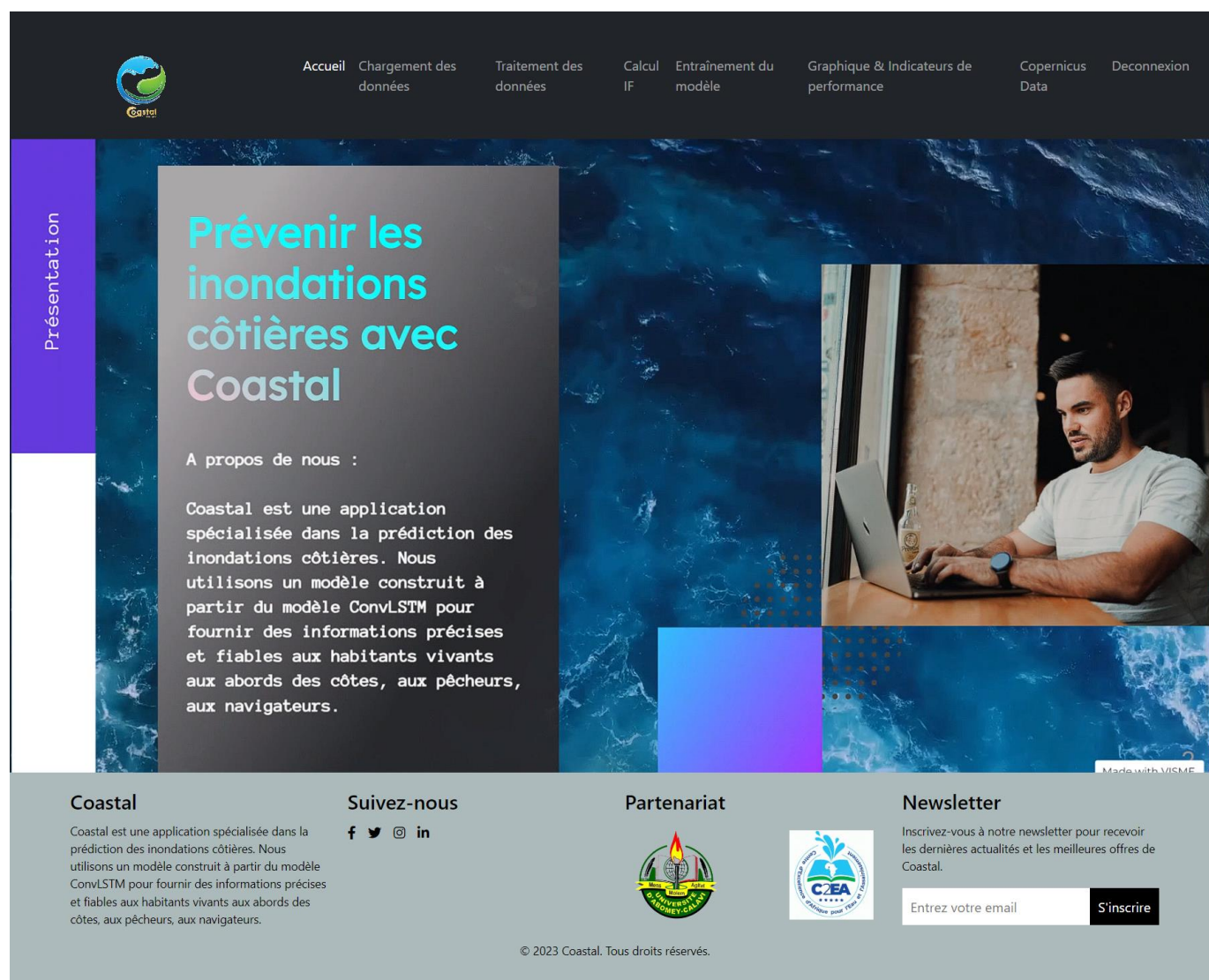


Image 3: Page d'accueil de l'application

- Au niveau de la page Chargement des données : l'utilisateur peut entrer le nombre de paramètres à utiliser pour le modèle, qui après validation lui ouvre une page lui donnant l'opportunité de sélectionner les fichiers Excel. Dans notre cas à titre d'exemple sont : les données de précipitation journalière, de la SSH, de marée. Nous avons utilisé la fonction request de Flask, et la fonction render template de Flask, qui ont permis de renvoyer la page HTML utilisant la route pour la page d'affichage des fichiers :

```
@app.route("/affichage", methods=["POST"])
```

avec la fonction **pd.read_excel** de **Pandas**, qui permet de lire un fichier Excel Pandas

Accueil

PARAMETRE

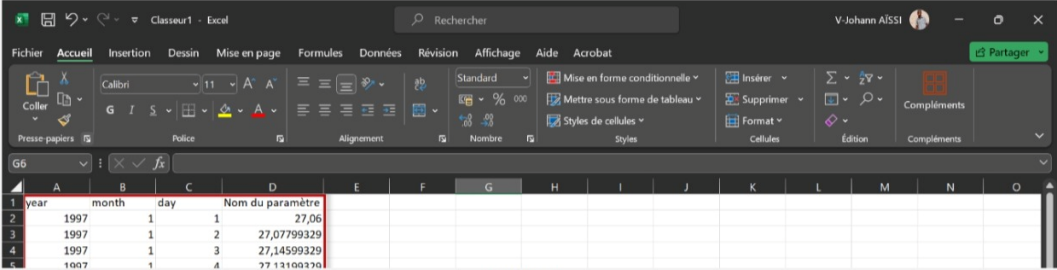
Saisissez le nombre de paramètres que vous voulez prendre en compte.

Nombre de paramètres :

Image 4: Initialisation du nombre de paramètres

Sélectionner les fichiers

Veuillez sélectionner les fichiers au format Excel ou CSV correspondant aux paramètres requis.
[NB]--> Assurez-vous que vos données respectent ce format !!<--
Suivez l'ordre suivant : Précipitation, Précipitation maximale, suivi des autres paramètres.



Fichier 1 :

Aucun fichier n'a été sélectionné

Fichier 2 :

Aucun fichier n'a été sélectionné

Image 5: Sélection des fichiers en fonction du nombre de paramètres fixés par l'utilisateur

Données Combinées

-> Regroupement des paramètres dans le tableau. <-

	year	month	day	Precip(mm/jour)
0	1996	10	1	6.471908
1	1996	10	2	14.765753
2	1996	10	3	23.521238
3	1996	10	4	7.805632
4	1996	10	5	0.527641
5	1996	10	6	10.422091
6	1996	10	7	27.376017
7	1996	10	8	0.342938

Image 6: Affichage de la liste des données

- Au niveau de la page Données combinées : l'utilisateur peut visualiser les données de chaque fichier Excel, sous la forme d'une liste de tableau, avant de pouvoir passer à la page suivante. Une fois sur la page suivante 'Traitement de données', il est appelé à entrer la durée sur laquelle il souhaite faire sa prévision. Nous avons utilisé :

- la route pour la page d'affichage avec pour action 'POST' des données tel que :

```
@app.route("/affichage1_traitement_donnee", methods=["POST"])
```

en prenant le soin de stocker les variables avec la *fonction global*.

- la *fonction functools.reduce* pour la concaténation des datas en une seule et une fonction d'appel *filter_data* pour le filtrage, crée au départ de notre code, pour éviter des répétitions de code et éviter d'alourdir notre code.

A la sortie de cette fonction, nous obtenons une DataFrame sur la page suivante à partir de la fonction *render_template()* qui renferme tous les paramètres ensemble.

Accueil

Traitement des données

Saisissez l'année de début et l'année de fin pour filtrer les données selon la période d'étude.

Année de début :

Année de fin :

Filtrer

Image 7: Traitement des données

Données

Combinaison des paramètres dans le tableau

	year	month	day	Precip(mm/jour)	Precip_max(mm/jour)	Maree(m)	sea_surface_height_above_geoid(m)
0	1997	1	1	7.345926	7.345926	0.913504	0.304733
1	1997	1	2	0.078831	1.329821	0.905447	0.307467
2	1997	1	3	0.112728	0.560421	0.896749	0.307667
3	1997	1	4	0.000000	0.000000	0.888825	0.312633
4	1997	1	5	5.068938	11.168144	0.883401	0.316333
5	1997	1	6	7.521014	9.778662	0.882701	0.321067
6	1997	1	7	9.559408	13.863021	0.888645	0.321700
7	1997	1	8	12.513014	19.248175	0.901006	0.325367
8	1997	1	9	1.685849	7.724859	0.916347	0.330500
9	1997	1	10	18.763140	18.904619	0.929159	0.333267
10	1997	1	11	27.740759	27.740759	0.934438	0.334867

Image 8: Affichage de l'ensemble des données combinées

- Une fois les données concaténées en une seule, l'utilisateur se rend sur la page suivante à partir du bouton 'suivant' situé en bas de la dataframe. Ce qui le mènera sur la page

de 'Calcul IF' avec la fonction ***render_template()***. A ce niveau nous avons créé une route que voici :

```
@app.route("/calculer", methods=["POST"])  
def calculer():  
    ...
```

Avec l'utilisation de quelques fonctions telles que ***numpy*** et d'autres fonctions, l'utilisateur aura à spécifier l'indice/position de la colonne contenant les données de la précipitation et celle des précipitations maximales. Mais également il aura à spécifier le nombre de paramètres suivis d'un moins ; puis il appuie sur le bouton 'Calculer' pour que cela lui affiche directement les résultats sur la même page du 'Calcul IF'. Comme le montre l'image ci-dessous :

Accueil Suivant

Calcul IF

Indiquez l'indice de la colonne contenant les données de précipitation quotidienne :

Indiquez l'indice de la colonne contenant les données de précipitation maximale :

Entrez le nombre de paramètres, en tenant compte de l'indice d'inondation côtière (IF) (ex : -X) :

Calculer

Image 9: Page du Calcul de l'IF

Accueil

Suivant

Calcul IF

Indiquez l'indice de la colonne contenant les données de précipitation quotidienne :

Indiquez l'indice de la colonne contenant les données de précipitation maximale :

Entrez le nombre de paramètres, en tenant compte de l'indice d'inondation côtière (IF) (ex : -X) :

Calculer

Résultat

	year	month	day	Precip(mm/jour)	Precip_max(mm/jour)	Maree(m)	sea_surface_height_above_geoid(m)	IF
0	1997	1	1	-0.136270	-0.323666	-0.483044	-1.939817	-0.084556
1	1997	1	2	-0.909873	-0.916793	-0.602620	-1.898495	-1.352527
2	1997	1	3	-0.906264	-0.992648	-0.731714	-1.895471	-1.771241

Image 10: Résultat du calcul de l'IF sur l'application

- Une fois que le calcul est effectué, l'utilisateur passe à la page suivante à partir de la fonctionne **render_template()** pour l'entraînement du modèle à partir de la route suivante :

```
@app.route("/predict", methods=["POST"])
def predict():
    ...
```

A titre d'exemple, nous avons inséré nos hyperparamètres pour notre étude mais au niveau du nombre d'époques (Epoch) nous avons testé sur 3 à titre expérimental mais nous tenons à rappeler que dans notre étude nous avons pris exactement comme valeur 100. Nous rappelons qu'au fur et à mesure de l'exécution du modèle les graphiques s'afficheront successivement mais l'utilisateur doit les fermer pour que l'exécution puisse continuer. Comme nous le montre les photos ci-dessous :

Accueil Suivant

Prédiction avec ConvLSTM

Entrez vos choix de configuration et obtenez une prédiction.

Tailles des fenêtres temporelles (en jours)

2

4

8

16

Nombre de filtres de la première couche ConvLSTM2D

256

Nombre de filtres de la deuxième couche ConvLSTM2D

128

Fonction d'activation de la première couche ConvLSTM2D

relu

Fonction d'activation de la deuxième couche ConvLSTM2D

relu

Fonction d'activation de la couche Dense

tanh

Fonction d'optimisation

adam

Epoche

3

Prédire

Image 11: Prédiction avec le modèle ConvLSTM sur l'application

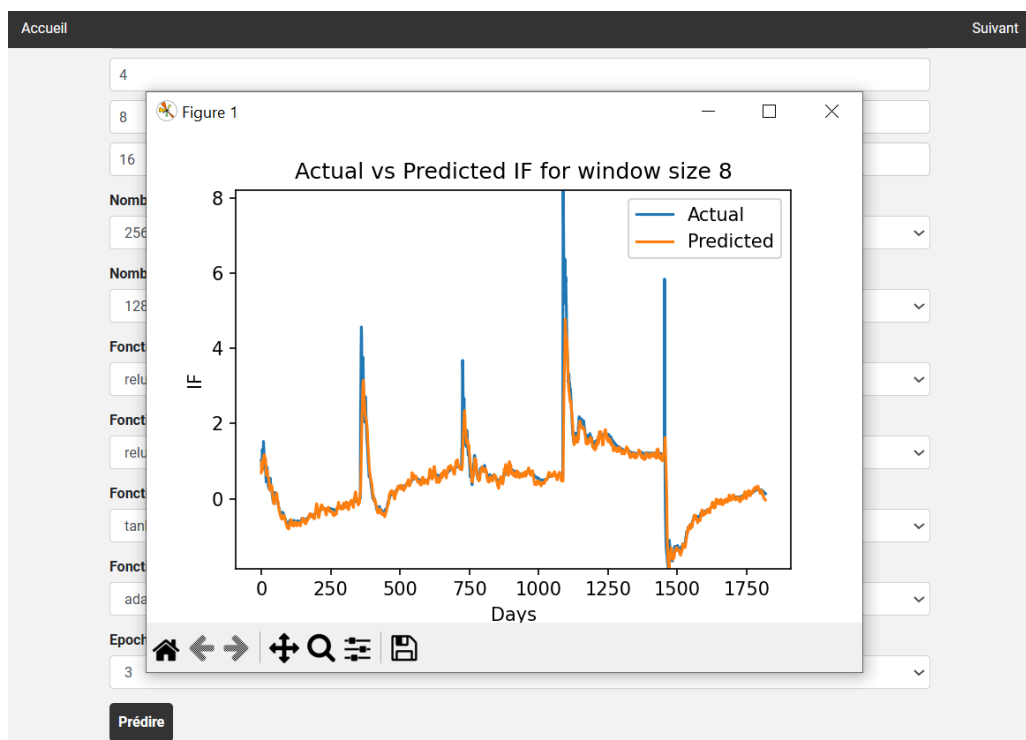


Image 12: Pendant le processus de prédiction

Accueil

Test loss: 0.0002684971841517836

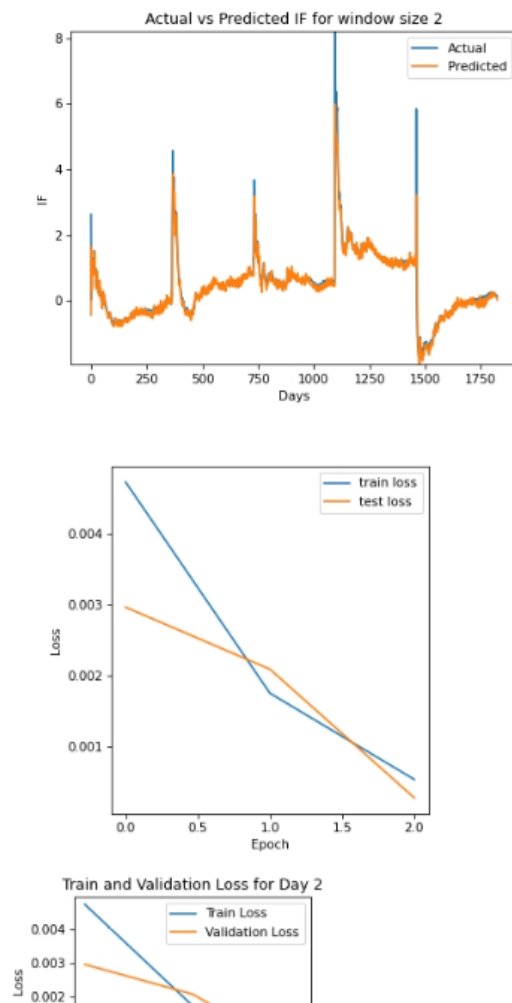


Image 13: Affichage de l'ensemble des graphiques et des indicateurs de performance

Nous avons prévu une fonctionnalité à notre application en cours de développement qui permettra à l'utilisateur de pouvoir télécharger des données de Copernicus. Il s'agit de la page Copernicus Data. Nous avons intégré des services du **Copernicus Climate Change Service (C3S)** pour des données de type climatique telles que la précipitation, l'humidité du sol et d'autres éléments. Nous avons aussi intégré les services du **Copernicus Marine Environment Monitoring Service (CMEMS)** pour des données marines et environnementales telles que : les vagues, le vent et d'autres paramètres. L'utilisateur aura la possibilité de pouvoir saisir les paramètres de son choix ne figurant pas dans la liste proposée, choisir la date de début et de fin, définir la zone géographique. L'ensemble des téléchargements seront effectués à partir d'une API que nous avons insérée dans l'application. Nous avons défini une route telle que :

```
@app.route('/coper_data', methods=['POST'])
def coper_data():
    ...
```

Quelques photos qui montrent la page Copernicus Data :

Accueil

Formulaire de téléchargement de données Copernicus

Veuillez choisir les données que vous souhaitez obtenir à partir du service Copernicus Climate Change Service (C3S) ou du service Copernicus Marine Environment Monitoring Service (CMEMS).

Service :
Copernicus Marine Environment Monitoring Service (CMEMS)

Couche :
Vagues océaniques

Autre :
Saisir le nom d'une couche qui n'est pas dans la liste

Date de début:
jj/mm/aaaa

Date de fin:
jj/mm/aaaa

Zone géographique :
ex: -180,-90,180,90

Format de fichier:
CSV

Envoyer

Image 14: Page Copernicus Data sur l'application

Accueil

Formulaire de téléchargement de données Copernicus

Veuillez choisir les données que vous souhaitez obtenir à partir du service Copernicus Climate Change Service (C3S) ou du service Copernicus Marine Environment Monitoring Service (CMEMS).

Service :
Copernicus Marine Environment Monitoring Service (CMEMS)

Couche :
Vagues océaniques

Copernicus Climate Change Service (C3S)
Copernicus Marine Environment Monitoring Service (CMEMS)
Température de l'océan
Salinité de l'océan
Niveau de la mer
Courants océaniques
Vagues océaniques
Vent océanique

jj/mm/aaaa

Date de fin:

Image 15: Quelques paramètres figurant dans les services CMEMS

Notre application devait être hébergée sur un serveur, afin de la rendre accessible à un large public et de faciliter son utilisation. Cependant, en raison du temps imparti pour la durée des travaux de recherche, nous n'avons pas pu l'intégrer.

Chapitre III

RESULTATS ET DISCUSSIONS

CHAPITRE III : RESULTATS ET DISCUSSIONS

Ce chapitre a pour objectif de présenter et de discuter les résultats obtenus par le modèle ConvLSTM que nous avons développé pour la prévision des inondations côtières de la zone d'étude. Nous présentons d'abord la courbe comparative de l'IF prédit et de l'IF réel de notre zone d'étude obtenu à partir du modèle sur les différents horizons. Ensuite, nous montrons ensuite les courbes de perte d'entraînement et de précision du modèle, ainsi que les valeurs obtenues des indicateurs de performance du modèle pour différents scénarios. Après avoir comparé les résultats du modèle avec ceux d'autres méthodes existantes, nous finirons par une discussion sur les résultats obtenus par le modèle ConvLSTM, que nous avons développé pour la prévision des inondations côtières en zone urbaine et côtière.

III.1 Résultats

3.1.1. Prévision de l'indice d'inondation côtière par le modèle ConvLSTM pour différents jours de prévision

Les figures 4-a, b, c, d ci-dessous présentent la courbe comparative de l'IF prédit et de l'IF réel de notre zone d'étude obtenu à partir de notre modèle construit à partir du ConvLSTM, en fonction du temps, pour les quatre jours de prévision (2j, 4j, 8j, 16j).

L'obtention des différents résultats obtenus, ont été possible à partir des données journalières sur la période respective de 1997 à 2021. Nous avons utilisé quatre paramètres d'entrée : la précipitation, l'IF, la marée, et la SSH. Nous avons simulé quatre scénarios (4 jours de prévision), en fonction des valeurs de ces paramètres, et nous avons comparé l'IF prédit avec l'IF réel.

Les figures 4-a, b, c, d se présentent comme suit :

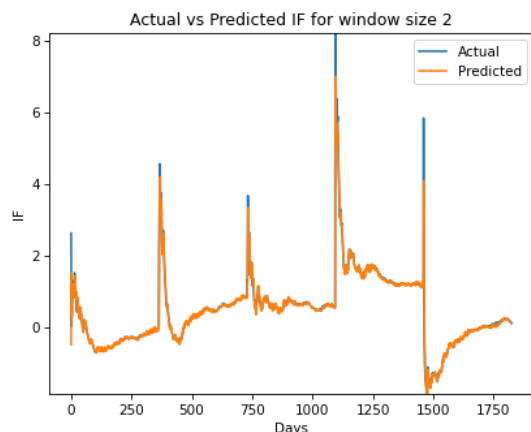


Figure 4- a: Pr vision IF pr dit Jour 2

Message obtenu : Le niveau de risque d'inondation c ti re pour le jour suivant est : Risque faible

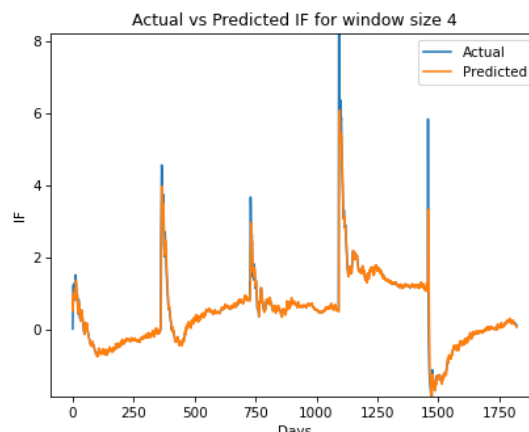


Figure 4- b: Pr vision IF pr dit Jour 4

Message obtenu : Le niveau de risque d'inondation c ti re pour le jour suivant est : Risque faible

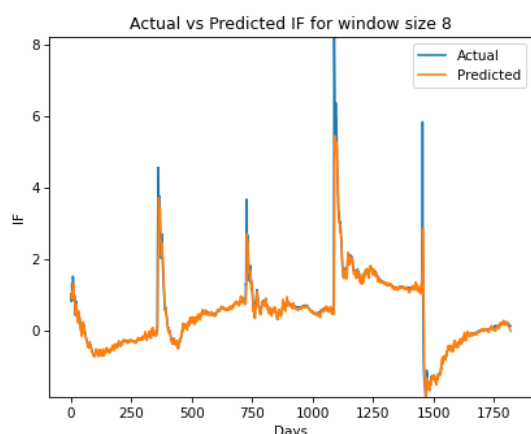


Figure 4- c: Pr vision IF pr dit Jour 8

Message obtenu : Le niveau de risque d'inondation c ti re pour le jour suivant est : Risque nul

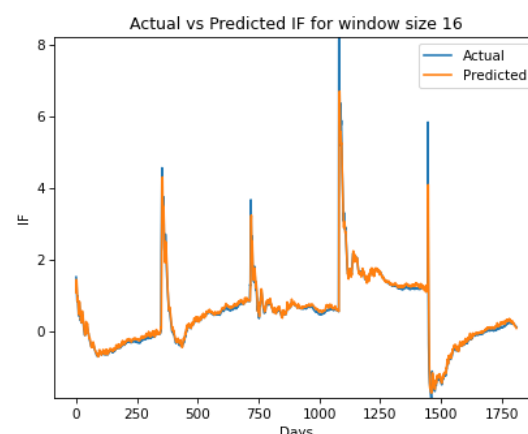


Figure 4- d: Pr vision IF pr dit Jour 16

Message obtenu : Le niveau de risque d'inondation c ti re pour le jour suivant est : Risque faible

On observe que le mod le construit   partir du mod le ConvLSTM reproduit fid lement l' volution de l'IF r el, avec une bonne pr cision et une bonne stabilit  durant toute la p riode. La courbe de l'IF pr dit suit de pr s la courbe de l'IF r el, sans pr senter de d calage ou de distorsion. Nous constatons que sur les 4 horizons de pr visions, une croissance  volutive des 2 courbes entre 0 et 115 jours, ce qui montre que le mod le anticipe les p riodes de fortes et de

faibles inondations en fonctions de nos paramètres d'entrées : la précipitation, l'IF, la marée, la SSH.

Nous observons également que les 4 scénarios présentent cinq pics, dont quatre se situent dans la période où il y a de grandes inondations, entre 438 et 1440 jours qui correspondent respectivement aux dates suivantes : [14-03-1998 ; 10-12-2000], et le dernier se situe dans la période où il n'y a pas d'inondation, après 1440 jours qui correspondent à la date du [04-12-2003 ; 16-10-2021]. Le plus grand pic qui se trouve au jour 1130 correspond à la date du [04-02-2000], où l'IF prédit, atteint la valeur de 6, ce qui indique un risque d'inondation très élevé. Le dernier pic se trouve au jour 1441 qui correspond à la date du [05-12-2003], où l'IF prédit, descend à la valeur de -1, ce qui indique un risque d'inondation nul. Nous pouvons traduire ces informations comme suit :

Tableau 3: Seuils de risques d'inondation

IF prédit	Risque d'inondation
< 0	Nul
0 – 0,5	Faible
0,5 – 1,5	Modéré
1,5 – 2,5	Élevé
> 2,5	Très élevé

Tableau 4: Message obtenu par horizon de prévision

	Risque d'inondation
Jour 2	Risque faible
Jour 4	Risque faible
Jour 8	Risque nul
Jour 16	Risque faible

En se basant sur ces analyses, nous pouvons déduire que notre modèle construit à partir du modèle ConvLSTM est capable de capturer les caractéristiques dynamiques et complexes des inondations côtières, en utilisant les données séquentielles de nos paramètres d'entrée. Il est également en mesure de détecter les changements brusques et les anomalies dans l'IF réel, et de les refléter dans l'IF prédit tout en restant sensible aux variations des paramètres d'entrée.

Nous pouvons également déduire que lorsque l'IF prédit est négatif, le risque d'inondation est nul, dans le cas contraire il y a risque d'inondation.

3.1.2. Perte d'entraînement du modèle ConvLSTM

Les figures 5-a, b, c, d ci-dessous présentent la courbe de la perte d'entraînement du modèle ConvLSTM, en fonction du nombre d'époques, pour les quatre jours de prévision (2j, 4j, 8j, 16j).

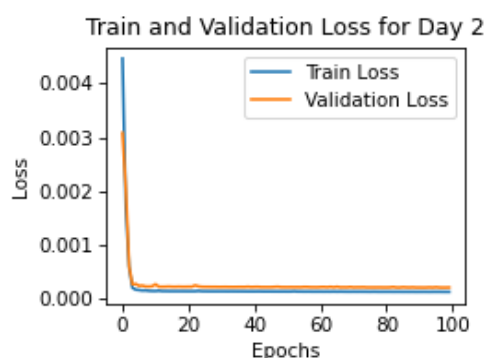


Figure 5- a : Perte au J2

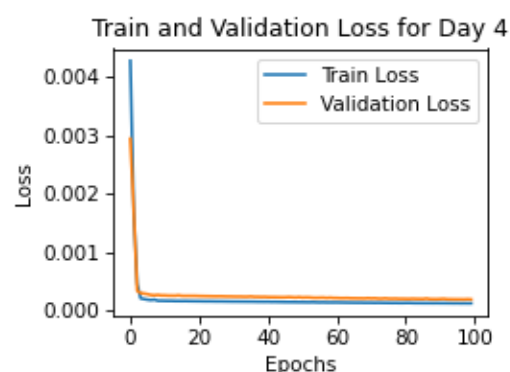


Figure 5- b: Perte au J4

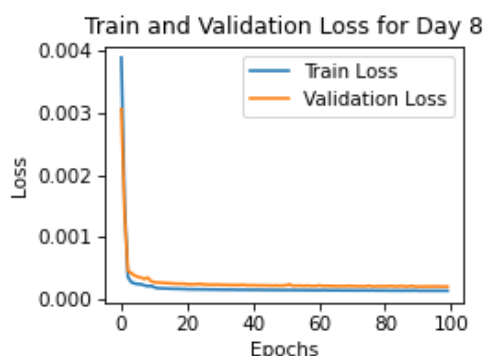


Figure 5- c: Perte au J8

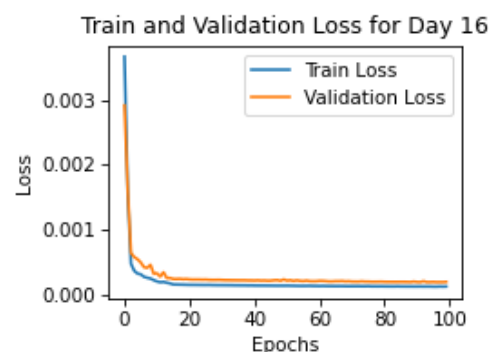


Figure 5- d: Perte au J16

On constate que la perte moyenne d'entraînement est de 0,0025, et la perte moyenne de validation est de 0,0027, ce qui indique une bonne adaptation du modèle aux données. Le modèle atteint sa perte minimale d'entraînement et de validation en moins de 3 époques, ce qui indique une bonne vitesse d'apprentissage.

Nous constatons également sur les 4 horizons le modèle est sur-ajusté entre 0 et 3 époques, mais au fil des époques s'ajuste bien ce qui justifie la trajectoire constante de la perte d'entraînement et de la validation.

En se basant sur ces analyses, nous pouvons déduire que notre modèle indique une bonne vitesse d'apprentissage et n'est plus soumis au surapprentissage qu'à partir de 3 époques, car la courbe de la perte de validation suit la courbe de la perte d'entraînement, sans s'écarter et demeure constante.

3.1.3. Analyse du diagramme de dispersion des résidus

Les figures 6-a, b, c, d suivant nous présentent le diagramme de dispersion des résidus.

Elle nous montre la tendance des résidus en fonction des valeurs prédites. Les résidus sont la différence entre les valeurs réelles et prédites qui dans notre cas sont : IF réel et IF prédit. Nous constatons au niveau des différents scénarios a, b, c, d, que la majorité des résidus sont dispersés autour de la ligne horizontale à zéro. Cela signifie que le modèle de régression est bien ajusté par rapport aux données. On peut déduire que le modèle est précis, mais également linéaire. L'ensemble des figures se présente comme suit :

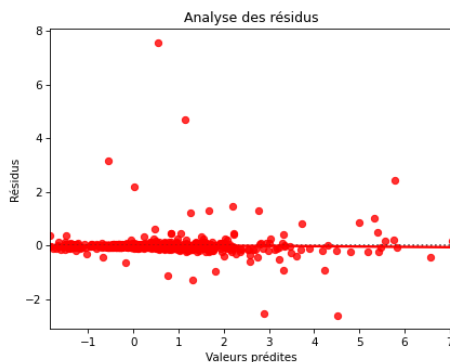


Figure 6- a : Dispersion des résidus au J2

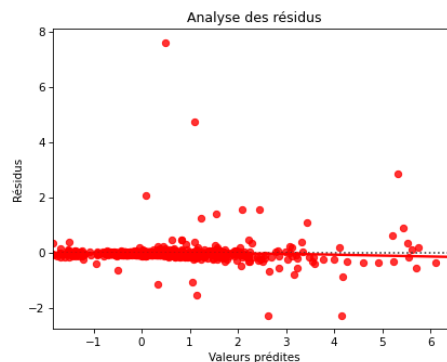


Figure 6- b: Dispersion des résidus au J4

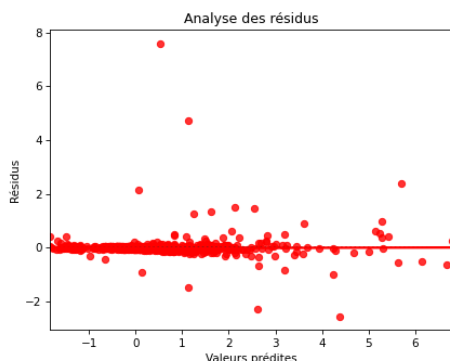


Figure 6- c: Dispersion des résidus au J8

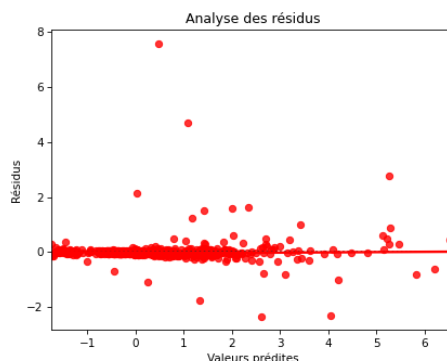


Figure 6- d: Dispersion des résidus au J16

3.1.4. Indicateurs de performance du modèle ConvLSTM, en fonction du nombre de jours prédits

La figure 7 ci-dessous présente les valeurs des indicateurs de performance du modèle ConvLSTM, en fonction du nombre de jours prédits, pour les différents paramètres d'entrée.

L'ensemble de ces données obtenus avec notre modèle peuvent être traduit dans le tableau ci-dessous :

Tableau 5: Récapitulatif des données des indicateurs de performances par horizon de prévision

Jour de prévision	d	NSE	LME	R2	MSE	MAE	RMSE	r
1	0,9	0,92	0,98	0,92	0,08	0,07	0,28	0,96
2	0,91	0,93	0,98	0,93	0,07	0,06	0,27	0,96
8	0,91	0,93	0,98	0,93	0,07	0,06	0,27	0,96
16	0,9	0,93	0,98	0,93	0,07	0,06	0,27	0,96

En utilisant les normes des indicateurs de performance liés au modèle en régression, nous pouvons interpréter les résultats obtenus pour les quatre jours de prévision. Les valeurs de NSE pour les quatre jours de prévision sont supérieures à 0,75, ce qui indique que le modèle est capable de prédire les valeurs observées avec une précision satisfaisante. Les valeurs de R2 pour les quatre jours de prévision sont supérieures à 0,75, ce qui indique que le modèle est capable d'expliquer la variance des données observées. Les valeurs de MSE et RMSE pour les quatre jours de prévision sont inférieures à 0,1, ce qui indique que le modèle a une faible erreur de prédiction. Les valeurs de MAE pour les quatre jours de prévision sont inférieures à 0,1, ce qui indique que le modèle a une faible erreur absolue de prédiction. Enfin, les valeurs de r pour les quatre jours de prévision sont supérieures à 0,9, ce qui indique une forte corrélation entre les valeurs prédites et les valeurs observées.

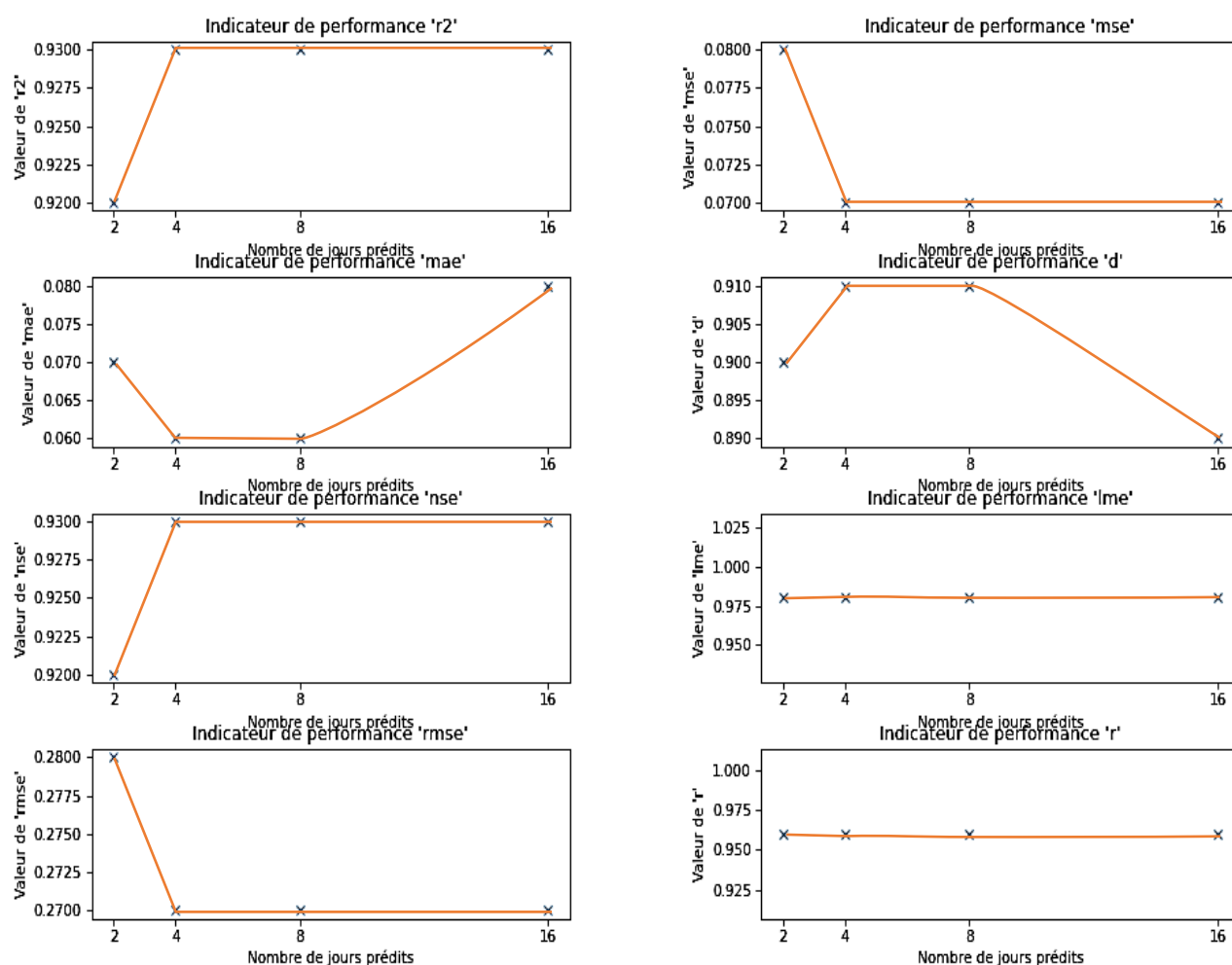


Figure 4: Indicateurs de performance pour différents horizons

Les résultats obtenus pour les indicateurs de performance liés au modèle en régression sont satisfaisants pour les quatre jours de prévision. Cela indique que le modèle est capable de prédire jusqu'au 16^{ème} jour.

En se basant sur les différentes déductions faites au niveau du 3.1.1., 3.1.2, 3.1.3, 3.1.4, nous pouvons dire que plus le modèle est capable de prévoir l'IF avec une faible perte, plus il est fiable pour anticiper le risque d'inondation. Et sur la base des indicateurs de performance liés au modèle en régression, qui indiquent que le modèle est capable de prédire jusqu'au 16^{ème} jour, nous pouvons conclure que notre modèle est un outil efficace pour la prévision et la gestion des inondations côtières, en utilisant les données séquentielles des paramètres d'entrée : la précipitation, l'IF, la marée, et la SSH.

III.2 Discussions

Dans cette partie, nous allons discuter des résultats obtenus par le modèle ConvLSTM, que nous avons développé pour la prévision des inondations côtières en zone urbaine et côtière. Le modèle ConvLSTM est un modèle d'apprentissage profond, qui combine la convolution et le LSTM, pour traiter les données séquentielles des paramètres d'entrée. Dans notre cas il s'agit de : la précipitation, la marée, l'IF, la SSH. L'objectif principal de l'étude est de construire un modèle de prévision des inondations côtières en zone urbaine à l'aide de ConvLSTM : étude de la région d'Akpakpa à Cotonou. La mise en place de ce modèle inclut le calcul de l'indice d'inondation côtière (IF) en plus des autres paramètres pour l'obtention de l'IF prédire suivant différents jours de prévision. Les résultats obtenus par le modèle ConvLSTM sont présentés sous forme de trois graphiques, qui montrent respectivement la courbe de l'IF prédit, la courbe de la perte du modèle, et les valeurs des indicateurs de performance. Nous allons discuter de ces résultats, en les comparant avec les travaux existants.

- Les résultats obtenus pour la prévision des inondations à l'aide de notre modèle, concordent avec ceux obtenus par Mohammed Moishin et al. en 2021, dans son étude portant sur “ Conception d'un modèle de prévision d'inondation basé sur l'apprentissage profond avec l'algorithme hybride ConvLSTM “
 - Notre étude se distingue de la leur par plusieurs aspects :
 - Le contexte géographique : nous avons étudié les inondations côtières en zone urbaine, tandis qu'ils ont étudié les inondations fluviales en zone rurale.
 - La durée des données utilisées : nous avons utilisé 24 ans de données historiques, tandis qu'ils ont utilisé 30 ans de données historiques.
 - Le type d'inondation : nous avons pris en compte les effets combinés de la précipitation, de l'IF, la marée, et la SSH, tandis qu'ils ont considéré uniquement la précipitation.
 - Le nombre et la nature des paramètres d'entrée : nous avons utilisé quatre paramètres d'entrée, dont deux sont liés à l'océan, tandis qu'ils ont utilisé deux paramètres d'entrée, tous deux liés à la pluie.
 - L'horizon de prévision : nous avons prévu le risque d'inondation jusqu'à 16 jours à l'avance, tandis qu'ils ont prévu le risque d'inondation jusqu'à 10 jours à l'avance.

- Nos résultats sont similaires à ceux de l'étude de Moishin et al. (2021), qui ont également obtenu des indicateurs de performance élevés pour la prévision du Jp, avec des valeurs supérieures à 0,89 pour le NSE, à 0,93 pour le R2, et à 0,96 pour le r, et des valeurs inférieures à 0,08 pour le MSE, à 0,28 pour le RMSE, et à 0,18 pour le MAE. Ils ont également utilisé des graphiques pour comparer le Jp prédit et le Jp réel, et pour montrer la perte du modèle ConvLSTM. Ces graphiques montrent que leur modèle suit également de près l'évolution du Jp réel, mais que la différence entre les valeurs prévues et réelles de Jp augmente avec l'horizon de prévision, et qu'elles atteignent une perte minimale en moins de 5 époques.
- Nous observons également quelques points de distinction en ce qui concerne les paramètres optimaux pour notre modèle développé et ceux présentés dans l'étude de Moishin et al. (2021) qui se présente comme suit :

Tableau 6: Différences observées des paramètres optimaux

Modèles de prévision	Nom des hyperparamètres	Hyperparamètre optimal
Modèle de prévision développé par Moishin et al. (2021)	Filter I	128
	Activation	ReLU
	Optimizer	Adam
	Batch Size	100
	Epochs	50
Modèle de prévision développé par notre étude	Filter I	256
	Filter II	128
	Activation I	ReLU
	Activation II	ReLU
	Activation III	Tanh
	Optimizer	Adam
	Batch Size	100
	Epochs	100
	Fonction de Perte	Huber

Cependant, notre étude présente des avantages et des inconvénients par rapport à la leur. D'une part, notre étude utilise plus de paramètres d'entrée que la leur, ce qui permet de prendre en compte plus de facteurs influençant le risque d'inondation

côtière, et d'améliorer la robustesse du modèle. D'autre part, notre étude utilise un horizon de prévision plus long que la leur, ce qui implique une plus grande incertitude et une plus grande sensibilité aux erreurs de prédiction.

L'étude de Moishin et al. (2021) présente également des avantages et des inconvénients par rapport à la nôtre. D'une part, leur étude utilise un modèle hybride qui combine les réseaux de neurones récurrents et convolutionnelles pour traiter des données séquentielles et spatiales. Ce modèle permet de capturer les caractéristiques dynamiques et complexes des phénomènes hydrologiques, en utilisant les données historiques et actuelles des paramètres d'entrée. Leur étude apporte donc une contribution innovante au domaine de la prévision des inondations, en proposant un modèle qui exploite les avantages des deux types de réseaux de neurones. D'autre part, leur étude utilise moins de paramètres d'entrée que la nôtre, ce qui limite la prise en compte des facteurs influençant le risque d'inondation. Leur étude utilise également un horizon de prévision plus court que le nôtre, ce qui réduit la portée et l'utilité de leur modèle pour la gestion des risques d'inondation. Leur étude pourrait donc bénéficier d'une extension du nombre et de la nature des paramètres d'entrée, ainsi que d'une extension de l'horizon de prévision.

- Les résultats obtenus pour la prévision des inondations à l'aide de notre modèle, concordent avec ceux obtenus par Tano et al. en (2023) dans son étude portant sur "Caractérisation des événements d'inondation côtière le long de la Côte d'Ivoire (Afrique de l'Ouest)"
 - Notre étude porte sur la région d'Akpakpa à Cotonou, au Bénin, tandis que celle de Tano et al. (2023) porte sur le littoral ivoirien, en Afrique de l'Ouest. Il s'agit donc de deux contextes géographiques différents, qui peuvent influencer les facteurs et les conséquences des inondations côtières. Nous avons utilisé quatre paramètres d'entrée pour prédire l'indice d'inondation (IF) : la précipitation, l'IF, la marée, et la SSH. Nous avons utilisé des données journalières sur la période 1997-2021, et nous avons simulé quatre scénarios (4 jours de prévision), en fonction des valeurs de ces paramètres. Par rapport à Tano et al. (2023), nous avons pris en compte la précipitation comme facteur d'inondation, ce qui peut améliorer la précision de la prédiction. En revanche, ils ont utilisé des données

de marée, de houle, de surcote et d'élévation du littoral sur la période 1993-2019, ce qui couvre une plus longue période et plus de facteurs physiques que nous.

- Nous avons évalué la qualité de notre modèle avec des indicateurs de performance, comme le NSE, le R2, le MSE, le MAE, le RMSE et le r. Ces indicateurs nous ont permis de mesurer la fiabilité de la prédiction et la variance expliquée par le modèle. En revanche, Tano et al. (2023) n'ont pas mentionné les indicateurs de performance qu'ils ont utilisés pour évaluer la qualité de leur méthode, ce qui rend difficile la comparaison des résultats.
- La principale différence entre les deux méthodes est que le modèle ConvLSTM est un modèle prédictif, qui utilise des données séquentielles et spatiales, tandis que le CFI est un indice descriptif, qui utilise des données ponctuelles et statiques. Le modèle ConvLSTM peut donc capturer les dynamiques temporelles et spatiales des inondations côtières, et simuler différents scénarios en fonction des valeurs des paramètres d'entrée. Le CFI, quant à lui, peut cartographier les zones inondées et les durées d'immersion, et déterminer les niveaux de risque d'inondation pour chaque sous-section du littoral ivoirien. Les deux méthodes présentent donc des avantages et des inconvénients, selon les objectifs et les données disponibles. Les deux méthodes présentent aussi quelques similitudes, comme le fait d'utiliser des données récentes, de couvrir une large zone d'étude, et de proposer une échelle de risque d'inondation.

Cependant, notre étude présente des avantages et des inconvénients par rapport à la leur. Elle se présente à travers le tableau ci-dessous :

Tableau 7: Avantages et inconvénients entre les méthodes utilisées

Avantages	Inconvénients
Méthode : ConvLSTM	
- Méthode avancée et adaptée pour traiter les données séquentielles et spatiales - Fiabilité de la prédiction grâce aux indicateurs de performance	- Complexité du modèle - Risques de surapprentissage ou de sous-apprentissage si les hyperparamètres ne sont pas bien optimisés

- Possibilité d'appliquer le modèle à d'autres zones d'étude et de simuler différents scénarios - Proposition de mesures de prévention ou d'adaptation face aux inondations côtières en zone urbaine	
Méthode : CFI	
- Méthode simple qui utilise un indice facile à calculer et à comprendre - Cartographie des zones inondées et des durées d'immersion - Usage de quatre paramètres d'entrée qui couvrent tous les facteurs physiques qui influencent les inondations côtières	- Faible qualité de la modélisation, qui utilise un indice descriptif, des données ponctuelles et statiques - Incapacité de capturer les caractéristiques dynamiques et complexes des inondations côtières, de détecter les changements brusques et les anomalies dans l'IF réel, de refléter les variations des paramètres d'entrée, et de prédire les valeurs futures de l'IF - Faible fiabilité de la prédiction, car pas d'indicateurs de performance mentionnés - Impossibilité d'utiliser la méthode pour simuler différents scénarios et proposer des mesures de prévention, de mitigation ou d'adaptation face aux inondations côtières en zone urbaine

- Les résultats obtenus pour la prévision des inondations à l'aide de notre modèle, concordent avec ceux obtenus par BLALOGOE en 2018, dans son étude portant sur "Dynamique spatio-temporelle des inondations de 1980 à 2010 dans le grand Cotonou"
 - Notre étude complète les résultats de l'article de BLALOGOE (2018) qui a analysé la dynamique spatio-temporelle des inondations de 1980 à 2010 dans Cotonou et les facteurs explicatifs. Notre étude couvre une période plus récente et plus court, et se focalise sur une zone plus restreinte et plus exposée aux inondations. Nous constatons également des paramètres différents de ceux de

l'article, qui sont : la pluviométrie, la densité de population, le taux d'urbanisation et le taux d'occupation du sol. Nous nous sommes basés sur un indicateur d'inondation plus précis et plus fiable, qui est l'IF, alors que l'article de BLALOGOE (2018) se base sur le nombre de ménages affectés par les inondations. En ce qui concerne la méthodologie, nous avons opté pour une modélisation plus avancée et plus performante, qui est le ConvLSTM, alors que l'article recourt à une méthode de cartographie plus classique et plus descriptive.

- Notre étude confirme les résultats de BLALOGOE (2018) qui a montré que les inondations sont plus importantes dans les agglomérations de Cotonou et de Sèmè-Podji, et qu'elles se manifestent surtout pendant les périodes de fortes pluies. En effet, dans nos résultats obtenus plus haut, nous avons observé que les quatre scénarios de prévision présentent cinq pics d'IF prédit, dont quatre se situent dans la période où il y a de grandes inondations, entre 438 et 1440 jours qui correspondent respectivement aux dates suivantes : [14-03-1998 ; 10-12-2000]. Ces dates coïncident avec les années où l'article a relevé les plus fortes précipitations et les plus grandes surfaces inondées dans Cotonou. Le plus grand pic d'IF prédit se trouve au jour 1130 qui correspond à la date du [04-02-2000], où l'IF prédit atteint la valeur de 6.7, ce qui indique un risque d'inondation très élevé. Cette date correspond également à l'une des plus grandes crues du fleuve Ouémé, qui a provoqué des inondations dans les quartiers riverains du lac Nokoué, comme Akpakpa. Le dernier pic d'IF prédit se trouve au jour 1441 qui correspond à la date du [05-12-2003], où l'IF prédit descend à la valeur de -1, ce qui indique un risque d'inondation nul. Cette date correspond à une période de faibles pluies et de faibles inondations dans le Grand Cotonou, selon l'article. Ainsi, nous pouvons affirmer que nos dates de prévision sont en accord avec les données historiques et les observations de terrain de l'article, et qu'elles reflètent la réalité des inondations côtières dans la région d'Akpakpa.
- Notre étude présente toutefois certaines limites. D'une part, nous n'avons pas pu évaluer le risque d'inondation d'autres quartiers, ni évaluer les impacts socio-économiques et environnementaux des inondations. Ces aspects mériteraient d'être approfondis dans de futures recherches. Par ailleurs, notre étude soulève la question de la gestion des risques d'inondation dans la région d'Akpakpa. Il est nécessaire de mettre en place des mesures de prévention, de protection et de

résilience face aux inondations, en impliquant les acteurs locaux et en tenant compte des spécificités du territoire.

CONCLUSION

La région d'Akpakpa à Cotonou, au Bénin, est une zone urbaine exposée aux inondations côtières, qui affectent les habitants et les activités socio-économiques. Pour modéliser et évaluer ces inondations, nous avons développé un modèle en s'appuyant sur le ConvLSTM, qui prédit l'indice d'inondation côtière (IF) en fonction de quatre paramètres d'entrée : la précipitation, l'IF, la marée, et la hauteur de la surface de l'eau. Nous avons simulé quatre scénarios (4 jours de prévision) sur la période 1997-2021, qui montrent que le risque d'inondation dépend de la valeur de l'IF prédit, de la perte du modèle et des paramètres d'entrée. Nous avons évalué la qualité de notre modèle avec des indicateurs de performance, comme le NSE, le R2, le MSE, le MAE, le RMSE et le r. Avec des valeurs du NSE et du R2 supérieures à 0,75, des valeurs du MSE, du RMSE et du MAE inférieures à 0,1 et une valeur du r supérieure à 0,95. Cela indique que le modèle est capable de prédire jusqu'au 16ème jour. Nous avons intégré notre modèle dans une petite application web développée avec Flask et Python. Cependant, notre modèle peut encore être amélioré et enrichi, en tenant compte des perspectives suivantes : optimisation des hyperparamètres du modèle ConvLSTM, en utilisant des méthodes plus efficaces et plus rapides, comme la recherche en grille ou la recherche aléatoire ; tester le modèle sur d'autres zones d'étude, en adaptant les paramètres d'entrée et les données spatiales selon les caractéristiques locales ; intégration d'autres paramètres d'entrée, comme la houle, la surcote ou l'élévation du littoral, qui peuvent influencer les inondations côtières ; amélioration de l'interface web afin qu'elle soit plus conviviale et plus interactive, permettant aux utilisateurs de visualiser les résultats du modèle, de simuler différents scénarios et de recevoir des alertes en cas de risque d'inondation.

REFERENCES BIBLIOGRAPHIQUES

- Afrique Conseil. (2006). Monographie de la commune de Cotonou (p. 8-13).
- Anoumou Rene Tano et al. (2023). Characterization of coastal flooding events along d'Ivoire (West Africa) Cote. Journal of Coastal Research, 1-8.
- Bourlard, H., & Kamp, Y. (1988). Auto-association by multilayer perceptrons and singular value decomposition. Biological cybernetics, 59(4-5), 291-294.
<https://www.semanticscholar.org/paper/Auto-association-by-multilayer-perceptrons-and-Bourlard-Kamp/f5821548720901c89b3b7481f7500d7cd64e99bd>
- Cocou Parfait BLALOGOE. (2018). Dynamique spatio-temporelle des inondations de 1980 à 2010 dans le grand Cotonou (p. 1, 55-60).
- Claude Touzet. Les réseaux de neurones artificiels, introduction au connexionnisme (p. 59-60, 66). EC2, 1992, Collection de l'EERIE, N. Giambiasi
- Dateandtime. (2021-2023). Les coordonnées géographiques de Cotonou, Bénin.
<https://dateandtime.info/fr/citycoordinates.php?id=2394819>
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning (Vol. 1, No. 10). Springer series in statistics.
- KOSTYA KRAVCHENKO. (2019). ConvLSTM : Tutoriel sur le réseau LSTM convolutif.
<https://www.kaggle.com/code/kcostya/convlstm-convolutional-lstm-network-tutorial>.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
- Moishin et al. (2021). Conception of a flood forecasting model based on deep learning with the hybrid algorithm ConvLSTM (p. 4, 6-44).
- QuinnRadich, Roudabeh et v-alje. (2023). Entraîner votre modèle d'analyse de données avec PyTorch. <https://learn.microsoft.com/fr-fr/windows/ai/windows-ml/tutorials/pytorch-analysis-train-model>
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k., & Woo, W.-c. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Dans : Advances in neural information processing systems (p. 802-810).
- Simon K. (2023). NUMÉRIQUE 2023 : BÉNIN. Magazine Datareportal.
<https://datareportal.com/reports/digital-2023-benin>
- Subramanian, S., Juarez, S. et al. (2023). Apprenez les bases des réseaux de neurones.
- Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. A. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. Journal of machine learning research, 11(Dec), 3371-3408.
<http://www.jmlr.org/papers/volume11/vincent10a/vincent10a.pdf>

Yakaranda Magazine. (2021). Qu'est-ce que la pression barométrique ?
<https://yakarandamag.com/fr/quest-ce-que-la-pression-barometrique/>.

TABLE DES MATIERES

FICHE DE CERTIFICATION	ii
DEDICACE.....	iii
REMERCIEMENTS.....	iv
RESUME.....	vi
ABSTRACT	vii
SOMMAIRE	viii
LISTE DES TABLEAUX	ix
LISTE DES FIGURES.....	x
LISTE DES IMAGES	xi
LISTE DES SIGLES ET ABREVIATIONS.....	xii
INTRODUCTION.....	1
CHAPITRE I : REVUE DE LITTERATURE	4
I.1 Clarification de certains concepts	4
I.2 Méthodes de modélisations existantes : Etat des connaissances.....	6
CHAPITRE II : DEMARCHE METHODOLOGIQUE	10
II.1 Cadre d'étude.....	10
2.1 Localisation de la zone d'étude : COMMUNE DE COTONOU	10
2.2 Cadre administratif et physique.....	11
2.2.1 Cadre administratif	14
2.2.2 Milieu physique.....	14
II.2 Matériel et approche méthodologique.....	16
2.1. Matériel	16
2.2. Méthodologie de travail	16

2.2.1 Recherche documentaire	16
2.2.2 Collecte des données	17
2.2.3 Prétraitement des données	17
2.2.4 Modélisation et simulation du risque d'inondation avec le modèle ConvLSTM ...	18
2.2.5 Interface web	18
2.3. Déroulement proprement des travaux	19
2.3.1. Volet 1 : Chargement des données - Réorganisation – Pré-traitement	19
2.3.2. Volet 2 : Construction du modèle, entraînement du modèle.....	21
2.3.3. Volet 3 : Mise en place d'une petite interface.....	30
CHAPITRE III : RESULTATS ET DISCUSSIONS	43
III.1 Résultats.....	43
3.1.1. Prévion de l'indice d'inondation côtière par le modèle ConvLSTM pour différents jours de prévision.....	43
3.1.2. Perte d'entraînement du modèle ConvLSTM	46
3.1.3. Analyse du diagramme de dispersion des résidus.....	47
3.1.4. Indicateurs de performance du modèle ConvLSTM, en fonction du nombre de jours prédits	48
III.2 Discussions.....	50
CONCLUSION.....	57
REFERENCES BIBLIOGRAPHIQUES.....	58
TABLE DES MATIERES	60